

UNIVERSITÉ DE STRASBOURG

École doctorale des sciences de la vie et de la santé

THÈSE

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE STRASBOURG

Discipline : Biophysique

par

Julien GODET

Étude du rôle de chaperon de la protéine
NCp7 de la nucléocapside du virus VIH-1

Soutenue publiquement le 19 octobre 2010

Membres du jury :

Pr. Frédéric DARDEL
Dr. Gilles MIRAMBEAU
Dr. Olivier PIOT
Dr. Daniel RIVELINE
Dr. Guy DUPORTAIL
Pr. Yves MÉLY

Rapporteur Externe, Paris
Rapporteur Externe, Barcelone
Examineur Externe, Reims
Rapporteur Interne, Strasbourg
Directeur de thèse
Directeur de thèse

UMR CNRS 7213, Faculté de Pharmacie, Illkirch

Demain, dès l'aube, à l'heure où blanchit la campagne,

Je partirai. (...)

Victor HUGO, *Contemplations*, 1847

Remerciements

Il faut quelquefois avoir un peu de chance . . .

C'est très certainement ce qui m'est arrivé quand, il y a un peu plus de cinq ans et alors que je venais de passer le concours de l'internat, j'ai décidé de pousser la porte du laboratoire de Yves. Débarqué un peu par hasard, juste avec l'espoir de me trouver, après un court passage ici, un boulot qui puisse m'amuser un peu.

Je n'aurais pas pu imaginer un environnement plus favorable pour effectuer ma thèse. J'ai eu la chance de travailler sur une des thématiques historiques du laboratoire. Disposer de l'héritage et du savoir faire rend nécessairement les choses plus facile. Dans cet environnement de travail exigeant, j'ai toujours bénéficié d'une situation protégée et d'une très grande liberté dans mon travail. J'ai eu l'occasion d'imaginer, de concevoir, d'essayer beaucoup et de rater aussi, en tout cas, de me faire la main. J'ai été fléché sur des travaux très différents menés dans le cadre de collaborations, ce qui m'a permis de mener de front de nombreux sujets de recherche. J'ai également eu la chance de découvrir l'enseignement, la rédaction des dossiers d'appels à projets, la mise en place de collaborations, d'entrevoir l'envers du décor et les rouages plus subtils des décisions "politiques"...

Je tiens bien évidemment à remercier mes deux directeurs de thèses.

De Guy, j'aurais appris qu'il est possible d'espérer mener une carrière complète sans jamais céder à la lassitude. Merci mil fois d'avoir accepté d'endosser le rôle de directeur de thèse.

De Yves, j'aurais appris la volonté insatiable de toujours aspirer à plus. C'est étonnant que ce genre d'énergie ait eu prise sur moi. C'est pourtant ce qui m'aura permis de repousser parfois mes que trop familières aisance et facilité au profit d'approches plus laborieuses et plus approfondies, et de constater la satisfaction qui peut en émaner. Merci en tout cas de m'avoir chaperonné au cours de ces années et d'avoir su peser sur les décisions qui ont

rendu mon parcours plus facile. Merci également de m'avoir permis de mener en parallèle de ma thèse mes obligations vis-à-vis de l'hôpital, deux DIU, un "second" master et depuis un an déjà mon mi-temps de praticien hospitalier. Toutes ces obligations ont amputé de manière évidente tout une partie de ce qui aurait pu être fait au cours de ma thèse.

Je tiens à remercier tous ceux dont je ne citerai pas le nom de peur d'en oublier, et avec qui j'ai eu l'occasion de travailler ou tout simplement de croiser au labo.

Je remercie également les plus rares qu'il m'a été donné de rencontrer également en dehors du labo, et dont certains sont devenus des amis.

De manière bien plus pragmatique, je tiens à remercier les "sponsors" pour leur soutien financier : TRIoH, le CHU de Nancy et le CHRU de Strasbourg, l'UDS ainsi que l'ANRS et l'ANR.

Etre resté cinq ans m'aura permis de constater combien ce labo a évolué. Chaque année a apporté sa nouveauté. C'est certainement cette dynamique qui a pesé dans mon choix de prolonger de quelques années encore cette aventure.

Aujourd'hui que ma thèse se termine, les années les plus faciles sont très certainement derrière moi. Il reste nécessairement beaucoup à faire. Je n'ai aucune idée de ce que sera la suite ni combien de temps cette "carrière" à venir m'accaparera si ce n'est qu'elle sera, quoi qu'il arrive, conditionnée par la possibilité de toujours trouver de quoi m'amuser un peu.

Il me faudra donc, certainement encore, avoir un peu de chance ...

Une vie sans ornements –
que le jour et la nuit multipliés
hors des flux, hors des liens –
l'air innocent ou *jaune*
dans le *pneu* de vent
et l'attitude appauvrie
de nos frères incessants :
ceux qui crurent montrer sans voir,
ceux qui ne mendient en rien
l'adhésion des couleurs, le reniflement des cailloux.
dans ou théâtre : fêlure
de l'apprivoisement.

- - - - -

L'autoroute ravive
le renard,
l'espionite.

Une nuit neigeuse
contourne la prose
de l'aventurier vieilli.

Michel MOUROT,
D'OÙ PARTIR, *Intimités du chaos*, 2007

Avant propos

La biologie moderne a l'ambition d'interpréter les propriétés de l'organisme par la structure des molécules qui le constituent. En ce sens, elle constitue un nouvel âge du mécanisme

C'est par ces mots dans l'introduction de *La logique du vivant* [1] que François Jacob caractérise la quête de la biologie du XX^e siècle dont le but avoué est la compréhension de l'architecture du vivant. La méthode consiste à tendre vers la décomposition maximale des constituants du vivant afin de caractériser un à un les éléments qui le composent et ainsi en décrire les rouages. Cette approche constitue un courant dominant au sein de la biologie que l'épistémologie qualifie de tomiste ou de réductionniste, et pour qui l'organisme est un tout qu'il convient d'expliquer par les seules propriétés de ses parties. En poussant à la caricature, la science de l'analyse de ces éléments devient donc une science de l'inerte. C'est donc tout naturellement que la biologie a constituée un nouveau terrain de jeux pour les physiciens et les chimistes au cours du siècle dernier (en particulier depuis Linus Pauling), permettant l'émergence de ce que l'on appelle la biophysique, une science d'interface où les outils d'observation des phénomènes physiques sont appliqués aux molécules d'origine biologique.

L'objectif de la biophysique est double. Tout d'abord elle doit permettre, par son approche "technique", l'étude des éléments les plus fondamentaux du vivant. Elle accompagne en effet *de facto* les progrès des outils de la physique qui permettent une étude toujours plus fine de ces éléments. En ce sens les expériences réalisées à l'échelle de la molécule unique présentées dans ce manuscrit en constituent certainement l'une des illustrations les plus abouties. D'autre part, les théories physiques servent à intégrer les observations au sein de modèles généraux visant à reconstituer de la complexité à partir des éléments les plus simples.

Les virus constituent certainement l'un des éléments du vivant les plus intéressants pour le biophysicien. En effet, le virus se situe à la frontière du vivant. Un virus est une particule

constituée d'un fragment d'acide nucléique enfermé dans une coque de protéines. Incapable de se multiplier seul, il nécessite une cellule hôte pour répliquer ses constituants. Cette particule biologique n'est donc qu'un assemblage de composants inertes, qui plus est en nombre limité. S'intéresser à l'étude des composants d'un virus revient donc à essayer de caractériser un assemblage simplifié de l'architecture du vivant.

Parmi les théories décrivant le vivant, la téléonomie de la relation structure/fonction mis en avant par Jacques Monod tient une place prépondérante. Conséquence probable du succès du livre *Le hasard et la nécessité* [2], on a tous été initié au paradigme selon lequel la structure biologique remplit une ou plusieurs fonctions, et toute fonction en biologie est assurée par une structure. Ainsi, la structure d'une protéine, c'est à dire sa conformation, devient responsable de son activité. Si c'est particulièrement vrai pour les enzymes ou les protéines structurales, cela l'est sans doute beaucoup moins pour les protéines dites chaperonnes, en particulier les ARN-chaperonnes, dont la fonction est de favoriser le repliement des acides nucléiques dans leurs conformations biologiquement actives. Nombre de ces protéines, de la classe des protéines intrinsèquement désordonnées, ne présentent en effet que peu voire pas de structure définie. Pour autant, elles sont douées d'activité. C'est le cas en partie de la protéine de la nucléocapside du VIH-1 (NCp7), pour qui près de la moitié de sa séquence ne présente aucune structure particulière. Pour autant, ces domaines non-structurés s'avèrent porteur d'une partie de l'activité de cette protéine absolument nécessaire à la réplication du VIH-1. Cette apparente contradiction est vite levée si le terme de "structure" est envisagée de manière plus intégrée. La structure ne doit pas concerner uniquement la notion de conformation particulière d'une molécule, mais au contraire s'élargir au sens d'architecture ou d'agencement des constituants. Dès lors, plus que la caractérisation des éléments fondamentaux du vivant, ce sont leurs organisations, leurs assemblages et leurs arrangements qu'il convient de caractériser. Il faut être capable de décrire leurs interactions puisque c'est dans la nature de leurs interactions, dans leur dynamique que naît la fonction. Les études biophysiques de systèmes dynamiques d'interaction, au sein de modèles de plus en plus reconstitués constitue un défi extrêmement intéressant. L'étude de petites protéines largement accessibles, comme la protéine de la nucléocapside du VIH-1, constitue dans cette optique, et nous le verrons au travers de multiples exemples dans les travaux réalisés dans

cette thèse, un excellent modèle, à la fois pour caractériser dans un système simple la nature des interactions de la NCp7 avec ses partenaires viraux, mais également pour développer des méthodologies aptes à appréhender ces dynamiques d'interactions.

Ayant commencé cet avant-propos par une citation de François Jacob, je terminerai en laissant à François Jacob le soin de rappeler que :

Chaque époque se caractérise par le champ du possible que définissent, non seulement les théories ou les croyances en cours, mais la nature même des objets accessibles à l'analyse, l'équipement pour les étudier, la façon de les observer et d'en parler.

Les travaux présentés dans ce manuscrit n'échappent bien évidemment pas à cette règle.

Table des matières

| | | |
|----------|--|------------|
| 1 | Introduction | 14 |
| 1.1 | Phylogénie | 14 |
| 1.2 | Le génome du VIH-1 | 17 |
| 1.2.1 | L'ARN viral | 17 |
| 1.2.2 | L'ADN proviral | 20 |
| 1.3 | La structure du VIH-1 | 22 |
| 1.3.1 | L'enveloppe | 22 |
| 1.3.2 | La matrice | 23 |
| 1.3.3 | La capside | 25 |
| 1.3.4 | La nucléocapside | 26 |
| 1.4 | Le cycle réplcatif du VIH-1 | 37 |
| 1.4.1 | La phase précoce de la réplcation de VIH-1 | 37 |
| 1.4.2 | La phase post-intégrative | 42 |
| 1.5 | L'infection à VIH | 50 |
| 1.6 | Traitements ARV | 53 |
| 1.7 | NCp7 | 57 |
| 1.7.1 | NCp7 : un produit de clivage | 57 |
| 1.7.2 | Structure | 58 |
| 1.7.3 | L'activité chaperonne de NCp7 | 61 |
| 1.7.4 | Rôle de la protéine NCp7 au cours du cycle viral | 66 |
| 1.8 | Objectifs de la thèse | 76 |
| 1.9 | Publication I | 78 |
| 2 | Matériels et méthodes | 114 |
| 2.1 | Fluorescence résolue en temps | 115 |

TABLE DES MATIÈRES

| | | |
|----------|--|------------|
| 2.1.1 | Technique du photoélectron unique | 115 |
| 2.1.2 | Déclins de fluorescence | 118 |
| 2.1.3 | Analyse des données | 119 |
| 2.2 | Anisotropie de fluorescence | 121 |
| 2.2.1 | Principe | 121 |
| 2.2.2 | Mesure d'anisotropie à l'état stationnaire | 122 |
| 2.2.3 | Mesure d'anisotropie résolue en temps | 125 |
| 2.3 | Transfert d'énergie non radiatif de fluorescence | 127 |
| 3 | Corrélation de fluorescence | 133 |
| 3.1 | Principe | 134 |
| 3.2 | Instrumentation | 137 |
| 3.3 | Applications | 138 |
| 3.4 | Quand la fluorescence n'est plus constante. | 156 |
| 3.5 | Cross-Corrélation ou corrélation croisée | 166 |
| 3.5.1 | Corrélation croisée à deux couleurs | 166 |
| 3.5.2 | Corrélation croisée à deux volumes focaux | 169 |
| 3.6 | Publication II | 170 |
| 4 | Rôle de NCp7 dans l'hybridation de cTAR/dTAR | 176 |
| 4.1 | Cinétique et mécanisme d'hybridation de cTAR/TAR | 177 |
| 4.1.1 | Contexte | 177 |
| 4.1.2 | Résultats | 178 |
| 4.2 | Publication III | 188 |
| 4.3 | Publication IV | 202 |
| 5 | Étude des propriétés chaperonnes de NCp7 à l'aide de la 2-Aminopurine | 220 |
| 5.1 | Caractérisation site spécifique de la liaison de NCp7 | 222 |
| 5.2 | Publication V | 225 |
| 5.3 | Rôle des doigts de zinc de NCp7 dans le second saut de brin | 237 |
| 5.4 | Publication VI | 239 |

TABLE DES MATIÈRES

| | | |
|----------|---|------------|
| 5.5 | Caractérisation de la déstabilisation locale de cTAR par la NCp7 | 271 |
| 5.6 | Publication VII | 273 |
| 5.7 | Analyse structurale de cTAR | 306 |
| 5.8 | Action de dTAR sur la stabilité de cTAR | 315 |
| 6 | Spectroscopie de fluorescence à l'échelle de la molécule unique | 318 |
| 6.1 | Matériels et Méthodes | 319 |
| 6.1.1 | Plateformes d'imagerie | 319 |
| 6.1.2 | Immobilisation des molécules | 320 |
| 6.1.3 | Validation de l'approche | 324 |
| 6.2 | Traitement du signal | 331 |
| 6.2.1 | Traitement des images | 331 |
| 6.2.2 | Traitement des signaux de fluorescence | 331 |
| 6.2.3 | HMM | 334 |
| 6.3 | Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique . . | 346 |
| 6.4 | Mesure de FRET à l'échelle de la molécule unique | 351 |
| 6.4.1 | Duplex d'ADN simplement marqués | 351 |
| 6.4.2 | PBS doublement marqué | 353 |
| 6.5 | Interaction de la Rh-NCp7 avec 5'Cy5-AATGCC | 355 |
| 6.5.1 | Mesures d'ensemble | 355 |
| 6.5.2 | Mesures en molécule unique | 357 |
| 6.6 | Interaction de la Rh-NCp7 avec (14-39)cTAR | 359 |
| 6.6.1 | Mesures d'ensemble | 359 |
| 6.6.2 | Mesures en molécule unique | 361 |
| 7 | Conclusions | 369 |
| 7.1 | Conclusions et perspectives | 370 |
| | Bibliographie | 375 |
| 8 | Annexe | 400 |
| 8.1 | Modèle de la liaison de n ligands sur un récepteur | 401 |

TABLE DES MATIÈRES

| | |
|---|------------|
| 8.2 Liaison de deux ligands sur un récepteur à deux sites de liaisons distincts . . . | 405 |
| 9 Autres travaux | 407 |
| 9.1 Publication VIII | 408 |
| 9.2 Publication IX | 411 |

Introduction

La partie introductive de ce manuscrit comporte une première partie s'attachant à décrire de manière succincte le Virus de l'Immunodéficience Humaine (VIH), la physiopathologie de l'infection à VIH ainsi que les traitements disponibles. Nous nous intéresserons dans un second temps aux propriétés de la protéine de la nucléocapside NCp7, objet de cette thèse.

1.1 Phylogénie et diversité génétique du VIH

Le VIH est un virus de la famille des *Retroviridae* du genre lentivirus. Il infecte préférentiellement les lymphocytes T CD4+ [3] ainsi que d'autres cellules du système immunitaire. Le VIH-1 est à l'origine d'une immunodéficience sévère, d'apparition progressive. Le stade terminal de l'infection, appelé SIDA (*Syndrome de l'Immunodéficience Acquise*), favorise l'émergence de maladies opportunistes ou tumorales qui finissent par être létales. Deux souches du VIH ont été identifiées jusqu'à maintenant : VIH-1 [4] et VIH-2 [5], qui se différencient notamment par leur distribution géographique (pandémie pour le VIH-1 alors que l'on parle encore d'endémie, principalement localisée en Afrique de l'Ouest, pour le VIH-2) ou par l'organisation de leur génome. Le VIH-2 apparaît moins pathogène.

Au sein du VIH-1, l'analyse des séquences génétiques des isolats viraux a montré une grande diversité de souches virales. Les VIH-1 sont classés en trois groupes phylogénétiques pour lesquels il s'agirait de trois transmissions indépendantes du virus simien vers l'humain :

le groupe M (*main*), représentant 90% des infections, est lui même divisé en 9 sous-groupes (A, B, C, D, F, G, H, J et K). La distribution géographique de ces sous-groupes ne semble pas aléatoire et on peut parler de zone d'endémie pour certains d'entre eux ;

1.1 Phylogénie

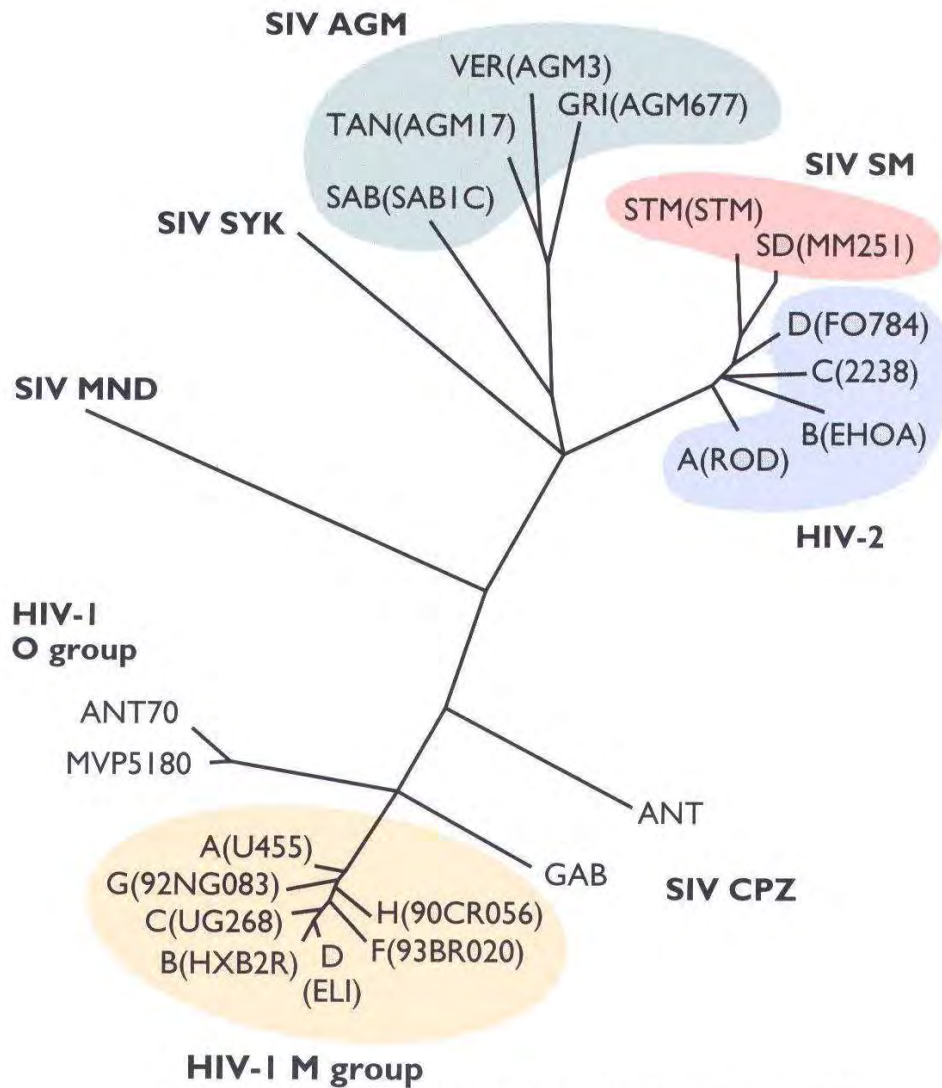
le groupe O (*outlier*), restreint à l'Afrique Centrale ;

le groupe N (*non-M non-O*) relativement anecdotique.

Certaines souches virales, considérées comme mosaïques, appartiennent à différents sous-types susnommés. Ces virus résultent de recombinaisons génétiques qui parviennent au moment de la phase de transcription inverse entre les ARN génomiques de différents sous-types co-encapsidés dans les mêmes particules virales. On appelle encore ces virus CRF (*Circulating Recombinant Forms*) dont les premières formes ont été mise en évidence dans les années 1990 [6]. Des évènements de recombinaison entre les groupes M et O ont été rapportés.

À titre d'illustration, l'arbre phylogénétique des lentivirus de primates, basé sur l'alignement des séquences du gène *pol* est présenté figure 1.1.

1.1 Phylogénie



S.J. Flint et al. 2000. Principles of virology, 1st Ed.

FIGURE 1.1 – Arbre phylogénétique des lentivirus de primate
Cet arbre est basé sur l'alignement des séquences du gène *pol*

1.2 Le génome du VIH-1

1.2.1 L'ARN viral

Le génome du VIH-1 est constitué de deux ARN de polarité positive, monocaténares, d'environ 9.2 kb, liés de manière non covalente. Cet ARN est coiffé et polyadénylé. Il est produit après l'intégration du provirus dans le génome de la cellule infectée. La transcription de l'ARN viral débute en 5' de la séquence R du 5'LTR et prend fin en 3' de la séquence R du 3'LTR (Figure 1.2).

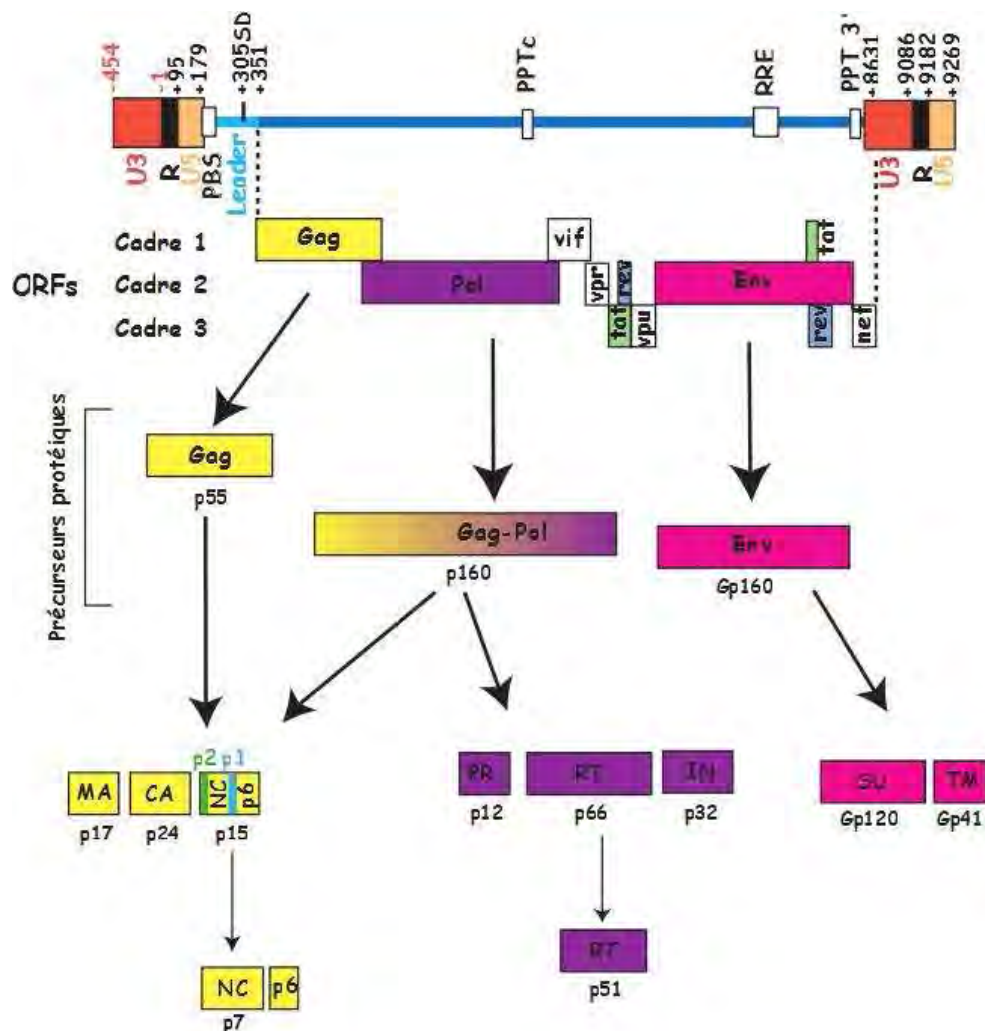


FIGURE 1.2 – Organisation générale du génome du VIH-1 [7].

1.2 Le génome du VIH-1

La région 5'

On retrouve dans cette région de nombreux motifs en tige-boucle [8, 9] impliqués dans différentes étapes du cycle réplcatif.

la séquence R en 5' et en 3'

Cette région se compose de deux séquences tige-boucles aux fonctions différentes :

- la tige-boucle TAR (*Trans-Activating Response element*) : composée des 55-60 premiers nucléotides de l'ARN viral. Cible de la protéine virale Tat, TAR joue, en 5' de l'ARNv, un rôle essentiel dans l'activation transcriptionnelle du génome proviral [10]. TAR joue également un rôle lors des étapes initiales de la transcription inverse [11] ou encore de l'assemblage du virus [12]
- une tige-boucle poly-A contenant le signal de polyadénylation (AAUAAA), en 3' de la séquence R. Ce motif n'est fonctionnel qu'en 3' de l'ARN génomique (ARNg) et permet le clivage du transcrit et l'addition de la queue poly-A [13]

la région U5

Le domaine U5 peut s'hybrider au domaine AUG, permettant la formation de la conformation dite BMH (*Branched Multiple Hairpin*). Il serait donc impliqué dans la régulation de l'orientation des ARN non-épissés soit vers la traduction soit vers la dimérisation et l'encapsidation (Figure 1.3).

la séquence PBS

Elle est immédiatement située en 3' de la région U5. Le *Primer Binding Site* (PBS) est une séquence de 18nt liant spécifiquement l'ARNt^{Lys3} cellulaire utilisé comme amorce lors de la transcription inverse.

la région leader

Elle est comprise entre la séquence PBS et le codon d'initiation de la traduction du gène *gag*. On retrouve dans cette région trois séquences en tige-boucle, respectivement nommées SL1 à SL3, et qui comporte :

- pour SL1, encore appelée DIS (*Dimerization Initiation Site*), un signal de dimérisation de l'ARN génomique [14]
- pour SL2, un site donneur d'épissage (SD). Le nombre important de sites donneur

1.2 Le génome du VIH-1

Les séquences internes d'intérêt

la séquence RRE (Rev-Responsive Element)

Séquence de 234 nt, située au sein de *env*. Elle est impliquée dans l'export nucléaire des transcrits mono ou non-épissés dont le mécanisme nécessite la protéine virale Rev et la machinerie de transport de la cellule hôte [18, 19].

les séquences PPT (PolyPurine Tract)

On distingue le PPT3' situé immédiatement en amont de la séquence U3 du LTR en 3' et le PPTc (central), localisé dans le cadre ouvert de lecture du gène *pol*. Ces séquences particulièrement riches en purines ne sont pas dégradées par l'activité RNase H de la RT. Elles servent d'amorces pour la synthèse du brin (+) de l'ADN proviral.

La région 3'

La région U3 contient les signaux nécessaires à l'intégration de l'ADN proviral [20], et comme mentionné plus haut, un signal de polyadénylation dans R.

1.2.2 L'ADN proviral

Le génome du VIH-1 est organisé en une région centrale codant pour les protéines de structure, des enzymes et les protéines accessoires, flanquée à ses deux extrémités par des régions non codantes.

Les régions non-codantes

Le génome du VIH-1 est bordé de deux séquences identiques ou LTR (*Long Terminal Repeat*), importantes pour l'intégration de l'ADN viral et pour sa transcription. Chaque LTR se divise en séquences U3, R et U5.

La séquence U3 contient les signaux nécessaires à la régulation de la transcription du provirus intégré par la machinerie cellulaire, à savoir

- le "core" du promoteur (de type eucaryotique), possédant trois sites de fixation pour le facteur de transcription (TF) Sp1 et une boîte TATA. Il permet la transcription d'ARNv à un taux basal.

1.2 Le génome du VIH-1

- le domaine activateur et le domaine modulateur qui fixent des TF cellulaires stimulant ou modulant la transcription.

Les régions codantes

L'ADN génomique code pour trois polyprotéines principales :

Le gène *gag*

Le gène *gag* ("*Group Specific Antigen*") code pour un précurseur myristylé de 55 kDa appelé Pr55^{Gag}. Après clivage par la protéase virale, Pr55^{Gag} donne naissance à la protéine de la matrice p17 (MA), à la protéine de la capsid p24 (CA), à la protéine de la nucléocapsid p7 (NCp7) et à la protéine p6. Par ailleurs, deux peptides p1 et p2, sont aussi générés, de part et d'autre de la région codante pour NCp7.

Le gène *pol* ("*polymerase*")

Il code pour un précurseur protéique de grande taille : le Pr160^{Gag-pol} qui provient d'un décalage -1 du cadre de lecture au niveau de la région codant pour p6 au cours de la traduction du gène *gag*. Le gène *pol* permet la synthèse d'une polyprotéine générant trois protéines enzymatiques : la protéase (PRp12) libérée par un mécanisme d'autocatalyse, l'intégrase (INp32) et la rétrotranscriptase (RTp66/p51).

Le gène *env* ("*enveloppe*")

Il code pour un précurseur polyprotéique Pr160^{env} transmembranaire dont la maturation se fait dans le réticulum endoplasmique. Son clivage par les protéases cellulaires permet la production d'une protéine de surface (SUgp120) et d'une protéine transmembranaire (TMgp41), deux glycoprotéines de membrane impliquées dans les interactions virus-cellule-hôte.

Le VIH-1 possède des gènes additionnels codant pour des protéines régulatrices (Tat et Rev) et auxiliaires (Nef, Vif, Vpu, Vpr). Les gènes codant ces protéines sont localisés de part et d'autre du gène *env*.

1.3 La structure du VIH-1

1.3 La structure du VIH-1

Les virions matures et infectieux ont une forme sphérique de 110 – 130 *nm* de diamètre. Ils sont enveloppés d'une membrane phospholipidique cellulaire. Directement sous cette bicouche lipidique se trouve un feuillet protéique (la matrice) à l'intérieur de laquelle se trouve un "core" conique central (la capside) contenant l'ARN génomique viral, des ARN d'origine cellulaire ainsi que des protéines virales et cellulaires. La particule infectieuse mature ne résulte pas de l'assemblage de chacune de ces protéines individuelles. Le virus résulte de l'assemblage de trois précurseurs polyprotéiques (Gag, Env, Pol), dont le clivage au cours de l'étape de maturation permet de générer l'architecture et l'infectiosité de la particule virale mature.

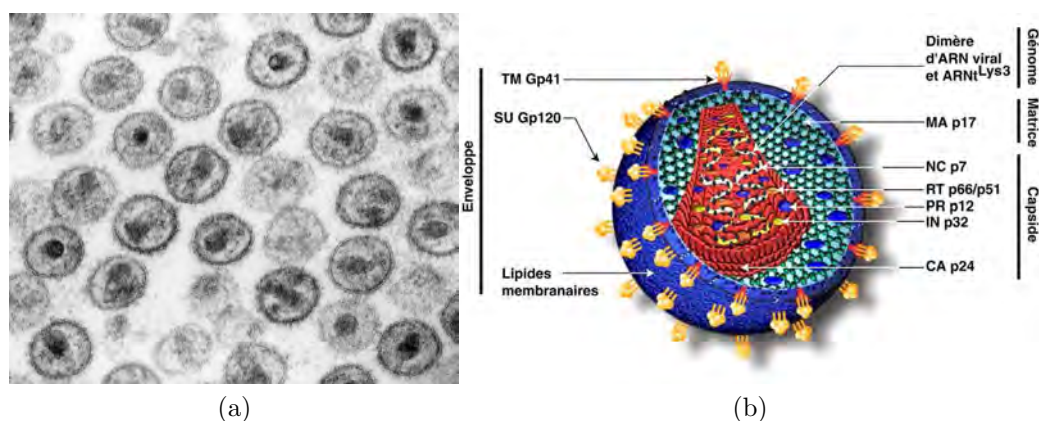


FIGURE 1.4 – Structure du VIH-1

- (a) Image par microscopie électronique du VIH-1
- (b) Schéma de la particule virale du VIH-1

1.3.1 L'enveloppe

L'enveloppe du virus est une bicouche lipidique provenant de la membrane plasmique de la cellule hôte. Elle a pour particularité d'être particulièrement riche en cholestérol et en glycosphingolipide [21], du fait de l'affinité particulière de Gag pour les "radeaux lipidiques", ce qui faciliterait son bourgeonnement [22]. Des protéines d'origine cellulaire (telles que les molécules HLA, ICAM-1 [23]), ou d'origine virale (TMgp41, SUgp120) y sont sélectivement insérées. Les protéines SUgp120 et TMgp41 sont associées en hétérotrimères à la surface

1.3 La structure du VIH-1

des virions [24].

La sous-unité SUgp120 comporte un site de liaison pour les récepteurs CD4, et pour les récepteurs des chimiokines CCR5 et CXCR4 [25]. C'est une protéine richement glycosylée avec 24 sites potentiels de glycosylation (séquences Asn-X-Ser/Thr) [26, 27] dont 13 sont conservés parmi les différents isolats de virus. Le profil de glycosylation dépend de la cellule infectée, ce qui pourrait jouer un rôle dans l'infectiosité du virus [28, 29]. La variabilité de la surface du virus est probablement un facteur important pour échapper au système immunitaire [30, 31, 32]. Sa séquence, d'environ 480 acides aminés, comporte cinq régions variables (V1 à V5) et cinq régions relativement conservées (C1 à C5).

La sous-unité TMgp41 (345 acides aminés) est impliquée dans la fusion des membranes au moment de l'entrée du virus dans la cellule. Elle se subdivise en un ectodomaine $N_{terminal}$, qui comprend notamment le peptide de fusion, un domaine transmembranaire, et une queue cytoplasmique qui interagit avec la matrice. L'ectodomaine est un trimère symétrique, dont chaque monomère consiste en deux hélices α antiparallèles reliées par une large boucle (les hélices N_{ter} se retrouvent au centre alors que les domaines C_{ter} se positionnent autour selon un arrangement anti-parallèle) (Figure 1.5). Contrairement à SUgp120, TMgp41 présente des domaines très conservés, notamment le domaine de fusion, des régions *heptad* permettant la trimérisation ou encore un domaine transmembranaire nécessaire à son insertion dans la membrane virale [33].

1.3.2 La matrice

La matrice (protéine p17) est obtenue à partir du clivage du précurseur polyprotéique Pr55^{Gag} en son extrémité N_{ter} (pour des revues, voir [36, 37]).

MAp17 est une protéine de 132 aminoacides, présentant un résidu Glycine en $N_{terminal}$ pouvant être myristoylé. Sa structure tridimensionnelle est composée de cinq hélices α , dont quatre se groupent pour former une tête globulaire [38, 34]. Cette protéine multifonctionnelle possède plusieurs signaux de ciblage (signal d'adressage à la membrane, signal de localisation nucléaire) régulés par des modifications structurales et biochimiques telles que la myristoylation, la multimérisation, ou la phosphorylation. Au sein du complexe Pr55^{Gag}, la MA permet d'adresser Pr55^{Gag} vers les compartiments cellulaires où s'opère l'assemblage.

1.3 La structure du VIH-1

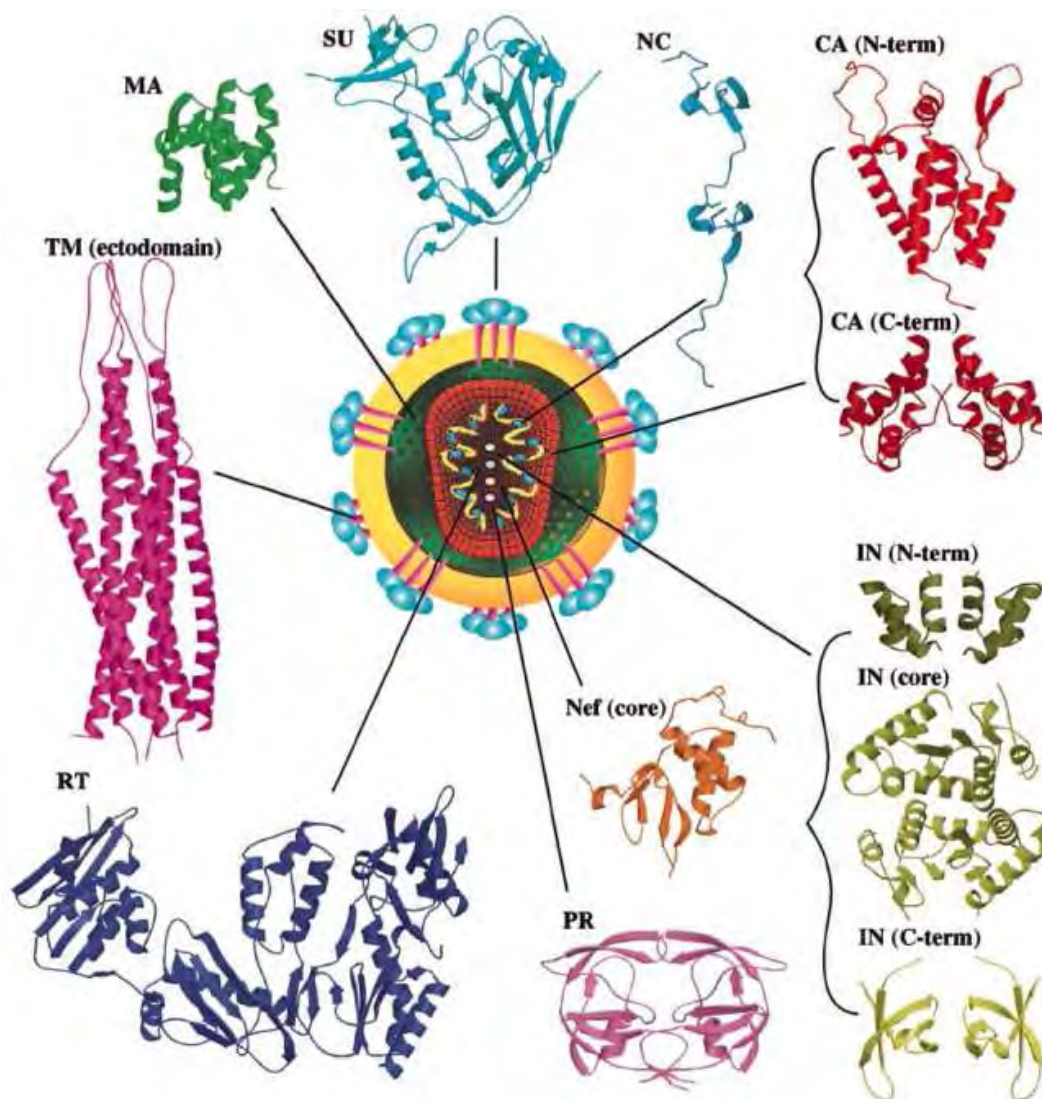


FIGURE 1.5 – Structure des protéines (ou fragments protéiques) de VIH-1. L'échelle est identique pour toutes les structures. L'ectodomaine TM correspond à celui du SIV.

L'exposition du myristate, nécessaire à l'association stable de p17 avec les membranes, serait déclenchée par l'oligomérisation de Pr55^{Gag} et son interaction avec le PI(4,5)P2 (un phosphatidylinositol phosphate) [39, 40]. La maturation protéolytique de la matrice [41] inhiberait cette exposition, provoquant une destabilisation de l'oligomérisation. Le domaine de la MA au sein de Pr55^{Gag} serait essentiel pour recruter la glycoprotéine Env [42]. La forme monomérique de MA serait, quand à elle, impliquée dans l'étape de décapsidation [43], et le transport du complexe de transcription inverse (RTC) dans le cytoplasme via son interaction avec l'actine. La MA possède un domaine basique de liaison aux ARN, avec un

1.3 La structure du VIH-1

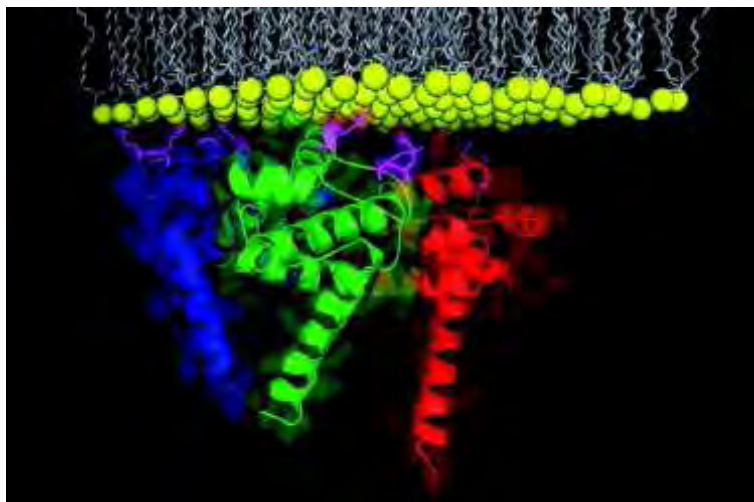


FIGURE 1.6 – Représentation de la structure cristalline d'un trimère de la matrice du VIH-1 interagissant avec une membrane lipidique (adapté de [34] par Adamson & al. [35])

site de liaison préférentiel pour une séquence homologue aux nucléotides 1433-1446 de *pol* [44], laissant supposer que la MA et la nucléocapside puissent agir de manière concomitante dans le recrutement de l'ARNg au cours de l'assemblage [23].

1.3.3 La capsid

Dans le virus mature, la capsid (CA) a une forme icosaédrique de type fullerène. Elle est constituée par l'assemblage de la protéine p24 (232 acides aminés), qui est, comme p17, produite à partir de la maturation du précurseur polyprotéique Pr55^{Gag}. Des résolutions partielles de p24 ont tout d'abord été déterminées par cristallographie et RMN [45, 46, 47, 48], permettant la modélisation de la totalité de la capsid [49, 50], puis des modèles de microscopie électronique et de RMN ont permis de déterminer la structure complète de la protéine [51, 52, 53, 54]. La p24 est composée de deux domaines indépendants reliés par une portion flexible : le domaine N_{term} , composé de 7 hélices α , et le domaine C_{term} , structuré en 4 hélices α (Figure 1.7). Le domaine N_{term} , qui constitue le corps de la protéine, est impliqué dans la condensation et la morphogénèse de la capsid. De plus, une région de ce domaine présente un site de liaison pour la Cyclophiline A, une enzyme cellulaire à activité protéinase, encapsidée au sein de la particule virale, et qui faciliterait l'étape de décapsidation [45]. Après maturation, les protéines p24 s'organisent principalement sous

1.3 La structure du VIH-1

forme d'hexamères dont la partie interne est constituée par six domaines N_{term} et la partie externe par les domaines C_{term} qui dimérisent pour lier les différents hexamères les uns aux autres. La capside résulte de l'assemblage d'environ 250 hexamères et d'exactly 12 pentamères (Figure 1.7). Ainsi, la CA joue un rôle fondamental dans la structuration de la particule virale.

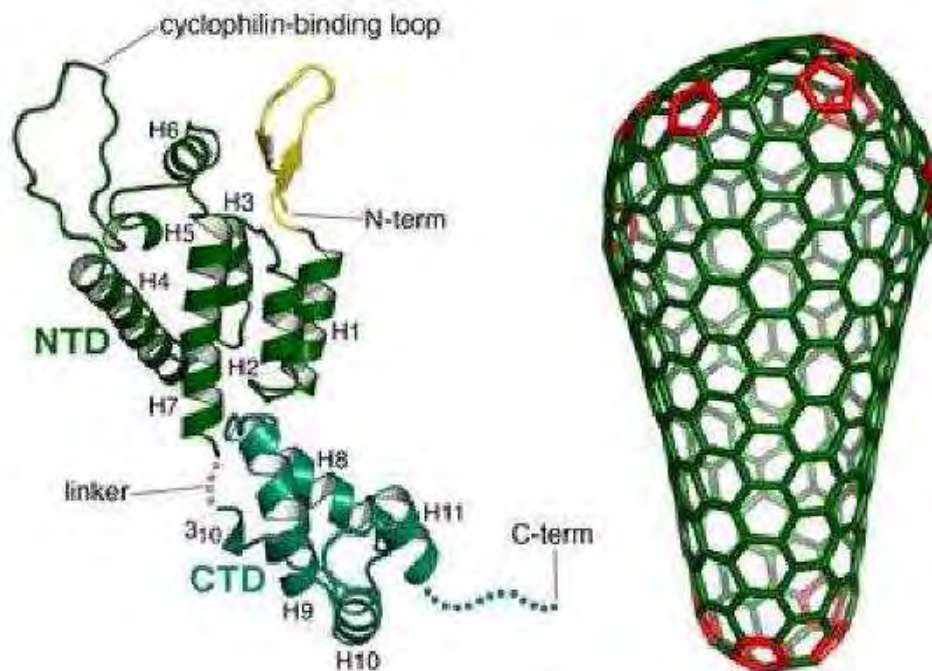


FIGURE 1.7 – Structure de la capside du VIH-1 [53]

L'antigénémie p24 peut être utilisée comme marqueur diagnostique de l'infection par le VIH-1 (tests combinés), même si les tests de dépistage reposent plutôt sur la détection d'anticorps anti-HIV (tout test ELISA positif est confirmé par Western Blot. Ce dernier est considéré positif lorsqu'au moins un anticorps dirigé contre p24, gp41, gp120 ou gp160 est détecté). En outre, il semblerait que le taux de p24 plasmatique constitue un indicateur précoce de la progression de l'infection à VIH-1 [55]

1.3.4 La nucléocapside

On retrouve dans la nucléocapside les deux molécules d'ARN génomique recouvertes et protégées par environ 1500 à 2000 unités de protéine de la nucléocapside (NCp7), les

1.3 La structure du VIH-1

enzymes virales nécessaires à la réplication comme la rétrotranscriptase (RTp66/p51), la protéase (PRp12) et l'intégrase (INp32) ainsi que des protéines et ARNt d'origine cellulaire.

La protéine de la nucléocapside

La NCp7 est une petite protéine basique de 55 acides aminés présentant des régions en doigts de zinc. Sa structure détaillée et sa contribution dans le cycle viral seront traitées section 1.7.

La protéine p6

La p6 est une petite protéine de masse moléculaire de 6 à 7 kDa (51 AA). Elle est issue du clivage de l'extrémité C_{term} du précurseur polyprotéique Pr55^{Gag}. Elle présente deux régions fonctionnelles importantes dont les séquences sont conservées, appelées région-L, L pour « *late* », puisque ces motifs jouent un rôle critique dans la phase tardive du cycle viral en recrutant les protéines cellulaires nécessaires au bourgeonnement du virus [56]. Une mutation à l'intérieur de ces motifs réduit significativement la production de particules virales [57]. De plus, ce type de motif est retrouvé chez de nombreux rétrovirus. Les motifs L sont reconnus par Tsg101 et AIP1 (ou Alix) [57, 58, 59], deux facteurs cellulaires impliqués dans la biogénèse des endosomes tardifs et le tri des molécules "cargo" à leur niveau. La mono-ubiquitinylation de deux résidus lysines au sein de la séquence de p6 pourrait également permettre (en synergie avec d'autres sites mono-ubiquitinylés de Pr55^{Gag}) le recrutement de facteurs cellulaires impliqués dans le relargage du VIH-1 [60]. D'autre part, p6 est capable de se lier à Vpr (*Viral protein R*), cette interaction permettant l'encapsidation de Vpr dans les virions [61, 62].

Les protéines enzymatiques

La protéase

La protéase (PR) est une aspartyl-protéase. Elle est générée par le clivage de Pr160^{Gag-Pol} selon un mécanisme d'autocatalyse [63]. La PR est une enzyme clé indispensable à la maturation du virus et dont l'inhibition induit la production de virions non-infectieux [64]. La

1.3 La structure du VIH-1

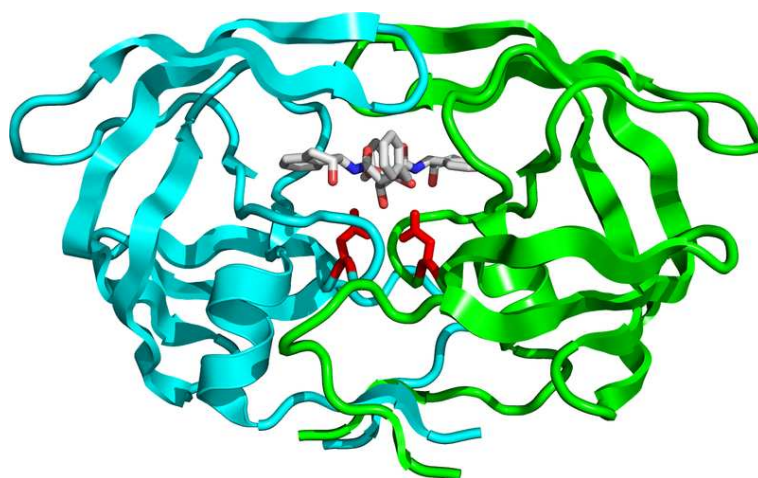


FIGURE 1.8 – Structure de l’homodimère de la protéase du VIH-1 (fichier PDB 1EBY)

PR est fonctionnelle sous forme d’homodimère. Les deux sous-unités de 99 AA interagissent par leur extrémité N_{term} et C_{term} , ainsi que par une région “flap” qui est un motif adaptatif réagissant à la liaison du substrat, pour augmenter l’affinité de l’enzyme pour celui-ci. Le site actif est constitué par l’arrangement de deux motifs Asp-Thr-Gly (Asp25, Thr26 et Gly27), caractéristiques des aspartyl-protéases. Ces motifs sont portés par chacune des deux sous-unités. Le site catalytique est situé au centre d’une poche de liaison au substrat au niveau duquel une molécule d’eau assure la protonation alternative des deux résidus Asp25 [65]. La reconnaissance des sites de clivage au sein de Pr55^{Gag} ou de Pr160^{Gag-Pol} semble se faire *via* des conformations spécifiques [66, 67] plus que par une spécificité de séquence. La protéase, de part son importance quand à l’infectiosité du virus, est une cible privilégiée des thérapies anti-rétrovirales.

L’intégrase

L’intégrase (IN) catalyse une série de réaction comme le *3’ processing* et le transfert de brin, aboutissant à l’insertion de l’ADN proviral au sein du génome de la cellule hôte. Elle est codée par Pr160^{Gag-Pol}. Sous sa forme fonctionnelle, elle se présente sous forme tétramérique. L’IN est constituée de 288 AA (32kDa) formant trois domaines distincts : un domaine N_{term} (1-50) qui lie le Zn^{2+} , un domaine catalytique (51-212) et un domaine C_{term} liant l’ADN [69]. La structure complète de cette structure enzymatique a été modélisée à

1.3 La structure du VIH-1

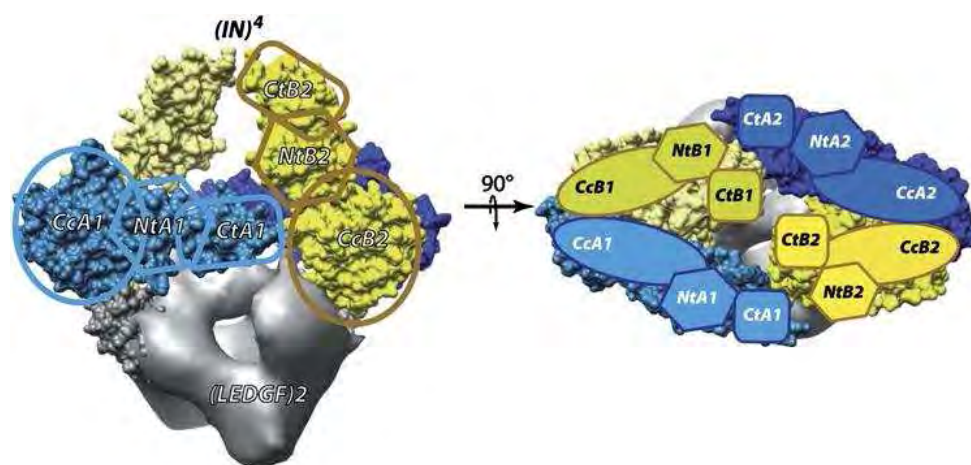


FIGURE 1.9 – Structure du complexe intégrase-LEDGF/p75

Ce complexe est formé de 4 IN organisées en 2 dimères asymétriques (A1-A2 et B1-B2) s'associant à 2 LEDGF/p75 [68]

partir de la résolution de différents domaines [70]. En fait, l'IN s'associe avec la protéine LEDGF/p75 (*human lens epithelium-derived growth factor/transcription co-activator p75*) pour former un complexe nucléaire (4 IN et 2 LEDGF/p75) capable de lier l'ADN viral et l'ADN cellulaire cible et de permettre l'intégration du génome viral [71, 68].

La réverse transcriptase

La réverse transcriptase (RT) est codée par le gène *pol*. Elle est un produit de clivage de Pr160^{Gag-Pol}. La formation de l'hétérodimère p66/p51 résulte de l'homodimérisation préalable de deux sous-unités p66. Ce dimère subit ensuite un clivage sélectif interne d'une seule de ses sous-unités.

La structure de la RT a été très largement étudiée, avec plus d'une centaine de structures déterminées : RT libre, RT complexée avec différents types d'inhibiteurs, RT associée au complexe amorce/matrice, en présence ou non de dNTP. L'obtention de ces structures a permis une grande avancée dans la compréhension du mécanisme de l'activité catalytique de la transcriptase inverse et de l'émergence des virus résistants aux drogues antirétrovirales.

1.3 La structure du VIH-1

La sous-unité p66

La sous-unité p66 est formée de 560 acides aminés et est organisée en 5 domaines. La structure de la RT présente une similitude grossière avec une main droite (Figure 1.11), d'où le nom de ses différents domaines :

- i.* Les doigts (AA 1-84 et 120-150) sont impliqués dans le contact de la RT avec le brin matrice et dans la processivité de la RT lors de la polymérisation [72, 73, 74].
- ii.* La paume (AA 85-119 et 151-244) contient le site catalytique de l'activité polymérase avec la séquence consensus YMDD (Tyr183-Met184-Asp185-Asp186) conservée dans les ADN polymérases [75, 76, 77]. Le site catalytique contient deux ions Mg^{2+} qui sont maintenus grâce à leur interaction avec les résidus Asp110, Val111, Asp185 et Asp186 [78, 57, 77]. Les résidus 229 à 232 forment le motif "primer grip" impliqué dans le positionnement du complexe matrice/amorce [79, 80, 81].
- iii.* Le pouce intervient dans la fixation du complexe matrice/amorce. Ce domaine contient également un motif conservé *helix clamp* riche en résidus basiques, qui participe à la liaison avec les acides nucléiques [82, 83, 73, 84].
- iv.* Le domaine de connection (AA 323-437) relie le domaine de polymérisation au domaine RNase H. Il contient un grand nombre de résidus hydrophobes responsables de la dimérisation, assurant une conformation correcte de l'enzyme et une activité RNase H efficace [85, 86].
- v.* Le domaine RNase H (AA 438-460) n'adopte sa conformation active qu'en interagissant avec la sous-unité p51 dont le pouce forme le plancher du site catalytique. Celui-ci est formé par les acides aminés Asp443, Glu478, Asp498, His539 et Asp549. Sa structure est très proche de celle de la RNase H de *E.coli* ou de *T. thermophilus* [87, 88]

La sous-unité p66 contient également un site de fixation des dNTP. Le nucléoside triphosphate destiné à être incorporé au brin d'ADN est stabilisé grâce à des interactions ioniques *via* les deux ions Mg^{2+} et les résidus Lys65 et Arg72 ainsi que par des liaisons hydrogènes formées avec les groupements $-NH$ des résidus Asp113, Ala114 et Tyr115 [57].

1.3 La structure du VIH-1



FIGURE 1.10 – Structure tridimensionnelle de la RT du VIH-1 complexée avec un acide nucléique.

La sous-unité p51 est en gris, les doigts sont représentés en bleu, la paume en rouge, le pouce en vert, le domaine de connexion en jaune et le domaine RNaseH en mauve.

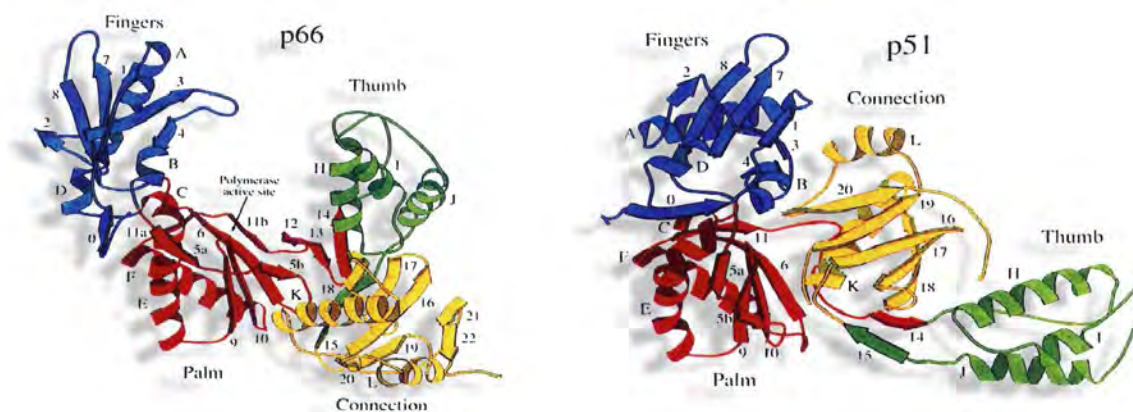


FIGURE 1.11 – Structure tridimensionnelle des sous-unités p66 et p51 de la RT du VIH-1. Les doigts sont représentés en bleu, la paume en rouge, le pouce en vert et le domaine de connexion en jaune.

1.3 La structure du VIH-1

La sous-unité p51

La sous-unité p51 est composée de 440 acides aminés. Elle possède les mêmes domaines que la sous-unité p66 à l'exception du domaine RNase H, mais sa conformation est très différente (Figure 1.11), expliquant sans doute pourquoi cette sous-unité ne possède pas d'activité enzymatique. Il a été suggéré que la sous-unité p51 jouerait un rôle dans le repliement de la sous-unité p66. Elle pourrait favoriser sa fixation sur la matrice et stabiliserait ainsi le complexe ternaire de transcription inverse [11, 89].

L'activité RNase H de la RT

L'activité RNase H (Ribonucléase H) permet l'hydrolyse d'un brin ARN engagé dans un hybride ARN-ADN. La RNase H des RT rétrovirales est similaire aux RNase H des bactéries ou des cellules eucaryotes [90], mais présente un mécanisme de clivage différent puisque la RNase H libère des extrémités 3'OH et 5'-PO₄⁻. Ces groupements pourront être utilisés pour former de nouvelles liaisons 3',5'-phosphodiester. Cette propriété est importante pour le rôle joué par la RNase H lors de la transcription inverse rétrovirale. En effet, le clivage de l'ARN rétroviral par la RNase H au cours de la synthèse du brin d'ADN(-) génère les amorces pour la synthèse du brin (+).

Deux types de clivage ont été mis en évidence : le clivage polymérase dépendant et indépendant [91, 92] :

- i.* L'activité RNase H polymérase dépendante clive l'ARN 17 à 18 nucléotides en amont de l'extrémité 3' terminale de l'ADN néosynthétisé [93]. Cette distance est dictée par la séparation spatiale entre les deux sites catalytiques de la RT. Dans la mesure où la vitesse de clivage de l'ARN est plus faible que celle de l'incorporation des dNTP, l'ARN n'est pas totalement dégradé et des oligonucléotides d'ARN restent hybridés à l'ADN néosynthétisé [93, 94].
- ii.* L'hydrolyse indépendante de l'activité de polymérisation permet de cliver les ARN de 18 nt en ARN de 9 nt. Ces ARN se dissocient probablement de l'ADN néosynthétisé.

1.3 La structure du VIH-1

Mécanisme de polymérisation par la RT

La RT est une ADN polymérase ARN et ADN dépendante, elle a donc la capacité d'allonger une amorce ARN ou ADN hybridée à une matrice. L'enzyme utilise de façon ordonnée deux substrats, l'acide nucléique et le nucléotide, selon un mécanisme de type bi-bi ordonné.

La réaction de polymérisation peut être séquencée comme suit [77] (Figure 1.12) :

1. Fixation de l'enzyme sur la matrice (ARN ou ADN), induisant un changement conformationnel de la RT, notamment un mouvement du pouce favorisant un placement correct du couple matrice/amorce.
2. Fixation du dNTP conduisant à un nouveau changement conformationnel de la RT qui adopte une position plus fermée.
3. Attaque nucléophile de l'extrémité 3'-hydroxyle de l'amorce sur le phosphate α du nucléotide entrant formant une liaison phosphodiester. L'amorce est allongée d'un nucléotide.
4. Translocation de la RT favorisant un nouveau cycle catalytique. La translocation se fait soit de manière concomitante à la libération du pyrophosphate (PPi), soit juste après. Le relargage du PPi pourrait favoriser l'ouverture des doigts et l'obtention d'un complexe ouvert.

La RT a la capacité de polymériser l'ADNc à très faible concentration en dNTP. C'est sans doute une des raisons expliquant pourquoi le virus peut se répliquer au sein de cellules quiescentes [95].

La synthèse d'ADN est dite processive lorsque la RT incorpore plusieurs nt sans se dissocier du complexe A/M. Elle est dite distributive lorsque la RT se dissocie du complexe A/M après addition de chaque nt. Le rapport k_{pol}/k_{off} détermine le type de polymérisation (processive si ce rapport est supérieur à 1, distributive sinon). Un troisième type de polymérisation correspond à la terminaison forcée, induite par la présence d'un seul nucléotide pouvant être incorporé, ou par l'incorporation d'un terminateur de chaîne. Dans ce cas, la RT ne peut plus incorporer de nucléotides.

1.3 La structure du VIH-1

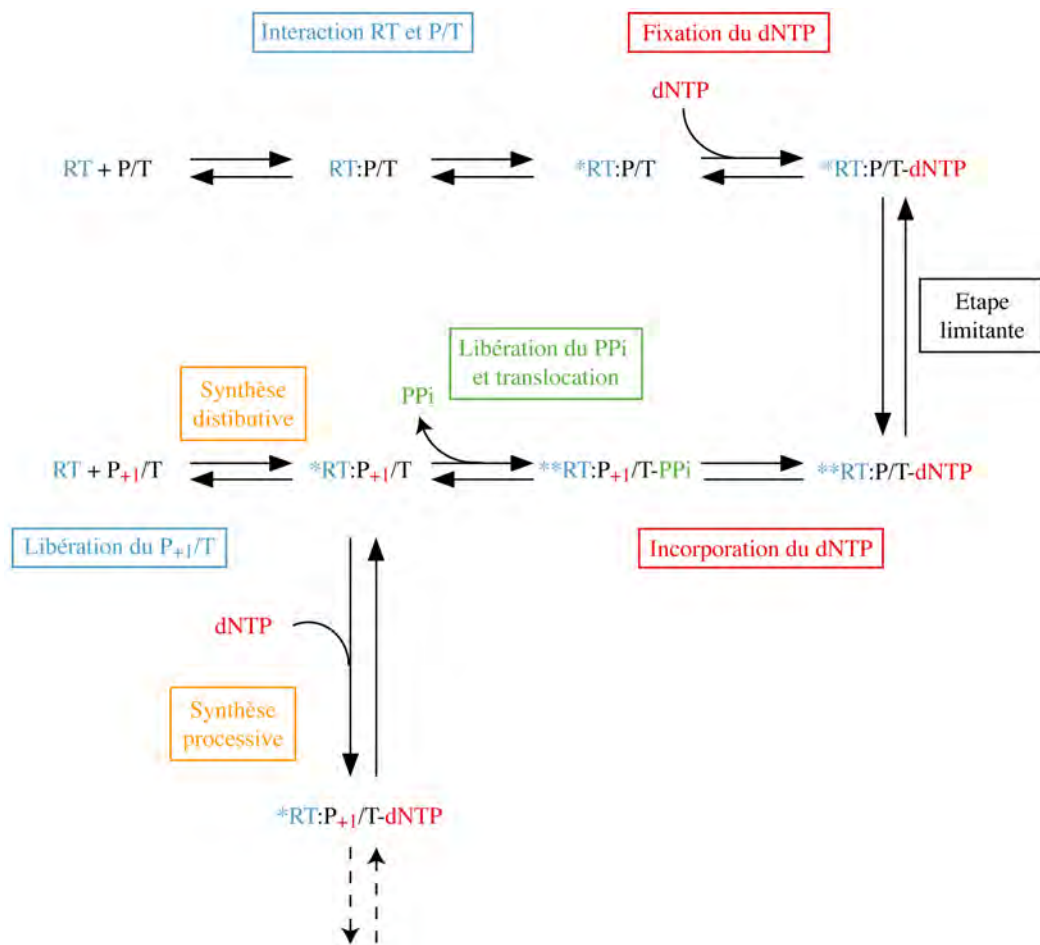


FIGURE 1.12 – Représentation schématique des étapes de polymérisation d'un nucléotide par la RT de HIV-1 [7].

1.3 La structure du VIH-1

Les protéines régulatrices et auxiliaires

Des protéines régulatrices (Tat et Rev) et auxiliaires (Nef, Vif, Vpr et Vpu) sont également codées par le génome du VIH-1.

Tat (*TransActivator of Transcription*)

Outre son activité de transactivation de la transcription qui sera décrite au cours de la description du cycle réplcatif [10], Tat serait également impliquée dans la régulation de la transcription inverse [96], l'élaboration de la coiffe en 5' de l'ARNm [97], et l'épissage [98]. Tat jouerait aussi un rôle de répression de l'expression des molécules du complexe majeur d'histocompatibilité (CMH-1) [99].

Rev (*Regulatory of Virion Expression*)

La protéine Rev (116 acides aminés) est un trans-activateur post-transcriptionnel qui possède un signal d'export nucléaire (NES pour *Nuclear Export Signal*) ainsi qu'un signal d'import nucléaire (NLS, *Nuclear Localisation Signal*) [100], lui permettant d'intervenir dans l'export nucléaire des ARNm non épissés ou mono-épissés, lors de la phase tardive de l'expression du génome viral. De plus, Rev augmente la stabilité des ARNm viraux non épissés [101].

Vpr (*Viral Protein R*)

La protéine Vpr (96 acides aminés) est encapsidée dans le virion [102] grâce à son interaction avec le domaine C-terminal de Pr55^{Gag} [62]. Vpr est localisée dans la nucléocapside en association avec l'ARN viral [103, 104]. Au cours du cycle viral, Vpr est impliquée dans la fidélité de la transcription inverse, dans la trans-activation à partir du LTR, et dans l'import nucléaire de l'ADN viral au sein du complexe de pré-intégration (PIC) [105]. Vpr permet au virus de se répliquer dans des cellules quiescentes. De plus, Vpr conduit à l'arrêt du cycle cellulaire en phase G2/M [106, 107], et a un effet direct sur la perméabilité de la mitochondrie en interagissant avec des protéines du pore mitochondrial, induisant l'apoptose des cellules infectées [108].

Nef (*Negative Regulatory Factor*)

La protéine Nef (206 acides aminés) augmente l'infectiosité du VIH-1. L'expression de Nef se fait rapidement après la transcription virale, suggérant qu'elle est nécessaire à la propagation efficace du virus. Nef favorise en effet la survie des cellules infectées notamment en

1.3 La structure du VIH-1

modulant à la baisse l'expression de plusieurs molécules de surface : *i.* les CD4 et CD28 évitant ainsi une surinfection détrimentale des cellules [109] et *ii.* des molécules du complexe majeur d'histocompatibilité I (CMH I) et du CMH II présentes sur les cellules présentatrices d'antigène (CPA) et impliquées dans la reconnaissance des cellules infectées par le système immunitaire [110, 111].

Vif (*Virion Infectivity Factor*)

La protéine Vif (192 acides aminés) est encapsidée en interagissant avec la région 5' de l'ARN génomique [112], ainsi qu'avec le précurseur Gag [113]. Vif stimule la dégradation par le protéasome de cytidines désaminases cellulaires de la famille des APOBEC (*APOlipoprotein B mRNA editing Enzyme, Catalytic polypeptide*). APOBEC3G est responsable d'hypermutations du génome viral au cours de la rétro-transcription. APOBEC3G constitue donc une défense de l'hôte contre l'infection virale. Cette défense est rendue inefficace par la présence de Vif. La présence de Vif permet la réplication virale dans des cellules non permissives. D'autre part, Vif stimulerait la transcription inverse en augmentant, en synergie avec NCp7, la processivité de la transcriptase inverse et en activant le premier saut de brin [114].

Vpu (*Viral Protein U*)

Vpu (81 acides aminés) est une protéine membranaire présente uniquement dans le VIH-1 et quelques isolats de SIV. Vpu induit la dégradation via le protéasome des récepteurs CD4 co-localisés avec les protéines Env au sein du réticulum endoplasmique [115]. Vpu favorise également le bourgeonnement des particules virales à la surface des cellules infectées [116], particulièrement dans les cellules exprimant la tetherine, un interféron responsable de l'inhibition du bourgeonnement viral.

1.4 Le cycle répliatif du VIH-1

1.4 Le cycle répliatif du VIH-1

La connaissance des différentes étapes du cycle répliatif est essentielle à la compréhension de la physiopathologie du VIH et à la mise au point des thérapeutiques. Les principales étapes sont :

Entrée : fixation à la surface des cellules portant le récepteur CD4. Le virus pénètre dans la cellule après interaction avec d'autres corécepteurs, en particulier les récepteurs des chimiokines CCR5 ou CXCR4.

Synthèse de l'ADN proviral lors de l'étape de rétro-transcription.

Intégration de l'ADN proviral au génome de la cellule infectée

Transcription de l'ADN proviral en ARNm et **synthèse des protéines virales** par la machinerie cellulaire

Assemblage, bourgeonnement et maturation conduisant à la formation de nouvelles particules virales infectieuses.

1.4.1 La phase précoce de la répliation de VIH-1

Reconnaissance, fusion et entrée

La reconnaissance de la cellule cible se fait par l'intermédiaire de la liaison de la glycoprotéine de surface SUgp120 sur un récepteur CD4 présent sur la surface de la cellule cible. Le cluster de différenciation 4 (CD4) est un marqueur protéique exprimée à la surface des lymphocytes T CD4+, des cellules régulatrices T, des monocytes, des macrophages et de certaines cellules dendritiques. L'interaction entre ces deux protéines de surface provoque un changement conformationnel de la gp120, permettant la liaison simultanée de celle-ci à un corécepteur. Plus d'une dizaine de co-récepteurs ont été identifiés, mais les principaux sont CXCR4 pour les lymphocytes T CD4+ et CCR5 pour les macrophages. La formation du complexe ternaire gp120/CD4/co-récepteur libère le domaine de fusion de la gp41 qui peut alors être inséré dans la membrane plasmique de la cellule. Cette dernière étape permet la fusion des membranes virale et cellulaire. La libération de la capsid virale au sein du cytoplasme requiert donc nécessairement le passage celle-ci au travers du réseau cortical

1.4 Le cycle réplcatif du VIH-1

cellulaire d'actine qui peut se comporter comme une barrière protectrice [117]. La protéine Nef, qui augmente l'infectivité de VIH-1, pourrait aider le virus à passer cette barrière [118].

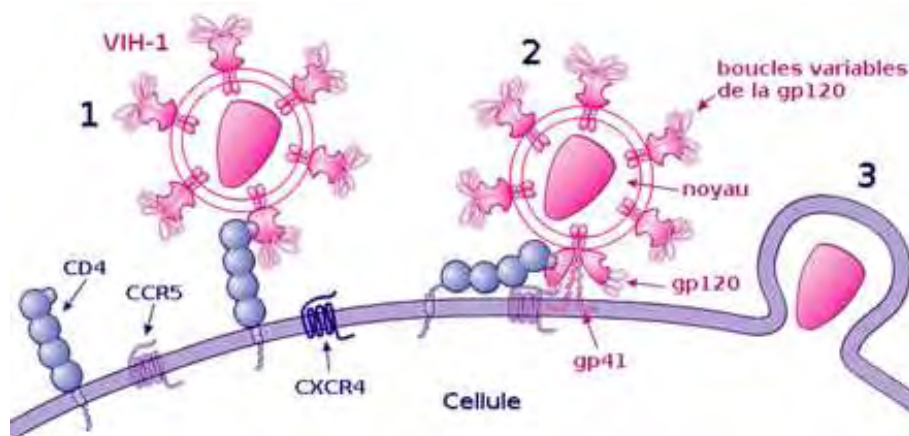


FIGURE 1.13 – Processus d'entrée du VIH-1

1. Fixation de la gp120 au récepteur CD4

2. Fixation d'une boucle variable de la gp120 au co-récepteur et fixation de la gp41 sur la membrane cellulaire

3. Pénétration dans la cellule.

Si le mode de reconnaissance, de fusion et d'entrée est principalement celui décrit plus haut, le virus peut également infecter des cellules non présentatrices de CD4, comme les lymphocytes B [119], les LT CD8+ [120] ou encore les trophoblastes [121], montrant que VIH-1 présente d'autres moyens pour s'adsorber à la surface de nombreux types cellulaires, en faisant intervenir des molécules de la cellule hôte qu'il a incorporé dans sa membrane et leurs récepteurs correspondants situés dans la membrane des cellules cibles [122, 123, 124].

Décapsidation et rétrotranscription

Pour permettre l'intégration du génome viral au sein du génome de la cellule infectée, l'ARN génomique doit être converti en ADN double brin, ce qui est réalisé au cours de la réverse transcription (RTion). La RTion est un processus complexe, catalysé par la Reverse Transcriptase (RT) et facilité par la NCp7. Elle nécessite deux transferts de brin. La RTion constitue une étape essentielle des étapes précoces de la réplication virale. Ainsi, l'organisation spatio-temporelle des événements permettant la réverse transcription du génome viral est représentée schématiquement par un processus séquentiel dans lequel la capsid

1.4 Le cycle répliatif du VIH-1

est, une fois relarguée dans le cytoplasme, déstructurée pour former le complexe de réverse transcription (RTC). La cyclophiline A, une protéine cellulaire présente dans les particules virales, jouerait un rôle dans le processus de décapsidation [125, 126]. Le RTC, au sein duquel se déroule la RTion, se retrouve alors exposé aux desoxyribonucleotides présents dans le cytoplasme. Ceci permet la rétro-synthèse d'ADN par la RT. La RT qui se trouve au sein du RTC n'en constitue pas le constituant majoritaire. Le RTC est composé principalement par l'ARN génomique viral couvert de quelques 1500 à 2000 molécules de NCp7, et contient également d'autres protéines comme MA, CA, Vpr ou encore IN [127, 128, 129]. En fait, cette vision schématique est remise en question par des observations récentes montrant que le core de HIV-1 resterait intact dans le cytoplasme des cellules infectées jusqu'au moment de l'import nucléaire et que la décapsidation se produirait au niveau des pores nucléaires [130, 131, 132, 133, 134] (illustration Figure 1.15). La synthèse complète de l'ADN proviral nécessiterait en outre un certain degré de stabilité et de structure du core. Enfin, il faut mentionner que de nombreux systèmes cellulaires de restriction, tel que TRIM5 α , peuvent intervenir au cours de cette étape de transcription réverse [135, 136].

Le processus de synthèse de l'ADN viral peut être résumé par le schéma séquentiel suivant :

1. *La synthèse du brin ADN strong-stop*

Débuté par l'hybridation du tRNA^{Lys,3} sur la séquence *Primer Binding Site* (PBS) de l'ARN viral qui génère une séquence double brin ARN/ARN de 18 nucléotides reconnue par la RT. Ce complexe d'initiation permet de débiter la synthèse du brin ADN(-). Le site d'hybridation du tRNA^{Lys,3} est proche de l'extrémité 5' de l'ARN viral. Un brin d'ADN(-) de 100 à 150 bases dit ADN *strong-stop* (-) ((-)ssDNA), correspondant au domaine 5' RU₅ du génome, est généré et la synthèse d'ADN(-) ne peut se poursuivre que suite au premier saut de brin.

2. *Le premier saut de brin*

Le premier saut de brin a lieu suite, ou de manière concomitante, à la dégradation du brin d'ARN au sein du duplex ARN :(-)ssDNA par l'activité RNase H de la RT. Il permet le transfert du (-)ssDNA vers l'extrémité 3' terminale de l'ARN viral géno-

1.4 Le cycle réplcatif du VIH-1

mique. Ce transfert est rendu possible par la présence de séquences redondantes (R) de part et d'autre de l'ARN viral. La séquence TAR est impliquée dans ce transfert. Cette étape est facilitée par NCp7.

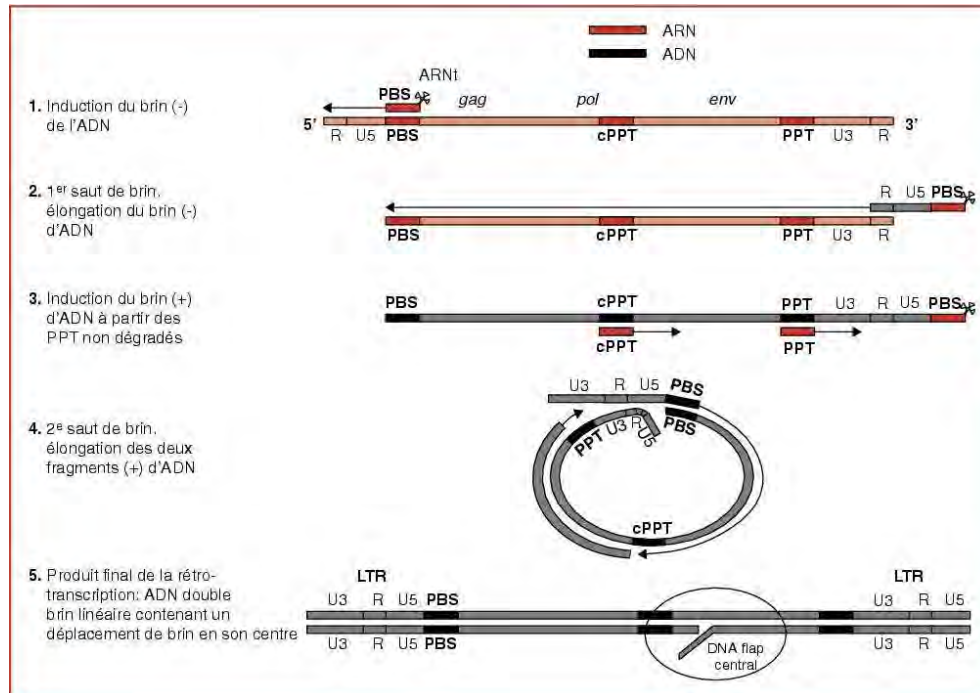


FIGURE 1.14 – Schéma des principales étapes de la transcription inverse

3. Synthèse complète du brin ADN(-)

Une fois le premier saut de brin effectué, la synthèse du brin ADN(-) peut se poursuivre, accompagnée par la digestion enzymatique de l'ARN par la RT. La dégradation du brin d'ARN n'est cependant pas complète. De courtes séquences PPT (Poly Purine Tract) résistent à la dégradation par la RNase H.

4. Synthèse du brin d'ADN(+) et le second saut de brin

Les duplex ADN(-)/PPT servent d'amorce pour le début de la synthèse du brin d'ADN(+) qui se poursuit jusqu'aux 18nt 3' terminaux du tRNA^{Lys,3}. Celui-ci est ensuite dégradé par l'activité RNase H de la RT. L'hybridation des séquences PBS(+) et PBS(-) constitue le second saut de brin. Cette étape permet la synthèse complète de l'ADN(+) avec *in fine* l'obtention de l'ADN proviral double brin. Tout comme le premier saut de brin, cette étape est stimulée par la NCp7. Il est à noter également

1.4 Le cycle répliatif du VIH-1

qu'en raison de l'initiation et de la terminaison centrale de la synthèse de l'ADN(+), un *flap* d'ADN simple brin subsiste au centre de l'ADN proviral.

Transport intranucléaire de l'ADN proviral

Le core enfermant le RTC est dirigé rapidement vers le noyau par l'intermédiaire des microtubules [132, 137]. Les lentivirus se distinguent des autres rétrovirus par leur capacité à infecter des cellules quiescentes. Ceci est rendu possible par un transport actif du matériel génétique au sein du noyau à travers les pores de l'enveloppe nucléaire. L'achèvement de la transcription donne naissance au complexe de préintégration (ou PIC pour *Pre-Integration Complex*) qui comprend l'ADN proviral néosynthétisé, différentes protéines virales, dont l'intégrase, la MA, la RT, Vpr et des protéines cellulaires HMG-I(Y) (*High Mobility Group*) [138], BAF (*Barrier to Autointegration Factor*) et LEDGF/p75 [139, 71]. Le mécanisme d'import implique probablement le système formé par les importines α et β et les protéines Ran-GTP et nécessite les signaux de localisation nucléaire portés par la protéine de la matrice, et Vpr [140]. Le mécanisme pourrait être comparable à celui décrit pour l'import dans le noyau de la protéine Rev. Il semblerait que la présence du *flap* d'ADN soit un déterminant majeur de ce mécanisme (Figure 1.15), puisqu'il constituerait le signal déclencheur de la décapsulation, permettant la formation du PIC [141, 131, 142].

Intégration de l'ADN proviral

Une fois dans le noyau, l'ADN proviral est intégré dans le génome de la cellule hôte (Figure 1.15, étape 10). Cette réaction est catalysée par l'intégrase virale qui reconnaît spécifiquement les deux LTR aux extrémités de l'ADN viral. L'intégrase clive un dinucléotide (GT) à chaque extrémité 3', libérant un dinucléotide CA fortement conservé et porteur d'une extrémité 3'-hydroxyl (Figure 1.16). C'est le *3' processing*. L'intégrase catalyse l'insertion de l'extrémité CA-OH 3' dans le génome cellulaire par une réaction de trans-estérification (*strand transfer*) [143]. Les enzymes cellulaires de maintien du génome réparent ensuite les discontinuités obtenues au niveau des sites de ligation cellulaires [144]. L'intégrase à elle seule permet de catalyser cette étape. Cependant, d'autres protéines, telles que la NCp7 [145], HMG-I(Y) [146] et LEDGF [147] en augmentent l'efficacité. Les sites d'intégration

1.4 Le cycle réplcatif du VIH-1

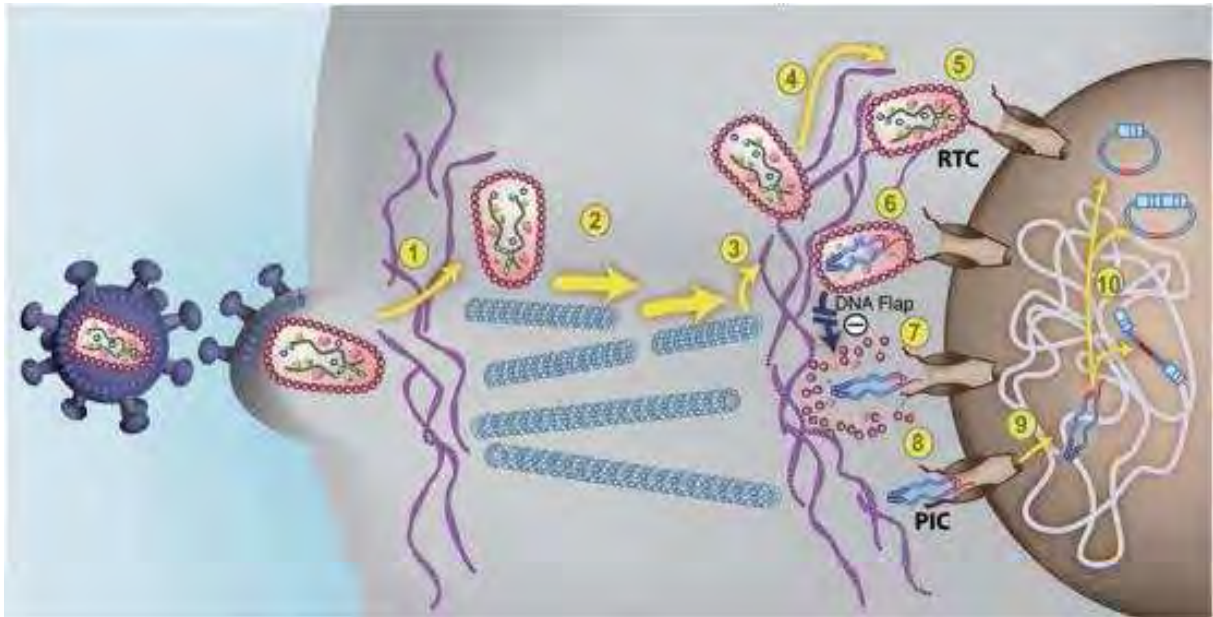


FIGURE 1.15 – Transport intracellulaire et import nucléaire du VIH-1

1. Entrée du VIH-1 et passage au sein du cytosquelette d'actine
2. Le core du VIH-1 est dirigé vers le noyau via les microtubules
- 3-6. Le core s'attache aux complexes des pores nucléaires
7. Décapsulation *DNA-flap* dépendante pour former le complexe de pré-intégration
- 8-9. Import nucléaire via les pores nucléaires
10. Intégration de l'ADN proviral à la chromatine de la cellule hôte ou circularisation

de l'ADN proviral sont multiples et la plupart des cellules infectées possèdent plus d'un provirus intégré [148]. Des sites préférentiels d'intégration (« hotspots ») pourraient exister dans les gènes qui sont activés à la suite d'une infection cellulaire par le VIH-1 [149]. Suite à l'étape d'intégration, il existe une étape de latence d'une durée variable et dont les mécanismes demeurent mal connus.

1.4.2 La phase post-intégrative

Cette phase débute par la synthèse d'ARN messagers viraux et se termine lors du bourgeonnement et de la maturation du virus.

1.4 Le cycle réplcatif du VIH-1

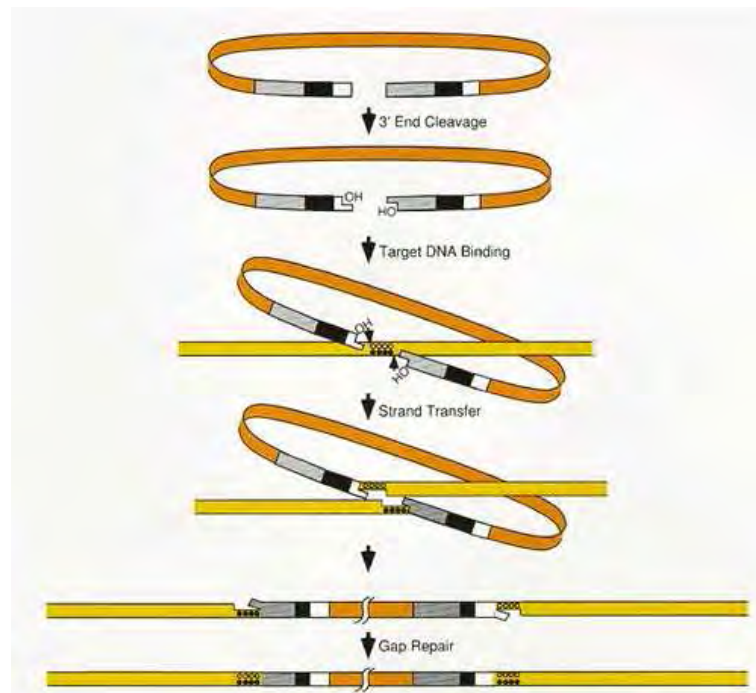


FIGURE 1.16 – Intégration de l'ADN viral dans le génome de la cellule

Transcription, export nucléaire et traduction

La transcription de l'ADN viral intégré dans le génome est la première étape vers la synthèse des protéines virales. L'expression du génome viral se déroule en deux étapes : une phase précoce au cours de laquelle sont exprimées des protéines régulatrices, suivie d'une phase tardive où les protéines structurales et accessoires seront traduites.

Phase précoce

Initiation de la transcription

Le VIH détourne la machinerie transcriptionnelle/traductionnelle de la cellule infectée. La transcription est assurée par l'ARN polymérase II. L'ADN proviral intégré fonctionne comme un gène eucaryote qui possède des promoteurs transcriptionnels favorisant le positionnement de l'ARN pol II au niveau du site d'initiation de la transcription. La faible processivité de l'ARN pol II entraîne la synthèse de transcrits courts, non polyadénylés. Une quantité minimale d'ARNv est transcrite, multi-épissée et traduite par la machinerie cellulaire permettant la synthèse des protéines régulatrices Tat, Rev et Nef. La protéine Tat

1.4 Le cycle réplcatif du VIH-1

active la transcription en se fixant sur la boucle interne de la séquence en tige boucle TAR en 5' de l'ARN viral en cours de transcription [150, 151, 152] (Figure 1.17 : étapes 1-2). La fixation de Tat permet de recruter différents facteurs comme la TATA-binding protéine (TBP), la cycline T1, la kinase 9 cyclin-dépendente (CDK9) et P-TEFb [153, 154, 155, 156], un facteur positif d'élongation de la transcription. Ce dernier hyperphosphoryle le domaine Cterm de l'ARN pol II, ce qui a pour effet d'augmenter la processivité de l'enzyme (Figure 1.17 : étape 3). La dissociation du complexe Tat-P-TEFb-TAR se fait après acétylation de Tat. L'interaction de Tat acétylée avec la protéine p300/CREB binding protein-associated factor (PCAF) (Figure 1.17 : étape 4), permet la synthèse de transcrits longs polyadénylés. Ces nouveaux transcrits codent pour des protéines de structure, des enzymes, des protéines auxiliaires mais également pour l'ARN génomique pouvant être intégré dans les nouveaux virions.

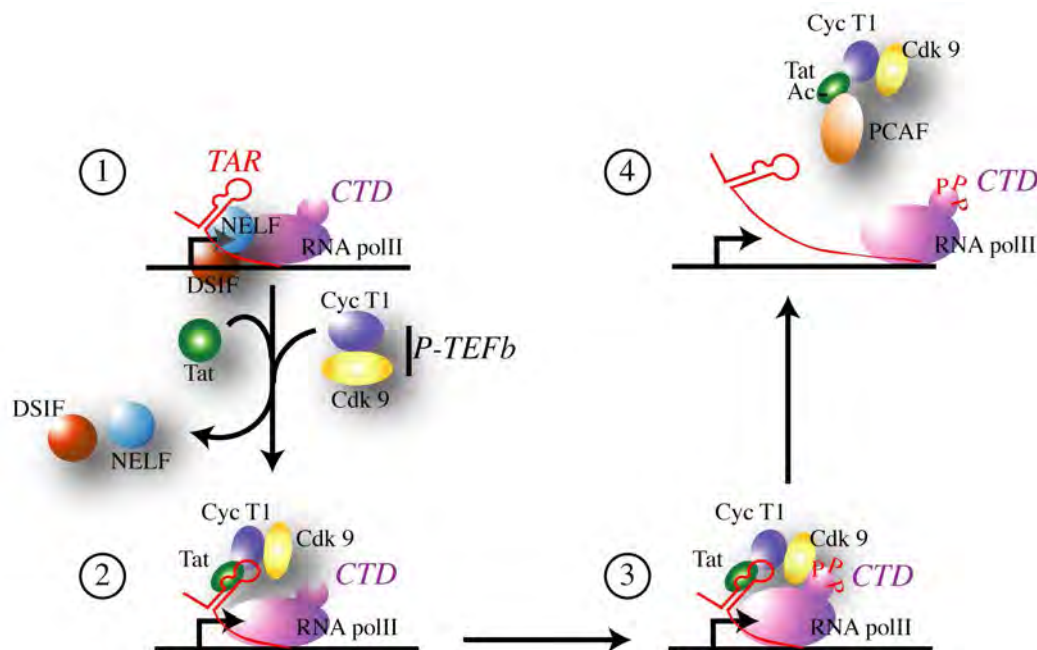


FIGURE 1.17 – Modèle de l'action trans-activatrice de Tat [7].

1-2. Fixation de Tat et recrutement de différents facteurs

3. Hyperphosphorylation du domaine Cterm de l'ARN pol II qui augmente la processivité de l'enzyme

4. Dissociation du complexe Tat-P-TEFb-TAR suite à l'acétylation de Tat

Epissage des ARNm viraux

À partir des transcrits longs, un grand nombre de d'ARNm viraux différents sont générés

1.4 Le cycle réplcatif du VIH-1

par un mécanisme d'épissage alternatif [157]. Les ARNm non-épissés codent pour Gag et Pol ou sont encapsidés, les ARNm mono-épissés (4 kb) codent pour Vif, Env, Vpr et Vpu, et enfin, les ARN multi-épissés (2.4 kb) codent pour Tat, Rev et Nef.

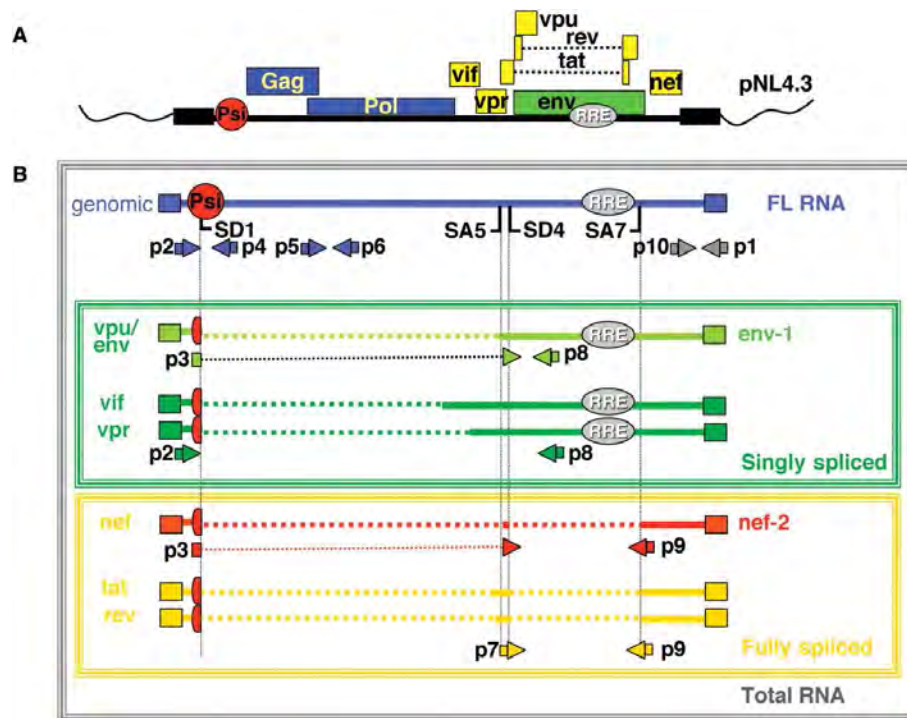


FIGURE 1.18 – Expression du génome de HIV-1.

A. ADN proviral du VIH-1 (pNL4.3).

Les rectangles noirs correspondent aux LTRs. Le signal d'encapsidation (*Psi*) est en rouge et le domaine RRE en gris

B. Représentation des principaux ARNm.

Phase tardive

Export nucléaire des ARNm

La régulation de la phase tardive se fait notamment par la protéine Rev. En effet, alors que les ARNm multi-épissés peuvent sortir librement du compartiment nucléaire, les ARNm non-épissés ou mono-épissés nécessitent l'implication de la protéine Rev pour l'export nucléaire [158]. La protéine Rev contient un domaine de liaison à l'ARN riche en arginines qui lie spécifiquement la séquence RRE (*Rev Responsive Element*) localisée dans le gène *env* [18] des ARNm non ou mono-épissés. Ce motif riche en arginines possède un signal NLS (*Nuclear Localization Signal*) ainsi qu'un domaine NES (*Nuclear Export Signal*). La

1.4 Le cycle réplcatif du VIH-1

protéine Rev agit au moyen d'un cycle répété d'import-export [100], permettant à chaque cycle, l'export nucléaire des ARNm (Figure 1.19). La multimérisation de Rev sur le domaine RRE masque le signal NLS et dévoile le NES, ce qui permet au multimère de Rev d'interagir avec des facteurs d'exports nucléaires (comme *Chromosome Maintenance gene1* (CRM1) et Ran GTP/GDP) et d'adresser les ARNm vers le cytoplasme de la cellule où ils pourront être traduits.

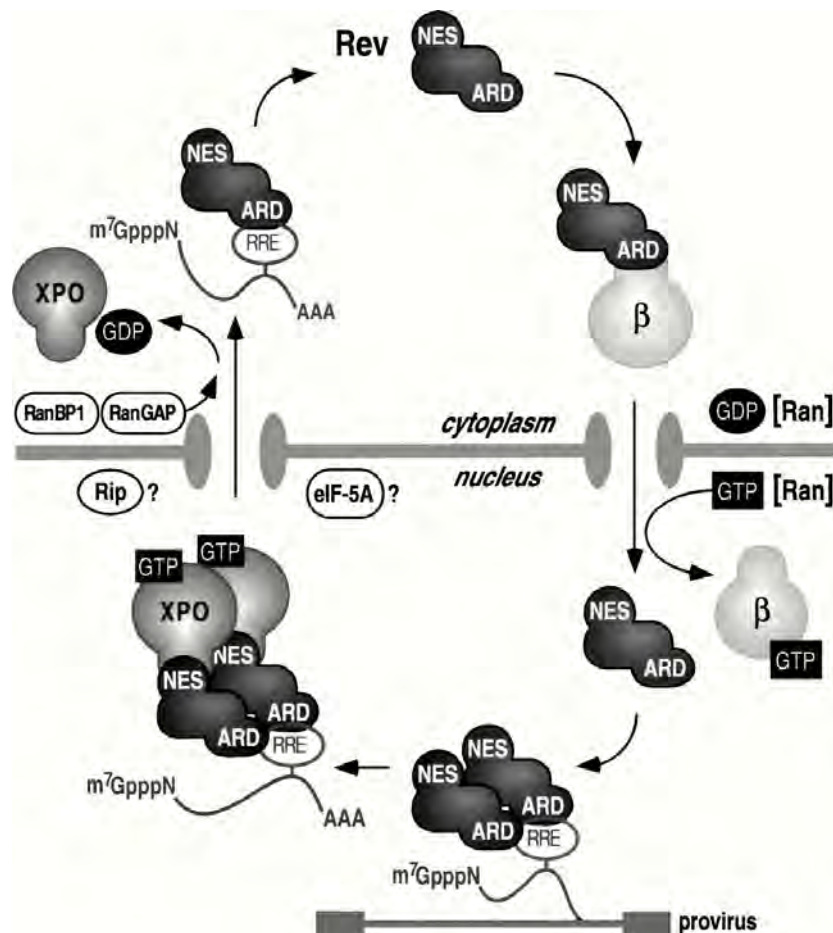


FIGURE 1.19 – Export nucléaire de l'ARNm viral au moyen de cycles import - export répétés de la protéine Rev.

Dans le cytoplasme, Rev s'associe aux importines et à Ran GDP pour être transloqué dans le noyau. Une fois dans le noyau, Rev lie l'ARNm par la séquence RRE, se multimérise et s'associe à l'exportine liée au GTP (XPO), pour être exportée vers le cytoplasme. Le signal d'export nucléaire (NES) de Rev permet l'interaction avec XPO. Dans le cytoplasme, le GTP va être hydrolysé, ce qui permet le relargage de l'ARNm

Traduction des ARNm

La traduction de ces ARNm dépend uniquement de la machinerie cellulaire. La synthèse des différentes protéines résulte de l'existence de différents transcrits d'ARNm. Dans le

1.4 Le cycle répliatif du VIH-1

cytoplasme, les ARNm mono-épissés sont traduits en précurseurs polyprotéiques gp160 à l'origine des glycoprotéines gp120 et gp41 [159], tandis que les ARNm non-épissés aboutissent à la production des précurseurs polyprotéiques Pr55Gag et Pr160Gag-Pol qui sont à l'origine de toutes les autres protéines structurales, fonctionnelles et enzymatiques du virus. Un mécanisme de type IRES-dépendant (*Internal Ribosome Entry Sites*) est également envisageable. Celui-ci, en permettant l'initiation de la traduction de manière interne, permettrait de contourner le fait que le codon AUG d'initiation de la traduction puisse être masqué suite à la liaison de molécules de Gag néo-synthétisées aux 5'-UTR des ARNm non épissés. Des éléments IRES ont été localisés dans le 5'-UTR et dans les régions codantes de l'ARN [160, 161].

Assemblage et relargage des virions

La multimérisation du précurseur de la polyprotéine Gag, Pr55^{Gag}, constitue l'événement central de la formation de particules du VIH-1. En effet, Pr55^{Gag} permet à lui seul de produire des particules de type viral (*virus-like particles* ou VLP) [162], même si la présence de l'ARN génomique, des enzymes codées par *pol* ainsi que des glycoprotéines d'enveloppe sont nécessaires à la production de virus infectieux [35]. Par l'intermédiaire d'interactions protéine-protéine, protéine-ARN, et protéine-lipide, Pr55^{Gag} assure l'orchestration de l'incorporation de chacun des composants majeurs constituant les particules virales, selon un modèle guidé par une double plateforme avec d'un côté l'ARN viral et de l'autre, les membranes cellulaires. En effet, l'ARN génomique du VIH-1 est un transcrit cappé et polyadénylé qui est sélectivement incorporé dans les particules virales sous forme de dimère. Son encapsidation dans les particules virales dépend de l'interaction de l'ARNg avec le domaine NC du précurseur Gag [163, 164]. La NC reconnaît la séquence d'encapsidation Ψ , localisée entre le site donneur d'épissage et le codon d'initiation du gène *gag*, ce qui favorise l'incorporation de l'ARN viral non épissé [165]. Il est en effet nécessaire que l'encapsidation de l'ARN génomique dans la particule rétrovirale soit un processus hautement spécifique qui permette la sélection d'un type unique d'ARN au milieu des ARNm totaux, cappés et polyadénylés présents dans la cellule infectée [166, 167]. La capture de l'ARN génomique se fait au niveau périnucléaire/centrosomale [168]. L'ARNg constitue un échaffaudage (*scaffold*)

1.4 Le cycle répliatif du VIH-1

pour l'assemblage du VIH-1 [14]. En effet, la fixation concomitante de nombreux précurseurs sur l'ARN favoriserait la multimérisation de Pr55^{Gag}, étape nécessaire à l'assemblage des particules [169]. D'autre part, ce processus inhiberait la traduction de l'ARN en stimulant sa dimérisation, ce qui permettrait son incorporation dans la particule virale [170]. De plus, l'acheminement vers les membranes des complexes ARN-Gag est assurée par la région MA de Pr55^{Gag}, qui dirige Pr55^{Gag} vers les rafts ou les endosomes tardifs incluant les MVB (*multivesicular bodies*). La partie N_{term} myristoylée de la matrice, associée à un cluster de résidus basiques permet l'ancrage de Pr55^{Gag} aux membranes cellulaires. Sur la base d'une homologie du processus de bourgeonnement viral avec la biogenèse des endosomes, il a été proposé que le bourgeonnement des particules infectieuses puisse avoir lieu au niveau des endosomes tardifs [171, 172, 173, 174, 175]. Des observations récentes tendent à montrer que l'assemblage est ciblé à la membrane plasmique et que la présence de particules virales dans les endosomes pourrait résulter de l'endocytose de bourgeons déficients [176, 177, 178, 179, 180], ou même que les compartiments associés aux endosomes tardifs pourraient correspondre à des invaginations de la membrane plasmique [181, 182]. Cette question reste donc très débattue.

La formation des particules infectieuses requiert ensuite le recrutement de Gag-Pol [183, 184], Env [35, 185], Vpr, Vif et Nef, ainsi que des protéines cellulaires [23, 186, 187]. Le recrutement de Gag-Pol, précurseur de la protéase, de la RT et de l'IN, se fait obligatoirement via son interaction avec Gag [188], par le domaine capsidique de Pr55^{Gag} [189], tandis que les protéines d'enveloppe sont recrutées par le domaine MA de Gag [190]. À noter également que le virion encapside des acides nucléiques cellulaires comme le tRNA^{Lys3} [191, 192]. Puis les virions assemblés quittent la cellule par bourgeonnement. Pour ce, Pr55^{Gag} recrute, *via* son domaine p6, la machinerie cellulaire ESCRT (*endosomal sorting complex required for transport*) [193]. Le bourgeonnement a lieu depuis la membrane plasmique et/ou par fusion membranaire des compartiments intracellulaires, vraisemblablement au niveau de radeaux lipidiques [194] ou de microdomaines membranaires enrichis en tétraspanines [195, 196]. Le schéma 1.20 résume les étapes tardives du cycle répliatif du VIH-1.

1.4 Le cycle réplcatif du VIH-1

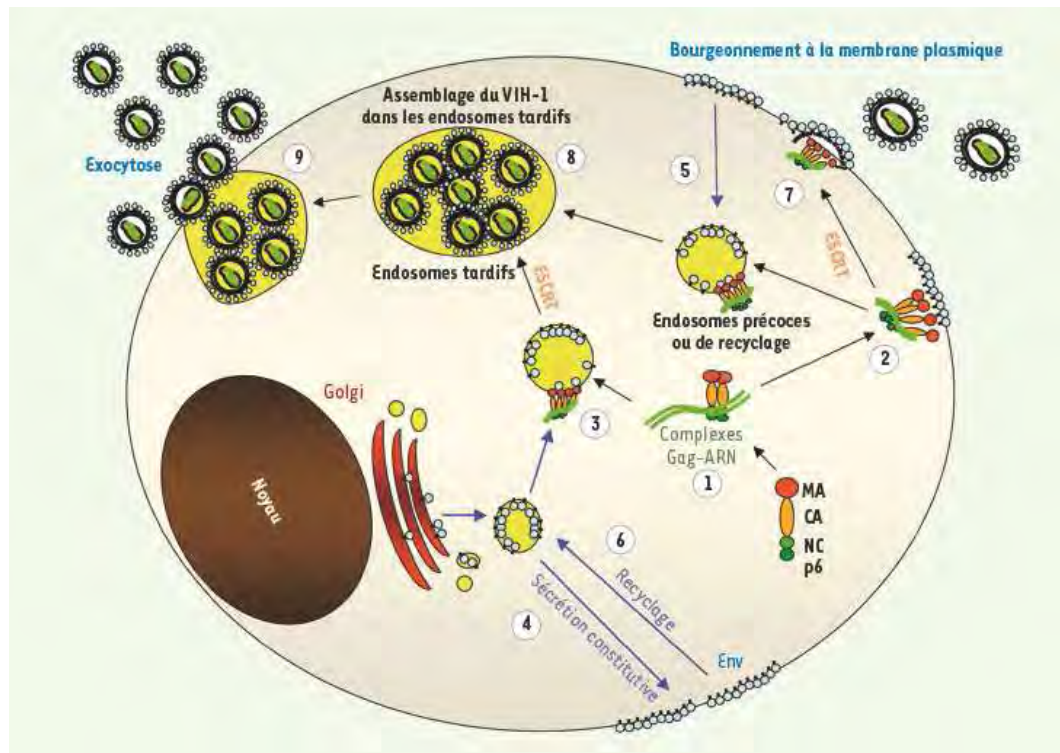


FIGURE 1.20 – Trafic et voies d'assemblage de VIH-1.

1. Synthèse et multimérisation de Pr55^{Gag} sur l'ARNv

2-3. Adressage des complexes Gag-ARN vers la membrane plasmique ou vers les compartiments endosomaux

4-6. Sécrétion et recyclage de l'enveloppe qui est ensuite recrutée par Gag au sein de la particule virale en formation

7-8. Bourgeonnement des virions stimulé par la machinerie cellulaire ESCRT

Maturation des virions

Une fois que le virion est libéré de la cellule hôte, le virus subit une étape de maturation [35]. Celle-ci correspond à une modification architecturale résultant du clivage du précurseur Pr55^{Gag}. Ces étapes de clivage sont assurées par la protéase codée par Pr160^{Gag-Pol}. Consécutivement au recrutement de Pr160^{Gag-Pol} lors de l'assemblage, la concentration intra-virion de ce précurseur devient suffisante pour permettre sa dimérisation, ce qui active la protéase et permet à celle-ci de s'auto-cliver au sein du virion libéré. Ces étapes de maturation permettent l'obtention d'une particule infectieuse.

1.5 Physiopathologie, histoire naturelle et aspects cliniques de l'infection à VIH-1

L'infection par le VIH évolue en plusieurs phases pouvant se succéder dans le temps :

Primo-infection (ou phase aiguë)

Première phase de l'infection après la contamination, au cours de laquelle le virus se réplique activement (on estime la production de virions entre 1 et 10 milliards par jour). La phase aiguë peut s'accompagner de symptômes aspécifiques, survenant dans les 1 à 8 semaines après la contamination, à titre de fièvre, myalgie, arthralgie, exanthème, dysphagie douloureuse, diarrhée, adénopathies ou encore ulérations buccales ou génitales. Des manifestations neurologiques aiguës, telles que méningite, encéphalite, paralysie faciale, myélopathie ou neuropathie périphérique peuvent apparaître. La sévérité clinique de la phase aiguë et la présence de signes neurologiques sont associés à un risque d'évolution accéléré de la maladie. Le diagnostic de cette phase est essentiel dans le but d'une prise en charge précoce et de prévention de la transmission. Du point de vue biologique, on observe une leucopénie puis un syndrome nucléosique avec une augmentation modérée des transaminases dans 50% des cas. La charge virale plasmatique (CV) est souvent très élevée puis celle-ci diminue ensuite spontanément pour atteindre un plateau d'équilibre en 3 à 6 mois (Figure 1.21). Les anti-corps anti-VIH spécifiques apparaissent de façon différée avec une séroconversion (positivité de la PCR) qui survient dans plus de 90% des cas dans les 2 à 8 semaines suivant la contamination, quel qu'en soit le mode.

Phase chronique pré-sida

Cette phase d'infection chronique cliniquement latente mais biologiquement active conduit au SIDA avec une médiane de 10 ans, même s'il existe de très grandes variations inter-individuelles. La réplication virale est constante, en particulier dans les organes lymphoïdes. Le système immunitaire, activé par la présence du virus, compense partiellement la destruction des lymphocytes CD4. Un syndrome de lymphadénopathie généralisée persistante peut s'observer chez les patients sans symptômes fonctionnels. Cela se traduit par des adénopathies symétriques cervicales, axillaires,

1.5 L'infection à VIH

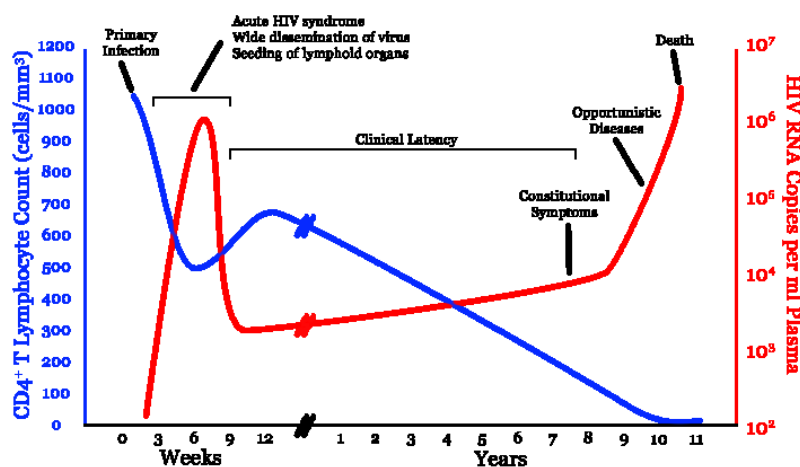


FIGURE 1.21 – Représentation schématique de l'évolution de la charge virale (rouge) et du taux de LT CD4 (bleu) au cours de l'infection au VIH-1.

sous-maxillaires ou occipitales. Histologiquement, on retrouve une hyperplasie folliculaire bénigne, non spécifique. Cette forme clinique n'a pas de valeur pronostique négative. Parallèlement, des manifestations cliniques ne répondant pas à la définition du stade SIDA témoignent de l'atteinte du système immunitaire (à titre de manifestations cutanéomuqueuses d'origines fongique ou virale, manifestations hématologiques avec thrombopénie, anémie ou leucopénie le plus souvent asymptomatiques ou encore manifestations générales avec altération de l'état général, fièvre persistante modérée, sueurs nocturnes abondantes ou encore diarrhée se prolongeant au delà d'un mois).

Phase SIDA

Le SIDA correspond à un ensemble de manifestations infectieuses opportunistes et/ou tumorales caractérisées, liées à la déplétion profonde de l'immunité cellulaire. Le tableau 1.1 présente une liste non exhaustive des infections opportunistes les plus fréquentes au cours du SIDA. Parmi les affections tumorales (ou onco-tumorales), on trouve en bonne place le sarcome de Kaposi (consécutif à une infection par le virus HHV-8) et les lymphomes non hodgkiniens (LMNH) (dont le lymphome de Burkitt). L'introduction de prophylaxies primaires, particulièrement contre la pneumocystose et la toxoplasmose, ainsi que l'augmentation de l'efficacité des multi-thérapies a permis de réduire très fortement la prévalence des infections opportunistes et tumorales. Celles-ci ne s'observent quasiment plus que dans deux cas de figure : chez les patients

1.5 L'infection à VIH

| | Agent | Localisations préférentielles |
|--------------------|--|--|
| Parasites | <i>Pneumocystis carinii</i> <i>Toxoplasma gondii</i> <i>Cryptosporidium</i> <i>Isospora belli</i> Microsporidies | Poumons SNC, rétines, poumons TD, voies biliaires TD TD, urine, sinus |
| Champignons | <i>Candida sp</i> <i>Cryptococcus neoformans gondii</i> <i>Histoplasma capsulatum</i> <i>Aspergillus</i> | Bouche, Oesophage SNC, poumons ou disséminée Disséminée Poumons ou disséminée |
| Bactéries | <i>Mycobacterium avium intracellulare</i> <i>Mycobacterium tuberculosis</i> <i>Salmonella non Typhi</i> | Sang, ganglions, TD, disséminée Poumons, ganglions ou disséminée Bactériémie |
| Virus | CMV Herpes simplex VZV Papovavirus (Polyomavirus et Papillomavirus) | Rétines, TD, SNC, poumons Peau, muqueuses, poumons, TD Peau, système nerveux Peau, muqueuses, poumons, TD |

TABLE 1.1 – Infections opportunistes les plus fréquentes au cours du SIDA.
TD : Tube Digestif; SNC : Système Nerveux Central

en échec thérapeutique avec des CD4 inférieurs à 100mm^{-3} , d'une part, et en tant qu'infection inaugurale du stade SIDA chez des patients s'ignorant VIH séropositifs.

1.6 Traitements anti-rétroviraux

S'il est des domaines thérapeutiques en perpétuelle évolution, la prise en charge des personnes infectées par le VIH en constitue un bon exemple. L'évolution des souches virales, la mise à la disposition régulière de nouvelles molécules, voire de nouvelle classe d'anti-rétroviraux ou encore la multiplicité des essais thérapeutiques dans le domaine font que les recommandations des groupes d'experts évoluent extrêmement rapidement. Aussi, et puisque vraisemblablement caduques d'ici quelques mois, les dernières recommandations quant à la prise en charge de ces patients ne seront pas détaillées ici (ces dernières sont décrites dans le [rapport Yéni 2008](#), avec addendum 2009, disponible sur le site internet du ministère de la Santé et des Sports). Les propriétés pharmacologiques des différentes classes de molécules ne seront pas non plus abordées. Un survol rapide des différentes classes de médicaments disponibles (AMM au 01/06/2010) (ou à venir) est proposé.

Les traitements actuels

L'objectif principal des traitements antirétroviraux anti-VIH est de diminuer la morbi-mortalité de l'infection par le VIH en restaurant un nombre de lymphocytes CD₄ (LT4) supérieur à 500mm^{-3} . Ceci est rendu possible par une réduction maximale de la réplication virale (charge virale plasmatique < 50 copies/mL), permettant une meilleure restauration immunitaire et une limitation l'apparition de virus résistants. Outre l'efficacité et la tolérance, la qualité de vie des patients, facteur contribuant à l'observance des traitements, est dorénavant largement pris en compte. Les molécules qui sont présentées ici par code anatomique, thérapeutique et chimique (ATC) correspondent aux traitements virustatiques actuellement disponibles (ou en phase avancée d'essai thérapeutique) (voir figure 1.22 pour un aperçu).

- **Les inhibiteurs nucléosidiques et nucléotidiques de la reverse transcriptase (INRT ou INtRT) (ATC J05AF)**

Première classe d'anti-rétroviraux utilisée dès 1987 dans le VIH avec notamment l'AZT. Il s'agit d'analogues de bases naturelles inhibant l'activité de la RT par compétition. Cette classe comprend les molécules *zidovudine*, *didanosine*, *zalcitabine*,

1.6 Traitements ARV

stavudine, lamivudine, abacavir, emtricitabine, apricitabine (phase III), *elvucitrabine* (phase II)

- **Les inhibiteurs non-nucléosidiques de la réverse transcriptase (INNRT)** (ATC J05AG)

Les INNRTs ont un mode d'action différents des INRT ou INtRT puisqu'ils inhibent la RT en se fixant sur d'autres sites de liaison que le site catalytique. Ce sont des inhibiteurs non-compétitifs. On trouve les molécules *nevirapine, efavirenz, etravirine (TMC-125), delavirdine* (ATU nominative), *rilpivirine* ou *TMC-278* (phase III).

- **Les inhibiteurs de protéase** (ATC J05AE)

Suite à leur mise à disposition en 1996, cette classe de médicament a "révolutionné" les traitements anti-HIV en permettant l'instauration des tri-thérapies (puis des multi-thérapies) qui ont considérablement réduit l'évolution vers le stade SIDA et la mortalité. Ces molécules inhibent l'action de la protéase virale permettant le clivage des produits du gène *gag* et *pol*, indispensable à l'obtention de virus infectieux. Bien qu'extrêmement puissantes, ces molécules sont nécessairement utilisées en association car la monothérapie aboutit très rapidement à la sélection de résistances, souvent croisées à l'ensemble des anti-protéases. Malgré les co-administrations, l'apparition de résistances est le facteur majeur d'échec. L'observance à moyen et long terme est souvent problématique tant par les contraintes des schémas de prise (à jeun, nombre de comprimés importants, ...) que par les effets indésirables engendrés (lipodystrophies, intolérance au glucose, hypertriglycémie, ...). Certains médicaments de cette classe comme le ritonavir, puissant inhibiteur du métabolisme induit par le CYP3A, sont utilisés comme potentialisateurs pharmacocinétiques d'autres IP co-administrés.

Les IP disponibles sont : *saquinavir, indinavir, ritonavir, nelfinavir, amprénavir, lopinavir, fosamprenavir, atazanavir, tipranavir, darunavir*

- **L'inhibiteur de fusion** (ATC J05AX)

L'*enfuvirtide*, encore connu sous le nom de T20. Ce peptide de 36 AA (nécessitant deux administrations SC quotidienne) est un inhibiteur du réarrangement structural de la gp41, qui agit en se liant spécifiquement à cette protéine dans le milieu extracellulaire, bloquant la fusion de la membrane virale avec la membrane de la cellule cible.

1.6 Traitements ARV

Indiqué chez des patients en échec thérapeutique, généralement en troisième ligne, en association systématique avec d'autres antirétroviraux. Cette classe sera probablement prochainement groupé à celle des inhibiteurs d'entrée.

– Les inhibiteurs d'entrée (ATC J05AX)

Le *maraviroc* est prescrit en association avec d'autres anti-rétroviraux dans le traitement de l'infection par le VIH-1 à tropisme uniquement CCR5. C'est un antagoniste des récepteurs CCR5, empêchant le VIH-1 à tropisme CCR5 de pénétrer dans la cellule. Le *maraviroc* ne possède pas d'activité antivirale *in vitro* contre les virus à tropisme CXCR4 ou à tropisme double CCR5/CXCR4.

Le *vicriviroc* : antagoniste non compétitif allostérique des récepteurs CCR5, en phase III.

Le *TBR-652* : antagoniste des récepteurs CCR5 et CCR2. Le CCR2 est un co-récepteur présent chez les monocytes, les cellules dendritiques et les lymphocytes T mémoires (en phase II).

Le *PRO 140* anticorps monoclonal humanisé anti-CCR5 (en phase II).

– Les inhibiteurs de l'intégrase (ATC J05AX)

Le *raltégravir* est un inhibiteur de transfert de brin qui bloque l'intégration de l'ADN proviral dans le génome de la cellule infectée.

L'*elvitegravir* est en phase III active, le *GSK1265744* en phase II.

L'espoir d'un vaccin ?

Il est inconcevable d'évoquer les traitements anti-VIH sans envisager la possibilité d'un vaccin, tant l'attente est grande et le projet mobilisateur. Néanmoins, il faut reconnaître que l'approche vaccinale anti-VIH-1 n'a, malgré des efforts considérables, fourni que des résultats très décevants. Les quatre essais cliniques d'envergure cherchant à démontrer l'efficacité de vaccins anti-HIV conduits jusqu'à présent, à savoir AIDSVAX 003 et AIDSVAX 004 de *VaxGen Inc* et HVTN 502 et HVTN 503 du *NIH-supported HIV Vaccine Trials Network* ont été des échecs, à tel point que le développement même d'un vaccin reposant sur des anticorps circulants neutralisants a été remis en question [197]. Les approches alternatives, reposant sur un vaccin dit

1.6 Traitements ARV

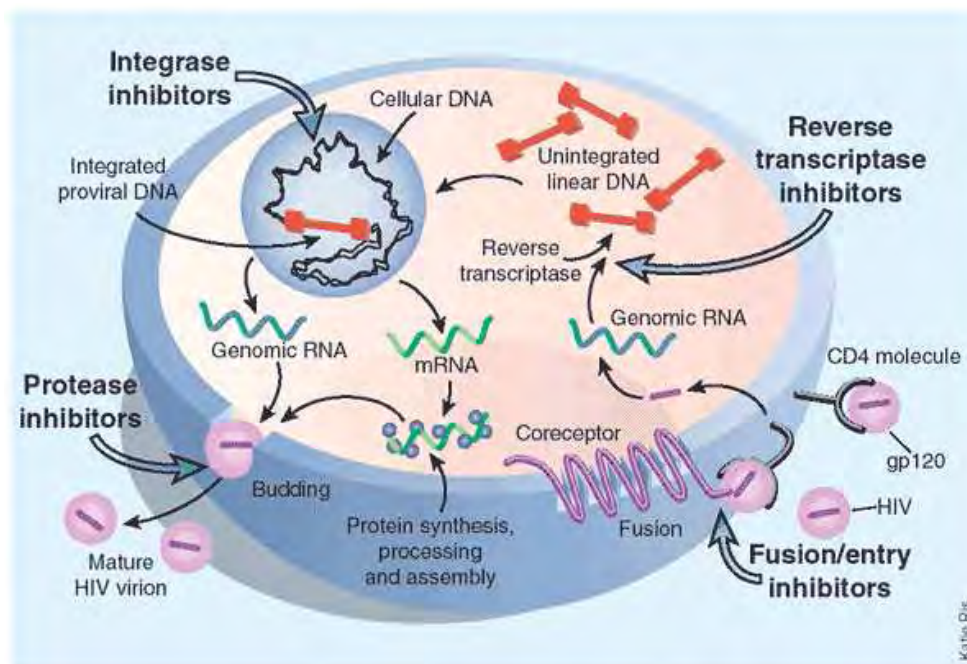


FIGURE 1.22 – Représentation des cibles des traitements anti-VIH actuels.

”thérapeutique”, dont l’objet est de stimuler les défenses immunitaires, ne se sont pas montré plus prometteuses. Plus récemment, un essai conduit en Thaïlande (RV144 trial) [198] utilisant une combinaison ALVAC-HIV et AIDSVAX B/E *versus* placebo a été présenté comme démontrant un faible niveau d’efficacité (malgré une méthodologie contestable puisque seule l’approche par *itt* modifiée (*Intent to treat*) s’est montrée significative, alors qu’à la fois l’approche par *itt* et l’approche *per protocole*, qui constituent les approches classiques d’analyse statistique de ce genre d’étude, n’ont pas montré de bénéfice associé au traitement), illustrant bien la volonté de maintenir l’espoir de pouvoir développer un vaccin anti-VIH.

1.7 La protéine de la nucléocapside NCp7

1.7.1 NCp7 : un produit de clivage

Chez les rétrovirus, la NC est synthétisée comme un domaine fonctionnel du précurseur Gag, $Pr55^{Gag}$ dans le cas du VIH-1. La forme mature de la NC est générée par une série de clivages protéiques discrets médiés par la protéase virale. Chez les *orthoretroviridae*, trois protéines principales sont libérées : la matrice (MA), la capside (CA), et la NC, ainsi que différents peptides supplémentaires, selon le type de rétrovirus [63].

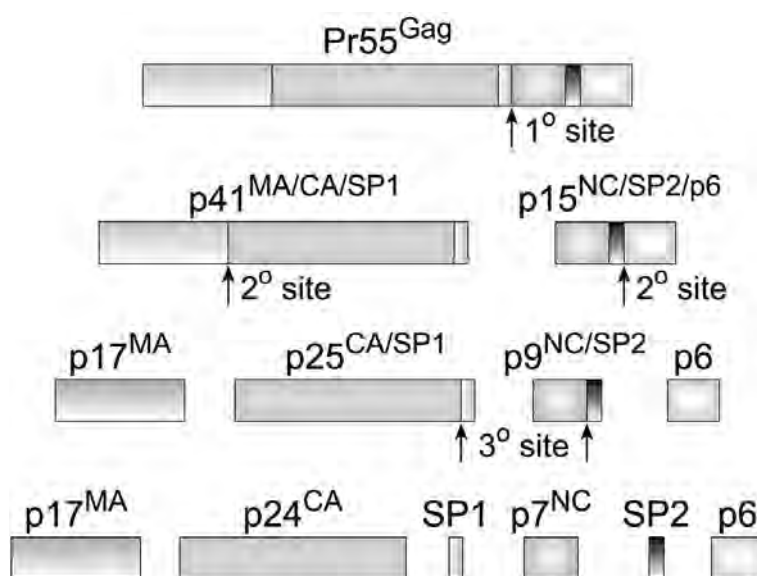


FIGURE 1.23 – Clivage protéolytique de $Pr55^{Gag}$ chez VIH-1. [199]

Le clivage initial par la PR de $Pr55^{Gag}$, entre SP1 et le domaine NC, résulte en la formation de la NCp15, un peptide comprenant la NCp7, les domaines p1 (SP2), et p6 [200]. La protéine NCp15 doit être liée à l'ARN pour pouvoir être ultérieurement clivée par la protéase [201, 202]. Le clivage de NCp15 conduit à NCp9 (NCp7 avec le peptide SP2 encore attaché). Le dernier clivage permet la libération de NCp7. Ainsi, au cours de la protéolyse, la NC existe sous deux formes intermédiaires, NCp15 et NCp9, avant d'adopter sa forme finale de 55 acides aminés, la forme prédominante détectée dans les particules matures du VIH-1. [203, 204, 205]. Il est important de noter que dans certaines études,

1.7 NCp7

la NCp7 est désignée comme possédant soit 55, soit 72 acides aminés [206, 207, 208]. On préférera parler ici de NCp9 pour la forme (1-72)NCp7. En outre, il semblerait que la forme native de la NCp9 ne comprenne que 71 acides aminés [204]. La protéolyse de Gag est rapide et efficace, de telle sorte que les intermédiaires NCp9 et NCp15 ne sont, habituellement, pas détectés par Western blot dans les particules du VIH-1 nouvellement libérées [64], même si celles-ci ont déjà pu être observées [209, 205]. Ceci laisse supposer que de petites quantités de NCp9 ou de NCp15 peuvent exister dans les virus matures. Toutes les formes de NC, de $Pr55^{Gag}$ à NCp7, présentent une activité chaperonne, bien qu'à des degrés divers [210].

1.7.2 Structure

La protéine de la NCp7 est une petite protéine caractérisée par des domaines N_{term} et C_{term} particulièrement riches en résidus basiques, flanquant deux motifs dits en doigt de zinc, de séquence $-Cys \cdot X_2 \cdot Cys \cdot X_4 \cdot His \cdot X_4 \cdot Cys-$ (où X désigne des acides aminés variables) [211]. Ces motifs CCHC sont capables de chélater le Zn^{2+} de manière stoechiométrique avec une forte affinité (aux alentours de $10^{-13}M$) [212, 213, 214, 215, 216]. Le cation Zn^{2+} est coordonné de façon tétraédrique par les groupements thiols des cystéines et amino de l'histidine [211, 217, 206, 218, 219] selon un mécanisme présentant plusieurs états intermédiaires [213]. La liaison du zinc est responsable de la structuration du doigt, permettant la transition d'une structure aléatoire vers une forme repliée, fortement contrainte, à l'origine de ses activités biologiques [206, 207, 219]. La complexation de ces motifs avec le zinc est un événement central dans la formation de la protéine active. De plus, la conservation du motif CCHC au sein des NC de différents rétrovirus témoigne de l'importance biologique de ces doigts de zinc. De nombreuses études ont montré que les mutations qui affectent les doigts de zinc ont une répercussion négative sur la réplication virale. La substitution de l' His_{23} par une Cys réduit l'affinité de la protéine pour l'ARNg et pour l'ARNt^{Lys3} conduisant à la production de particules virales non infectieuses [220, 205]. De plus, les deux doigts semblent jouer des rôles asymétriques [221, 222], et leur agencement spatial, notamment la proximité entre certains résidus des doigts de zinc semble être un facteur important pour l'activité biologique de la protéine [207]. Ceci est à relier avec le fait que la séquence *spacer* inter-digitale est très fortement conservée (Figure 1.24). Au sein de cette séquence *spacer*, le

1.7 NCp7

résidu Pro_{31} joue un rôle critique dans le rapprochement des doigts [223, 207, 224]. La for-

| | N-Terminus | N-terminal zinc finger | Central | C-terminal zinc finger | C-Terminus |
|-------------------|-------------------|------------------------|----------|------------------------|------------|
| M-Group Consensus | MQRGNFKGQKRILK | CFNCGKEGHIARNC | RAPRKRKG | CWKCGKEGHQMKDC | TERQAN |
| Consensus-A1 | -----R----- | -----L----- | ----- | ----- | ----- |
| Consensus-A2 | -----R----- | -----L----- | ----- | ----- | ----- |
| Consensus-B | -----RN-RKTV----- | -----K----- | ----- | ----- | ----- |
| Consensus-C | ---S---P---V--- | ----- | ----- | ----- | ----- |
| Consensus-D | -----PRK--- | -----K----- | ----- | ----- | ----- |
| Consensus-F1 | --KS---R---V--- | -----K----- | ----- | ---R----- | ----- |
| Consensus-G | --KS---PR-T--- | -----L----- | ----- | ----- | ----- |
| Consensus-H | --K---PRK-V--- | ----- | ----- | ---R----- | ----- |
| Consensus-K | -----RK--- | ----- | ----- | ----- | ----- |
| Consensus-01-AE | -----L----- | -----L----- | ----- | ----- | ----- |
| Consensus-02-AG | -----R--RT--- | -----L----- | K----- | ----- | ----- |
| Consensus-03-AB | --KS--R-P---L--- | -----D--L----- | ----- | ----- | ----- |
| Consensus-04-CPX | --KS---R--- | -----L----- | ----- | ----- | ----- |
| Consensus-06-CPX | --KS---P--S--- | -----L----- | ----- | ----- | ----- |
| Consensus-07-BC | ---S---S---V--- | ----- | ----- | ----- | ----- |
| Consensus-08-BC | ---S---S---V--- | -----K----- | ----- | ----- | ----- |
| Consensus-10-CD | -----P-K--- | -----K----- | ----- | ---R----- | ----- |
| Consensus-11-CPX | ---S-----L--- | -----L----- | ----- | ----- | ----- |
| Consensus-12-BF | --KS---R---V--- | -----K----- | ----- | ---R----- | ----- |
| Consensus-14-BG | --KS---PR-N--- | -----L----- | ----- | ----- | --SK--- |
| NL4-3 | I---RN--KTV--- | -----K----- | ----- | ----- | ----- |

FIGURE 1.24 – Alignement des séquences de NC du VIH-1 du groupe M. [199]

L'alignement des séquences consensus provient du *Los Alamos HIV Sequence Database* (<http://www.hiv.lanl.gov>)

mation et l'agencement des doigts de zinc confèrent à la protéine un repliement particulier permettant l'apparition d'un plateau hydrophobe à la surface des deux doigts impliquant les résidus Val_{13} , Phe_{16} , Thr_{24} , Ala_{25} et Trp_{37} [225, 226, 227, 228, 229, 17]. Parmi ces résidus, la Phe_{16} et le Trp_{37} jouent un rôle primordial puisqu'ils sont largement impliqués dans l'interaction de la protéine avec les acides nucléiques, en s'empilant avec les bases *via* des interactions hydrophobes [230, 231]. Une mutation de ces résidus provoque la formation de particules virales non infectieuses [232].

Les structures RMN de la NCp7 complexée avec les séquences SL2 [233] et SL3 [234] du VIH-1, ainsi que les structures de la NC(12-53) ou NC(12-55) complexées avec les séquences ACGCC et $\Delta P(-)$ PBS, respectivement, sont disponibles. Dans les complexes NCp7-

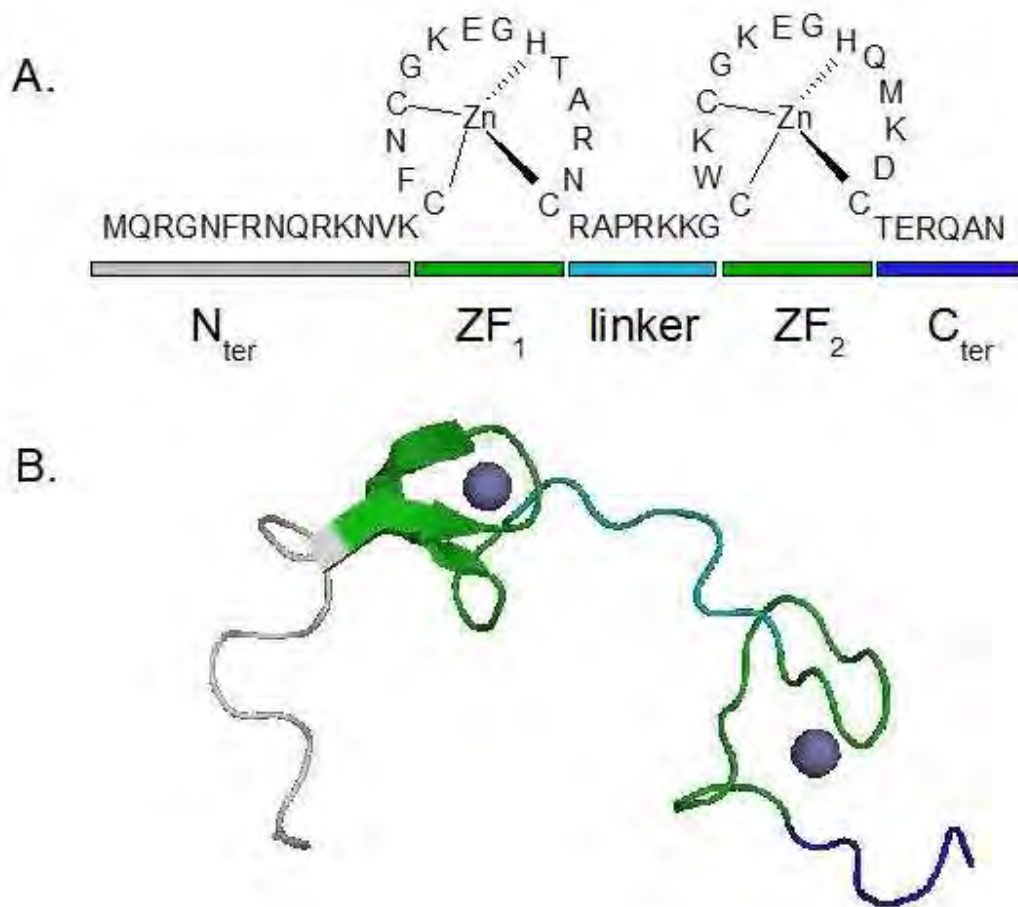


FIGURE 1.25 – Structure secondaire et tertiaire de la NCp7.

oligonucléotide, la proximité spatiale entre Phe16 et Trp37 est stabilisée. Le plateau hydrophobe est largement impliqué dans la liaison *via* des liaisons hydrophobes et des liaisons hydrogènes avec les bases de l'ADN ou de l'ARN. Il permet également d'orienter le résidu Trp. Les résidus basiques des chaînes latérales stabilisent la liaison par des interactions électrostatiques avec les groupements phosphates des oligonucléotides. Les structures disponibles montrent que les séquences ARN lient la NCp7 selon une orientation opposée par rapport aux séquences ADN. Ainsi pour SL2 et SL3, les deux extrémités de la protéine sont renversées par rapport à l'orientation de la boucle en comparaison de la structure de PBS (Figure 1.26). Il sera intéressant de voir si cette observation est généralisable à d'autres séquences d'acides nucléiques.

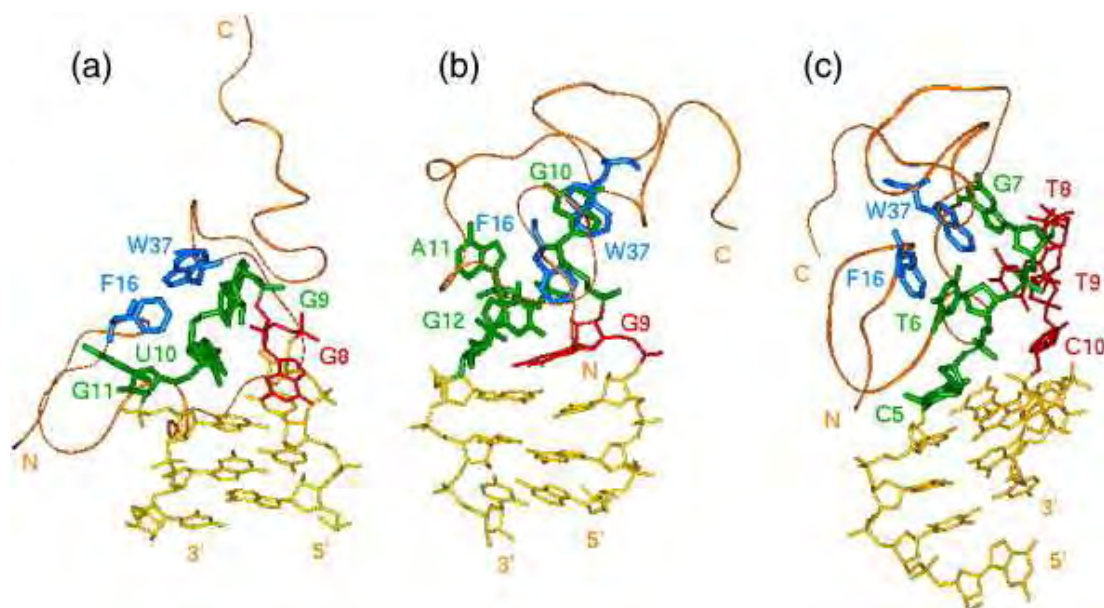


FIGURE 1.26 – Structures tridimensionnelles des complexes

(a) NCp7-SL2 ARN

(b) NCp7-SL3 ARN

(c) (12-53)NCp7- $\Delta P(-)$ PBS ADN [227]

Les oligonucleotides sont en jaune. Les bases impliquées dans l'interaction sont en vert, les autres bases de la boucle sont en rouge

1.7.3 L'activité chaperonne de NCp7

La notion de protéine chaperonne découle conceptuellement de la théorie du paysage énergétique (*foldings energy landscape*) proposée pour expliquer le repliement des protéines [235]. Ce paysage correspond à l'hyperespace des énergies libres d'une protéine associées à ses différentes conformations potentielles. En raison d'un très large nombre de degré de liberté, une chaîne polypeptidique possède un nombre astronomique de conformations possibles (on estime à $\sim 10^{30}$ conformations différentes pour une petite protéine de 100 AA par exemple). Ainsi, si une protéine devait explorer séquentiellement l'ensemble de ces conformations, le temps nécessaire à l'obtention de sa conformation native serait beaucoup trop long, même dans le cadre d'interconversions conformationnelles extrêmement rapides. Or la plupart des protéines se *foldent* spontanément dans une échelle de temps de la milliseconde. C'est ce que l'on appelle le paradoxe de Levinthal [236]. L'hypothèse d'un repliement "dirigé" (*foldings funnel*) a été proposé pour résoudre ce paradoxe [237, 238, 239, 240]. Dans le cadre de cette théorie, les conformations natives des protéines correspondent aux conforma-

1.7 NCp7

tions les plus stables parmi l'ensemble des conformations possibles généralement inactives. Ainsi, même si le paysage énergétique est accidenté (*rough*), avec de nombreux *minima* locaux correspondant à des repliements partiels dans lesquels la protéine peut se retrouver piégée, l'état natif d'une protéine est associé à un *minimum* très marqué avec des parois très abruptes. Ce *minimum* correspond à une structure tertiaire unique bien définie. Le repliement "dirigé" permet à la protéine de se structurer selon des chemins préférentiels au sein du paysages énergétique. Il existe un large nombre de chemins et d'intermédiaires permettant d'atteindre l'état natif. Cette théorie a tout naturellement été étendue aux acides nucléiques et plus particulièrement aux ARN, notamment pour les ARN dits "fonctionnels" nécessitant une conformation bien définie pour être biologiquement actifs. Les ARN, de part leur structure monocaténaire, ont tendance à se structurer spontanément par l'intermédiaire d'hybridations intramoléculaires. Les structures résultantes sont souvent sub-optimales du point de vue thermodynamique. De ce fait, les paysages énergétiques des ARN sont considérés comme plus accidentés que ceux des protéines [241, 242, 243]. C'est sans doute pour cette raison que des protéines chaperonnes d'ARN sont apparues "nécessaires" au cours de l'évolution. Les protéines chaperonnes d'ARN sont définies comme des protéines dont le "rôle" est de permettre aux conformations métastables de se réarranger pour atteindre les conformations thermodynamiquement les plus stables [244].

Cette activité repose sur trois propriétés :

1. une capacité à se lier à des structures d'acides nucléiques très différentes, de manière séquence et structure non-spécifique afin de résoudre le maximum d'états métastables
2. une capacité à déstabiliser les structures d'AN afin de faire "respirer" les repliements d'acides nucléiques sub-optimaux
3. une capacité à favoriser l'hybridation de séquences parfaitement complémentaires afin de permettre un réappariement optimal et de générer les conformations les plus stables thermodynamiquement

La NCp7 possède ces trois activités vis à vis des ARN, qu'elle exerce également à l'égard des ADN. La NCp7 est donc une ARN et ADN chaperonne.

Une protéine qui se lie aux acides nucléiques

La NCp7 est capable de se lier à de très nombreux acides nucléiques, avec des affinités très variables, à la fois de manière spécifique et non-spécifique. La liaison de la NCp7 dépend notamment fortement des conditions salines [245, 246, 247, 248, 249]. Cette forte composante électrostatique, qui lui permet de se lier de manière non-spécifique à quasi n'importe quelle séquence d'au moins 6 nucléotides, explique en grande partie son absence de spécificité vis à vis des séquences nucléiques [250, 246]. Pour autant, la littérature décrit nombre de liaisons site-spécifiques [251, 252, 249, 253]. Les séquences simple-brin riches en UG- ou en TG- [254, 255, 245, 249, 256] ou les séquences GXG au sein des boucles simple-brin constituent des sites de liaison préférentiels [225, 228]. Ce type de liaison spécifique implique très fortement le plateau hydrophobe à la surface des deux doigts et l'empilement des AA *Phe*₁₆ et *Trp*₃₇ avec les bases. Il semblerait également que l'affinité pour les séquences d'ARN en tige-boucles soit supérieure à celles des analogues d'ADN [257, 258, 259, 249]. La capacité de NCp7 à se lier à différentes structures d'acides nucléiques suggère que la protéine est largement mobile et flexible à l'état lié [260, 261, 262, 263, 264, 265].

Une protéine qui déstabilise les structures des acides nucléiques

La NCp7 est capable de déstabiliser les acides nucléiques. Cette propriété a fait l'oeuvre de nombreuses études notamment biophysiques (pour une revue voir article I). Les doigts de zinc de la protéine sont responsables de l'activité de déstabilisation [226, 266]. La capacité de déstabilisation de la NCp7 est fortement corrélée à la stabilité locale des acides nucléiques, notamment à la présence de *bulges* ou de mésappariements, suggérant une co-évolution entre les séquences génomiques virales et la NCp7 [267, 268, 269]. La déstabilisation des AN se traduit de différentes manières, allant de la diminution significative de la stabilité de séquences double-brin [252, 270], à de subtiles modifications de structures tertiaires [271, 272, 273, 264] ou encore à une modification conformationnelle de séquences simple-brin comme la boucle de (-)PBS [227] (article V). L'efficacité de déstabilisation reste cependant modeste et la NCp7 n'est pas capable de déstabiliser des séquences très stables. Ceci prend tout son sens dans le cadre de la théorie du paysage énergétique. La NCp7 déstabilise donc principalement les états métastables, ce qui se traduit par une augmentation de leur énergie

1.7 NCp7

libre (comparer 1.27a et 1.27c figure 1.27 à titre d'illustration). Par contre, elle n'affecte que très peu les états stables et le niveau d'énergie de l'état natif reste donc inchangé. Ces modifications énergétiques remodelent profondément le paysage énergétique qui ne présente plus alors qu'un puit de potentiel marqué correspondant au repliement optimal (Figure 1.27).

Une protéine qui agrège les acides nucléiques et qui favorise l'hybridation de séquences complémentaires

La NCp7 promeut l'hybridation de séquences complémentaires, de manière concentration-dépendante. L'hybridation est optimale pour des concentrations saturantes de NCp7 [274, 275, 276, 250, 277, 252, 278]. La promotion de la réaction induite par NCp7 est d'autant plus marquée que les acides nucléiques sont initialement structurés [279]. Puisque la NCp7 se montre systématiquement plus efficace que la NC(11-55), un mutant de NCp7 délété de la partie N_{ter} basique, il faut en conclure que le domaine N_{ter} de la NCp7 joue un rôle plus important que les doigts dans la promotion de l'hybridation [280, 281, 282]. La nucléation des acides nucléiques est une réaction dépendant de la diffusion. Celle-ci est ralentie par l'existence de forces de répulsion induites par les charges des chaînes phosphates des acides nucléiques, ainsi que par la nécessité d'un positionnement adéquat des nucléotides pour pouvoir s'hybrider. La vitesse de nucléation d'AN structurés dépend également de la probabilité de fusion des éléments les moins stables de cette structure. En se liant sur les AN, la NCp7 induit, *via* ses AA basiques, un écrantage efficace des charges des chaînes phosphates. De plus, il a été montré que la NCp7 pouvait induire une vitesse d'hybridation supérieure de 3 *log* à la vitesse optimale obtenue à 68°C à 1M de NaCl [276], suggérant que la NCp7 est capable de "concentrer" les acides nucléiques. Cette capacité d'attraction, résulte de la capacité de NCp7 d'agréger de manière non-spécifique les acides nucléiques [276, 283, 284, 285]. La partie N_{ter} de NCp7 est responsable des propriétés d'agrégation. En effet, des études par diffusion dynamique de lumière (ou DLS pour *Dynamic light scattering*) ainsi que par microscopie électronique ont pu montrer que la protéine NCp7, mais également NCp9, sont capables d'agréger les acides nucléiques dans des conditions salines physiologiques avec un ratio nucléotide/protéine au voisinage de 8 au sein des agrégats. Ce

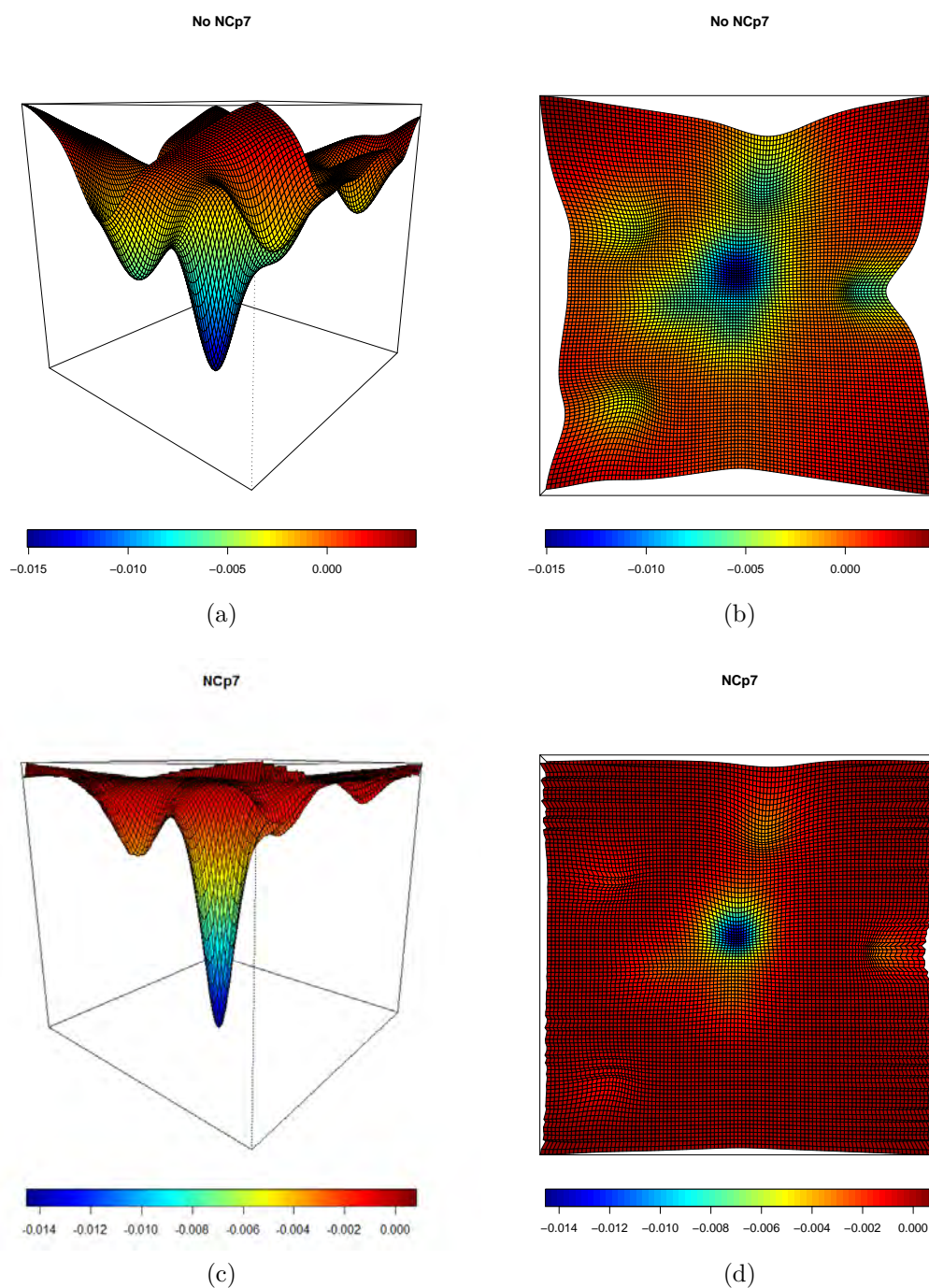


FIGURE 1.27 – Illustration d’un paysage énergétique (en 3D et en 2D) d’un acide nucléique en absence (a) et (b) ou en présence de NCp7 (c) et (d).

phénomène gouverne pour grande part les propriétés d’hybridation de la NCp7 [282].

Pour résumer, la NCp7 est capable de lier de nombreuses structures d’acides nucléiques, de déstabiliser les conformations les moins stables, et de favoriser, par sa propriété d’agrégation,

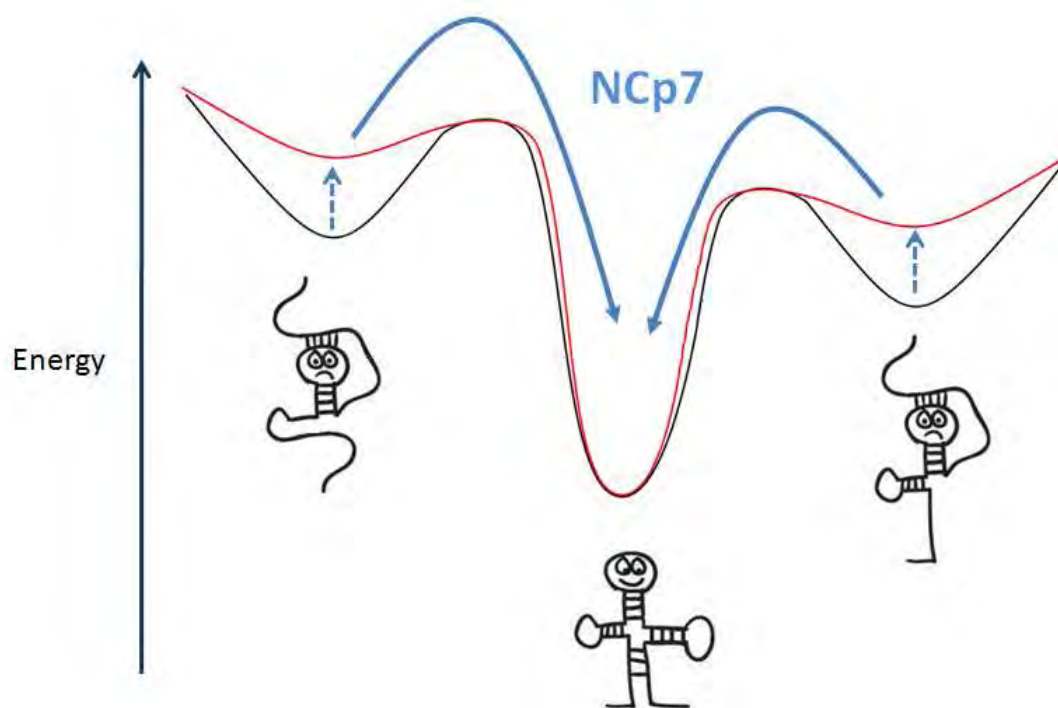


FIGURE 1.28 – Modèle de l'activité RNA- DNA-chaperonne de NCp7

Le paysage énergétique de l'ARN est représenté en noir. NCp7 déstabilise principalement les états métastables et dirige le repliement de l'ARN vers sa conformation la plus stable en modifiant le paysage énergétique (rouge).

gation, la formation des conformations les plus stables. Ceci est résumé par la figure 1.28.

1.7.4 Rôle de la protéine NCp7 au cours du cycle viral

Les propriétés chaperonnes de NCp7 vis à vis des acides nucléiques expliquent en grande partie le rôle critique que peut jouer la NCp7 au cours des étapes précoces et tardives du cycle viral.

Rôle de la NCp7 dans la dimérisation du génome viral

La dimérisation du génome viral semble être une étape préalable à son encapsidation [14]. La séquence DIS est le site majeur d'initiation de la dimérisation [286, 287, 288, 289], et sa conservation au sein des lignées du VIH semble montrer l'importance de ce motif. DIS peut former des complexes boucle-boucle grâce à l'autocomplémentarité de sa boucle. Un dimère étendu peut se former à 55°C [290]. La NCp7 permet la formation de ce duplex

1.7 NCp7

étendu à température physiologique [291]. La caractérisation de ces duplex *in vitro* s'est

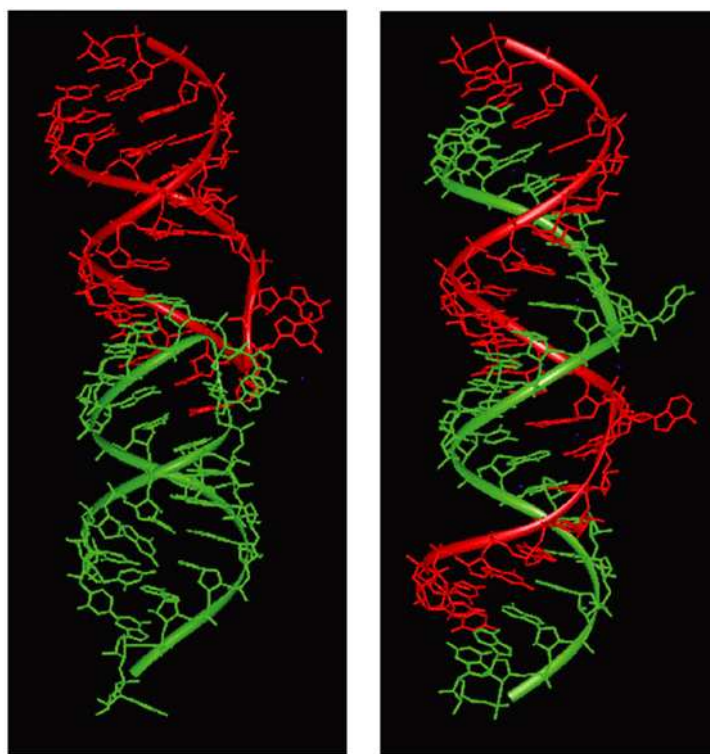
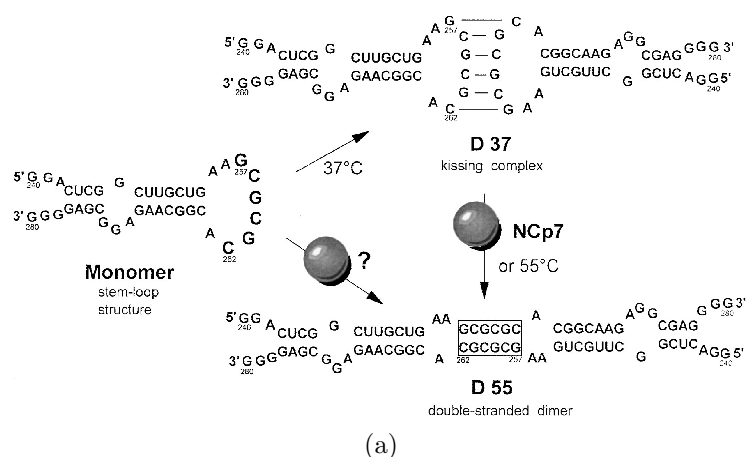


FIGURE 1.29 – (a) Dimérisation de DIS [292].
 (b) Structure RMN du “kissing complex” (gauche) et du duplex étendu (droite) [14].

montrée précieuse pour définir la relation entre la dimérisation de l'ARNv, l'encapsidation et la maturation de l'ARNg. La structure de ce dimère a également été caractérisée dans le but de constituer le rationnel pour la recherche de molécules pouvant inhiber cette étape [293, 294, 295]. Pour autant, il semblerait que les modèles *in vitro* ne reflètent pas com-

1.7 NCp7

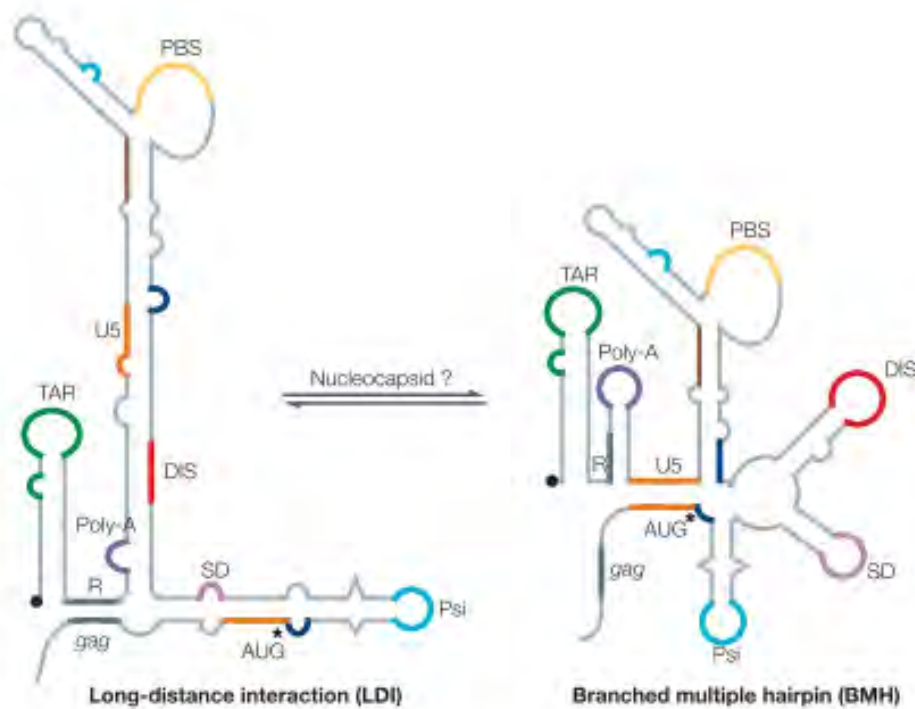


FIGURE 1.30 – Equilibre entre les conformations LDI et BMH de la région 5'UTR du VIH-1 permettant de masquer ou d'arborer, respectivement, le signal de dimérisation [14].

plètement ce qui existe *in vivo* [14]. De plus, l'existence au sein de la séquence TAR d'une séquence palindromique, révélée en présence de NCp7, pourrait constituer un autre site de dimérisation [296]. Enfin, il est à noter que DIS n'existe pas systématiquement sous forme de tige-boucle au sein de l'ARNv. L'existence de deux conformations principales du 5'UTR du génome du VIH-1 pourrait constituer un mécanisme de régulation. La forme LDI (*long distance interaction*) orienterait l'ARNg non-épissé vers la traduction de protéines alors que la forme BMH (*branched multiple hairpin*) permettrait la dimérisation de l'ARNg et son encapsidation (Figure 1.30). De part ses propriétés chaperonnes, la NCp7 jouerait un rôle dans la réorganisation de la structure du 5'UTR et pourrait de fait participer à la régulation de la dimérisation du génome viral [297, 298, 299].

Rôle de la NC au cours de l'encapsidation du génome viral

L'encapsidation du génome viral au sein des particules virales est une étape obligatoire à la formation de virions infectieux. Parmi tous les ARN coexistants au sein de la cellule infectée, l'ARNg est spécifiquement encapsidé. La sélectivité de l'encapsidation semble résulter de l'interaction spécifique du domaine NC de Gag, notamment *via* les doigts de zinc, avec le signal d'encapsidation Ψ situé en 5' UTR du génome viral [300]. La séquence Ψ est composée de quatre structures en tige-boucle SL1 à SL4. Des mutations affectant des éléments de cette séquence réduisent significativement l'encapsidation du génome viral [301]. L'intégrité des doigts de zinc est indispensable pour l'encapsidation du génome viral [232, 221, 302, 205], en accord avec un mode d'interaction spécifique impliquant le plateau hydrophobe décrit par les structures RMN de NCp7 avec SL2 ou SL3 et se traduisant par de fortes affinités [233, 225, 303, 234, 247]. La structure secondaire de la séquence Ψ semble prépondérante dans l'interaction [304, 305]. De plus, l'affinité de Gag est plus élevée pour l'ensemble de la séquence Ψ que pour chacune de ses boucles prises séparément [300, 301]. L'ensemble de ces éléments suggèrent que l'affinité du domaine NC de Gag pour le signal d'encapsidation est essentiellement structure-dépendant. Il est intéressant de noter que le rôle de reconnaissance spécifique de la NC au cours de l'encapsidation tranche singulièrement avec la nécessité de la NCp7 de se lier à de nombreux acides nucléiques différents pour assurer ses fonctions au cours des autres étapes du cycle viral. Il faut néanmoins rappeler que l'on discute ici de NC au sein de Gag et que la protéine non processée possède des propriétés différentes de la forme mature. Gag présente notamment des propriétés chaperonnes diminuées (particulièrement pour l'activité de déstabilisation) par rapport à NCp7 [210]. De plus Gag se lie de manière coopérative aux AN, avec une dynamique de liaison très différente de NCp7 [306], laissant supposer que la dynamique avec laquelle NC se lie aux acides nucléiques joue un rôle important pour expliquer les propriétés chaperonnes.

Rôle de NCp7 au cours de la reverse transcription

La transcription inverse est sans doute l'étape du cycle viral qui met le mieux en évidence les propriétés chaperonnes de la NCp7.

Rôle au cours de l'initiation de la reverse transcription

L'étape d'initiation de la transcription inverse requiert l'hybridation de l'amorce ARNt^{Lys3} cellulaire, sélectivement encapsidé [192, 307, 308], aux dix-huit nucléotides complémentaires de la séquence PBS (*Primer Binding Site*) située dans la région 5' du génome viral. Cette hybridation est stimulée par la présence de NC, au sein du précurseur de Gag [309, 310] ou sous sa forme mature NCp7 [271, 223, 311, 273, 312, 250, 313, 314, 263, 265]. La forme mature apparaît plus efficace pour l'hybridation de l'ARNt^{Lys3}, suggérant que les propriétés chaperonnes de la NCp7, plus marquées que celles de la forme non processée, sont requises [315] au cours de cette étape. L'hybridation ARNt^{Lys3} / PBS est une étape cinétiquement déterminante, d'ordre 2 [312]. La liaison de la NCp7 sur l'ARNt^{Lys3} ne déstabilise que légèrement la structure de celui-ci [271, 272, 273, 316, 317, 264]. La stimulation induite par NCp7 résulte principalement de son pouvoir agrégant vis-à-vis des acides nucléiques et la NCp7 intervient en facilitant l'hybridation de l'ARNt^{Lys3} sur PBS principalement par l'intermédiaire des résidus basiques de ses chaînes latérales. La réaction d'hybridation nécessite ensuite la fusion partielle de la structure de PBS et de l'ARNt^{Lys3}, avec la rupture des interactions tertiaires au sein de la boucle TΨC et entre la boucle TΨC et la boucle D, ainsi que des paires de bases, G6-U67 et T54-A58 au centre de l'ARNt [312, 264, 265] (Figure 1.31).

L'hybridation des 18nt en 3' de la séquence de l'ARNt^{Lys3} avec PBS ne constitue pas nécessairement la seule interaction permettant la formation d'un complexe stable entre l'ARNt^{Lys3} et l'ARNg viral. Il semblerait que *in vivo*, la formation de ce complexe implique des interactions spécifiques supplémentaires. Il a ainsi été montré qu'il existe également une interaction entre la boucle riche en A en amont du PBS et la boucle anti-codon de l'ARNt^{Lys3} dans le cadre de l'isolat *Mal* [318, 319, 320, 321]. De même, l'interaction de la partie 5' de la tige du bras TψC avec une région en amont de PBS appelée PAS (*Primer Activation Signal*) semble nécessaire pour activer la synthèse de l'ADN [322, 323, 324, 325]. L'efficacité de l'initiation de différents mutants semble corrélée à l'accessibilité du PAS [326]. Pour autant, cette interaction n'a pas été systématiquement retrouvée [327]. De plus, un mécanisme relativement différent, impliquant uniquement la liaison de la séquence PBS a été décrit pour les isolats *Hxb2* [328]. Ces interactions multiples et différentielles au sein des

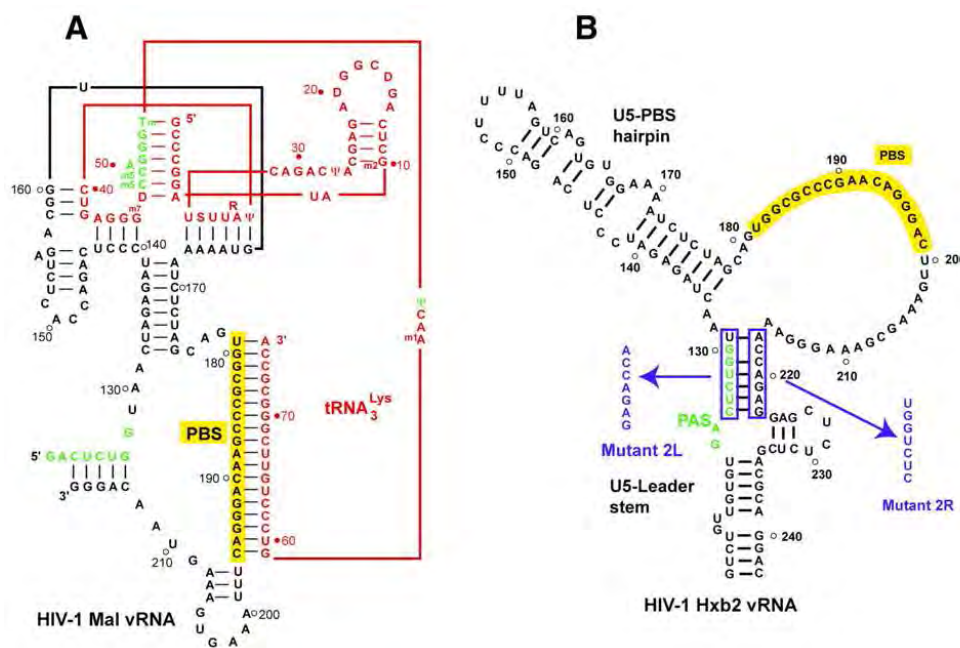


FIGURE 1.32 – Modèle du complexe ARNv-ARNt^{Lys3} pour les souches *Mal* (A) et *Hxb2* (B). [327]

souches virales traduisent la complexité des interactions existant entre les acides nucléiques au cours de la reverse transcription et la “nécessité” pour la NCp7 d’être peu spécifique pour les structures d’AN afin d’“accompagner” cette variabilité.

La synthèse de l’ADN(-) débute suite à cette étape d’initiation. Le passage de l’étape d’initiation à l’étape d’élongation de la synthèse d’ADN viral n’est pas immédiat. La RT effectue une pause après la mise en place du deuxième nucléotide. Cette étape serait nécessaire à l’enzyme pour acquérir une meilleure processivité. La NCp7 diminue également les pauses lors de l’initiation de la rétrotranscription [329, 330], probablement en stabilisant le complexe RT/amorce/matrice [331, 332]. En fait, la RT effectue même deux pauses successives aux positions +1 et +3 lorsque l’ARNt^{Lys3} est hybridé avec l’ARN viral par la chaleur [333]. Lorsque l’hybridation est réalisée en présence de la NCp7 la pause au nucléotide +1 disparaît [333], ce qui montre que la NCp7 favorise la formation d’un complexe ARNt^{Lys3} / ARNg dynamique dont la conformation est biologiquement efficace, s’adaptant au déroulement de l’initiation, par opposition au complexe thermodynamiquement stable. Ceci confirme le caractère dynamique de l’interaction de la NCp7 avec les AN.

1.7 NCp7

Suite à l'hybridation de l'amorce ARNt^{Lys3} sur le PBS, NCp7 stimule l'extension de l'amorce par la transcriptase inverse et la synthèse de l'ADN "strong stop (-)" (ou (-)ssDNA) impliqué dans le premier saut de brin.

Rôle au cours du premier saut de brin

Le premier saut de brin est une étape obligatoire à la formation complète de l'ADN(-), complémentaire à l'ARN génomique. Le (-)ssDNA néosynthétisé est transféré de l'extrémité 5' vers l'extrémité 3' de l'ARNg lors d'une réaction impliquant l'appariement des bases complémentaires des régions R [334, 335, 94]. Le premier saut de brin peut être intermoléculaire [336] ou intramoléculaire [337, 338, 339]. La NCp7 joue un rôle majeur lors du premier saut de brin en intervenant lors de l'hybridation de TAR avec cTAR. Ceci sera développé dans le chapitre consacré. Par ailleurs, la NCp7 augmente la processivité de la RT en déstabilisant les structures secondaires d'ARN en amont du site de polymérisation [340, 341, 168, 342]. La NCp7 stimule aussi l'activité RNase H de la RT, principalement au niveau d'un site de pause de la RT, à la base de la tige-boucle TAR à l'extrémité 5' de la séquence R de l'ARN viral. Il en résulte une augmentation de la coupure de la matrice ARN en position -18 par rapport au site de pause et l'apparition de sites de coupure additionnels, en positions -24, -28 et -38. Par cette stimulation, l'ARN viral est davantage clivé, créant un plus grand nombre de sites d'invasion et rendant le premier transfert de brin plus efficace

Rôle de NCp7 au cours du second saut de brin

La synthèse de l'ARN (+) débute au niveau du PPT, une séquence qui résiste à l'activité de l'ARNase H de la RT. Cette synthèse se poursuit jusqu'au niveau du duplex PBS/ARNt^{Lys3}. L'adénosine méthylée en position 58 (au sein de la boucle Ψ de l'ARNt^{Lys3}) pourrait constituer le signal de fin de synthèse [334] même si d'autres signaux ont été décrits [343, 344]. L'ARNt^{Lys3} est ensuite déplacé. La NCp7 pourrait jouer un rôle dans cette étape. Le brin d'ADN *strong stop* (+) est transféré et hybridé au PBS(-) situé en 3' de l'ADN. Cette étape, qui constitue le second saut de brin, permet de reconstituer la séquence du LTR 5'. L'investigation du mécanisme du transfert de brin (+) est basée essentiellement sur des systèmes modèles *in vitro*. La mutation E478Q au sein de la RT, abolissant l'activité

1.7 NCp7

RNase H, réduit dramatiquement le taux de transfert [345, 346, 347], ce qui met en évidence le rôle important de la dégradation de l'ARN dans ce transfert. La NCp7 joue également un rôle important. L'interaction de PBS avec NCp7 a été caractérisée au laboratoire. NCp7 ne déstabilise que très peu la tige de PBS [348], en accord avec la grande stabilité thermodynamique de la tige, et ce malgré sa taille limitée à 4 paires de bases et l'existence d'un *bulge*. PBS peut lier 3 NCp7. Un site de liaison a été identifié au niveau de l'extrémité protrudente simple brin, les deux autres au niveau de la boucle. Le trinuécléotide 6-TGC-8 lie préférentiellement NCp7 [227]. La liaison de la NCp7 sur la boucle induit un changement conformationnel de celle-ci, principalement en étirant la boucle [227]. En ce sens, NCp7 déstabilise PBS. La NCp7 stimule le saut de brin. La réaction d'hybridation (+)/(-)PBS se fait spontanément. En absence de NCp7, (+)PBS s'hybride principalement à (-)PBS par l'intermédiaire d'une étape de nucléation au niveau de l'extrémité simple brin. La vitesse de réaction est accélérée en présence de NCp7 d'un facteur 60. De manière intéressante, l'ajout de NCp7 résulte en un changement important du mécanisme de la réaction, puisque l'étape de nucléation implique alors les boucles. Une explication de ce *switch* mécanistique est donnée dans l'article V.

Malgré ces travaux, les données concernant le mécanisme du transfert de brin (+) restent limitées en comparaison avec les résultats obtenus sur le mécanisme du premier transfert de brin.

Rôle au cours de l'achèvement de la réverse transcription

Au cours de l'extension du brin (+) d'ADN, la RT fait également des pauses au niveau des domaines structurés de l'ADN (-) [349, 350, 351, 352]. La présence de NCp7, de part son activité de déstabilisation de ces structures, induit une augmentation de l'efficacité de la synthèse du brin (+) [341]. Une fois la réverse transcription achevée, il en résulte un ADN double brin linéaire possédant un segment simple brin central appelé *DNA flap*. En effet, le brin d'ADN(+) n'est pas continu [353, 354, 355]. L'initiation de la synthèse du brin d'ADN(+) se fait au niveau du PPTc en direction de l'extrémité 3' terminale. Après le second saut de brin, la synthèse de l'ADN(+) se poursuit au delà du PPTc, impliquant un déplacement de brin du segment d'ADN(+) initial d'environ 99 bases, formant le *DNA*

1.7 NCp7

flap [356, 357, 358]. Le *flap* apparaît important pour la réplication virale au sein des cellules qui ne se divisent pas en favorisant l'import nucléaire. En effet, les mutations du *flap* compromettent l'import nucléaire de l'ADN proviral [141], alors que certains vecteurs dérivés du VIH contenant ce *flap* central présentent des propriétés de transfection améliorées [359, 360, 361]. Le *flap* aurait également un rôle dans la décapsidation du virus au niveau de la membrane plasmique [130, 137, 131]. L'activité particulière de ce segment d'ADN pourrait résulter de dynamiques ou de conformations particulières [362, 363, 364], notamment induites par la NCp7 liant préférentiellement les ADN simple-brin. En effet, il a été montré que la NCp7 stabiliserait les fluctuations structurales de ce *flap* [358]. La NCp7 pourrait donc jouer, de part ses propriétés chaperonnes, un rôle important dans l'import du PIC. La présence de NCp7 liée au *flap* suggère que la NCp7 entre dans le noyau, elle pourrait donc également jouer un rôle dans l'étape d'intégration du génome viral.

Enfin, la synthèse du *flap* est stimulée par la présence de NCp7. Ceci constitue donc un exemple supplémentaire suggérant que la RT et la NCp7 agissent de manière synergique au sein des différentes étapes de la réverse transcription. De nombreux travaux ont émis l'hypothèse que les activités concertées de la NC et la RT résulteraient pour partie, d'une interaction directe entre les deux protéines [365, 331, 366, 274, 367, 368, 369, 340, 370, 205, 371]. Cette interaction a été mise en évidence *in vitro* [372, 373], indiquant que la NCp7 se lie sur la RT, probablement via ses doigts de zinc, avec une stœchiométrie 1 :1 et une affinité de $6 \times 10^{-7} M^{-1}$. Cependant, les détails moléculaires de cette interaction et ses implications fonctionnelles restent pour l'instant largement inconnus.

1.8 Objectifs de la thèse

Les propriétés chaperonnes de la NCp7 constituent un ensemble de caractéristiques permettant à la NCp7 de réarranger les acides nucléiques de manière à favoriser leurs conformations les plus stables. Ces propriétés sont nécessaires à la réplication virale du VIH. Elles reposent sur trois propriétés particulières vis-à-vis des acides nucléiques : une capacité de liaison à des substrats variés, une activité de déstabilisation et une activité d'agrégation des acides nucléiques. Bien qu'étudiés depuis déjà près d'une vingtaine d'années, les mécanismes responsables de ces activités restent largement méconnus à l'échelle moléculaire. L'objectif du travail de cette thèse est d'utiliser et de développer différentes approches expérimentales basées sur de techniques de fluorescence de pointes afin d'approfondir la compréhension de ces mécanismes au niveau moléculaire. Un effort tout particulier a été porté sur l'étude des aspects dynamiques de ces propriétés.

Plusieurs approches complémentaires ont été utilisées :

- Nous avons tout d'abord cherché à exploiter la méthode de spectroscopie à corrélation de fluorescence afin de *i.* caractériser la formation d'homodimères induits par NCp7, *ii.* de caractériser l'influence de la NCp7 sur la dynamique de repliement de la sonde marquant un acide nucléique et *iii.* nous avons utilisé les propriétés de la liaison de la NCp7 pour valider une méthode de spectroscopie à corrélation de fluorescence à deux volumes focaux permettant la détermination de constantes absolues de diffusion de particules fluorescentes.
- Dans un second temps, nous avons cherché à utiliser les propriétés du transfert d'énergie de fluorescence (FRET) au sein d'un acide nucléique doublement marqué pour décrire la cinétique et le mécanisme de l'hybridation de cTAR/TAR au cours du premier saut de brin.
- Nous avons ensuite cherché à exploiter les propriétés de la fluorescence de la 2-Aminopurine, une sonde structurale fluorescente, afin de caractériser l'influence de la liaison de la NCp7 sur la dynamique des acides nucléiques. Nous avons ainsi pu caractériser de manière site-spécifique la liaison de la NCp7 sur un oligonucléotide possédant plusieurs sites de liaison. Nous avons également pu montrer que les doigts

1.8 Objectifs de la thèse

de zinc de la NCp7 permettent d'induire une spécificité de mécanisme dans l'hybridation des séquences (+)PBS au (-)PBS, impliquées dans le second saut de brin. Enfin, la 2-Ap nous a permis d'étudier localement la déstabilisation de cTAR.

- Nous avons également développé une approche par spectroscopie de fluorescence à l'échelle de la molécule unique de manière à étudier la dynamique de la liaison de la NCp7 sur les acides nucléiques.

L'ensemble de ces travaux sera présenté à la suite d'une brève revue de la littérature sur les études biophysiques qui ont été menées pour étudier les propriétés chaperonnes de la NCp7. Cette revue présente l'état des lieux des principales connaissances concernant les propriétés de déstabilisation et d'agrégation obtenues par des méthodes biophysiques. Les résultats antérieurs obtenus au sein du laboratoire y sont rapportés. Elle constitue également *de facto* un état de l'art sur les méthodes biophysiques actuellement disponibles pour ces études. Cette revue est actuellement soumise à publication.

1.9

Publication I

1 **Biophysical studies of the nucleic acid chaperone properties of the**
2 **HIV-1 nucleocapsid protein**

3

4 Julien Godet^{1,2} and Yves Mély^{1*}

5

6 ¹ Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de
7 Strasbourg, Faculté de Pharmacie, 74 route du Rhin, 67401 Illkirch, France

8 ² Medical Information and Biostatistics Department, Strasbourg University Hospital, 1 pl. de
9 l'Hôpital, 67000 Strasbourg, France

10

11

12

13

14

15

16

17

18

19

20

21

22

23 * Corresponding author

24 Tel : +33 (0)3 90 24 42 63 Fax : +33 (0)3 90 24 43 12

25 **e-mail:yves.mely@unistra.fr**

26 **The authors declare no conflict of interest**

27

1

2 **Key words:** HIV-1, nucleocapsid protein, RNA-chaperone, destabilization, annealing, biophysical
3 techniques, fluorescence, single-molecule, FRET, NMR

4

5 **Acknowledgments:** This work was funded by the Agence Nationale de Recherche sur le SIDA and
6 Sidaction

7

8 **Abbreviations:** HIV-1: Human Immunodeficiency Virus- type I, NCp7: nucleocapsid protein of
9 HIV-1, ZF: zinc finger, NA: Nucleic Acid, TAR: TransActivation Response element, PBS: Primer
10 Binding Site, FRET: Fluorescence Resonance Energy Transfer, TCSPC: time correlated single
11 photon counting, FCS: Fluorescence Correlation Spectroscopy.

1 **Abstract**

2 The nucleocapsid protein NCp7 of HIV-1 possesses nucleic acid chaperone properties that are
3 thought to be crucial throughout the viral life cycle. These properties promote the rearrangement of
4 nucleic acids into their most thermodynamically stable conformations. These NCp7 properties
5 involve two components, namely nucleic acid destabilization and activation of the annealing of
6 complementary sequences. Biophysical techniques have been found extremely powerful to decipher
7 the molecular mechanisms underlying these two components. We propose here to overview the
8 recent reports that examine the nucleic acid chaperone properties of NCp7 by these techniques.

9

10 **Running title:** Biophysical studies of the NCp7 chaperone properties

11

12

1 Introduction

2 The HIV-1 nucleocapsid protein is a small (55 aminoacids) basic protein, characterized by two
3 strictly conserved zinc fingers (ZFs) that chelate zinc ions with high affinity¹ (Fig. 1). The binding
4 of zinc through the cysteine and histidine residues of the ZFs provides a conformational transition
5 from a flexible unfolded to a folded constrained structure, in which a hydrophobic plateau forms at
6 the top of the ZF domain. This plateau is composed of the residues Val13, Phe16, Thr24 and Ala25
7 of the proximal ZF and the residues Trp37, Gln45 and Met46 of the distal ZF.²⁻⁴ This hydrophobic
8 plateau directs the nucleic acid (NA) binding properties of NCp7⁵⁻⁹, with notably a strong
9 contribution in the stabilization of the complexes due to the stacking of Trp37 with the guanine
10 bases.¹⁰ The complexes are further stabilized by electrostatic interactions with the basic amino acids
11 of the protein. NCp7 binds both specifically and nonspecifically to NA sequences. Specific binding
12 is notably required for the selection of the genomic viral RNA among the large excess of cellular
13 RNAs.^{11, 12} Selectivity is thought to be mediated by the hydrophobic platform of NCp7.^{7, 13, 14} In
14 contrast, the extensive coating of the viral RNA by approximately 1500 copies of NCp7 in the inner
15 core structure of the viral particle relies largely on the ability of NCp7 to bind to most if not all
16 sequences of seven nucleotides in length.¹⁵⁻¹⁸ Importantly, zinc ejection (for a review see references
17^{19, 20}) or mutations affecting the ZF folding and formation of the hydrophobic plateau, cause a
18 complete loss of virus infectivity.²¹⁻²⁶ This has been related to the central role of NCp7 in HIV-1
19 replication, notably in the reverse transcription, integration and assembly steps. In fact, the
20 functions of NCp7 in the viral life cycle are thought to largely rely on its NA chaperone properties
21²⁷⁻³², which direct the structural rearrangement of NAs into their most stable conformation and
22 promote the annealing of complementary NA sequences.^{31, 33} Two main components account for the
23 NCp7 chaperone properties. The first one corresponds to the transient destabilization of the NA
24 secondary structure upon binding of NCp7 molecules.³⁴⁻⁴⁰ The second component corresponds to
25 the NCp7-promoted annealing of complementary NA sequences.⁴¹⁻⁴⁸ Numerous efforts have been
26 performed to understand the molecular mechanisms underlying these two components. In this
27 respect, our aim will be to overview the progresses in this field that have been achieved by using
28 biophysical techniques.

29

1 **Nucleic acid destabilization activity of NCp7**

2

3 One of the first evidence of the NA destabilizing activity of NCp7 was obtained using thermal
4 melting experiments. NCp7 was shown to depress the melting temperature of short duplexes by
5 about 10°C in low salt conditions⁴⁹, suggesting a weak ($\Delta G = -0.23$ kcal/mol/bp) NCp7-induced
6 duplex destabilization.⁴³ The destabilizing activity of NCp7 was further characterized by
7 absorbance and fluorescence spectroscopy on the transactivation response element TAR RNA and
8 its cDNA copy, cTAR.^{34-36, 50, 51} The TAR and cTAR sequences are extended but imperfect stem-
9 loops with regularly spaced destabilizing motifs (bulges, mismatches and internal loop) that delimit
10 five double stranded segments (Fig. 2A). The TAR and cTAR sequences play a critical role in the
11 obligatory first strand transfer during reverse transcription. Due to the sensitivity of optical
12 techniques on light diffusion, the wt NCp7 was substituted by NC(12-55), a truncated NCp7
13 derivative, lacking the N-terminal basic domain that is responsible of the strong NA aggregating
14 properties of NCp7.⁵² A melting of 7-8 base-pairs of the cTAR at saturating concentrations of
15 peptide (1 peptide for 5 nt) was inferred from the hypochromicity changes in the absorption spectra.
16 In contrast, NC(12-55) induced only a limited melting (~1bp) in the more stable TAR RNA
17 sequence.

18 To investigate more specifically the effect of NC(12-55) on the cTAR stem, a cTAR sequence
19 labelled at its 3' and 5' ends by a fluorescence resonance energy transfer (FRET) donor/acceptor
20 couple was used. Time-resolved fluorescence measurements were performed on this doubly labelled
21 sequence to gain information on the different cTAR conformational states within the sample.
22 Indeed, the energy transfer efficiency, E , between the two dyes is strongly dependent of the inter-
23 dye distance r and can be expressed as:

$$24 \quad E = 1 - \frac{\tau_{DA}}{\tau_D} = \frac{R_0^6}{R_0^6 + r^6} \quad \text{Equation (1)}$$

25 where τ_{DA} and τ_D correspond to the lifetime of the donor in the presence or the absence of the
26 acceptor, respectively, and R_0 is the Förster distance for which the energy transfer is equal to 0.5.
27 As a consequence, time-resolved fluorescence measurements can provide the average dye-to-dye
28 distances associated to the different conformations of the molecules, together with their relative
29 amplitudes. The time-resolved fluorescence decay of the doubly-labelled cTAR was recorded using
30 the time correlated single photon counting (TCSPC) technique (Fig. 2B). In TCSPC, a fluorescent
31 sample is excited with a pulsed laser light and the fluorescence decay is reconstituted by measuring
32 the arrival time of individual photons relatively to the excitation pulses (Fig. 2C). The distribution
33 of the most probable lifetimes can be extracted from these decays (Fig. 2D).⁵³ In the absence of

1 NCp7, the time-resolved fluorescence decay of the doubly-labelled cTAR appeared more complex
2 than that of cTAR singly labelled by the donor, with the appearance of short-lived components^{36, 54}
3 and a significant proportion of « dark-species » with lifetimes shorter than ~50ps (Fig. 2E).³⁶ This
4 fluorescence decay suggested an equilibrium between cTAR species with a closed stem, exhibiting
5 a non fluorescent excitonic heterodimer between the two dyes^{51, 55} and partly melted cTAR species,
6 with a FRET mechanism between the two dyes.^{36, 54} The energy transfer efficiencies calculated from
7 the lifetimes (Equation 1) provided the average end-to-end distances of the different open species in
8 thermal equilibrium with the closed species. This in turn allowed estimating the number of melted
9 base pairs in each of these open species. Moreover, the amplitudes associated to the different
10 lifetimes being proportional to the populations of the species, the time-resolved fluorescence
11 provide thus, a clear view of the nature of the species in solution and their relative proportions. This
12 technique appears thus as a powerful technique to monitor the mechanism of fraying in
13 oligonucleotides (ODNs) and its perturbation by proteins.

14 Addition of NC(12-55) to cTAR at a ratio of one protein per five nucleotides does not generate new
15 species but shifts the fraying equilibrium towards the open species, favouring mainly the species
16 where the lower half of the stem is melted (Fig. 2E and F). The NCp7-induced melting of the stem
17 strongly relies on the presence of the terminal bulges within the stem, which cooperatively
18 destabilize the cTAR secondary structure.^{34, 37, 56} In contrast, the top half of the sequence is much
19 more stable and thus, resistant to the NCp7-induced destabilization activity.³⁵ The terminal bulges
20 being well conserved among HIV-1 strains, a co-evolutionary relationship between the TAR stem
21 stability and NCp7 destabilizing activity can be envisioned. NCp7 structural determinants for cTAR
22 destabilization were identified using a series of NCp7 mutants.⁵⁷ NCp7 destabilization properties
23 are supported by the two ZFs in their proper context and the basic 29-RAPRKKG-35 linker. The
24 strict requirement of the properly folded ZFs suggests that the hydrophobic platform on their top is
25 crucial for binding and destabilization of the TAR species. In fact, this platform and notably, the Trp
26 residue at position 37 in the distal finger appear as the major determinant for propagating the initial
27 melting of cTAR from the ends, up to the middle of the stem.

28 The molecular mechanism by which NCp7 destabilizes the cTAR stem was further characterized by
29 fluorescence correlation spectroscopy (FCS).⁵⁸ This technique is based on the analysis of the
30 fluorescence intensity fluctuations observed within the sub-femtoL focal volume defined either by a
31 confocal microscope or two-photon excitation (Fig. 3A). Analysis of these fluctuations by an
32 autocorrelation function allows retrieving the apparent diffusion time and the average number of
33 fluorescent molecules within the excitation volume, as well as the kinetics of the physico-chemical
34 events that modify the quantum yield of the fluorescent species diffusing through the observation

1 volume. As an example of such physico-chemical events, the fraying of the doubly labelled cTAR
2 stem results in dynamic changes of the dye-to-dye distance, giving rise to additional fluorescence
3 fluctuations as compared to the corresponding singly-labelled DNA sequence (Fig. 3B). These
4 dynamic changes in the end-to-end distances could be characterized from the ratio $G^*(\tau)$ of the
5 auto-correlation curves of the doubly- and singly-labelled cTAR (Fig. 3C). The $G^*(\tau)$ curve
6 appeared as a single exponential decay, from which the kinetic rate constants governing the fraying
7 of cTAR stem could be deduced.⁵⁰ NC(12-55) was found to increase mainly the opening rate
8 constant of cTAR. In addition, NCp7 changed the shape of the $G^*(\tau)$ curve, which became a
9 stretched exponential decay (stretched parameter ~ 0.5), indicating a broad distribution of energy
10 barriers between the closed and open states. Thus, NCp7 likely augments the number of pathways
11 between the open and closed states of cTAR, suggesting that it initiates the melting of base-pairs at
12 different locations within the terminal segments of cTAR.

13 Similar approaches have been used to characterize the NCp7-induced destabilization of the (+)PBS
14 and (-)PBS DNA sequences that form short but highly stable stem-loops with a single strand
15 overhang at their 5' and 3' end, respectively.^{38, 59} Both sequences play a key role in the second
16 strand transfer during reverse transcription. NCp7 was found to induce only a marginal
17 destabilization of the PBS stems, limited to its terminal base pair.

18 The conformational distribution and fraying dynamics of the cTAR sequence in the presence and
19 absence of NCp7 was also explored by single-molecule FRET experiments, which give access to a
20 wider range of observation times (from ms to 10th of sec) as compared to FCS. In these
21 experiments, the fluorescence intensities of the FRET donor and acceptor of single immobilized
22 doubly-labelled DNA hairpins are recorded as a function of time. The immobilization of the DNA
23 hairpins onto the glass is achieved using a biotin linker attached to a dT inserted within the hairpin
24 loop and a streptavidin-coated glass surface (Fig. 4A). The simultaneous detection of both donor
25 $I_D(t)$ and acceptor $I_A(t)$ fluorescence intensity time trajectories can be used to calculate the time
26 trajectory of the FRET efficiency (Fig. 4B) according to:

$$27 \quad E = \frac{I_A(t)}{I_A(t) + \gamma \times I_D(t)} \quad \text{Equation (2)}$$

28 where γ is a correction factor which accounts for the differences in quantum yield and detection
29 efficiency between the donor and the acceptor intensities.

30 These experiments confirmed that NCp7 shifts the equilibrium towards the open species and that
31 the opening of the stem is strongly dependent on the number of internal bulges within the stem.⁵⁶
32 Moreover, they reveal a complex mechanism for secondary structure fluctuations, with a strong
33 dependence of the FRET dynamics on both the NCp7 and Mg^{2+} concentrations. Finally, the

1 dynamics of the fraying process were found to occur over a wide time range (~ 5 to >250 ms), with
2 the involvement of long-lived intermediates,³⁷ confirming that the NCp7-chaperoned opening of
3 the cTAR hairpins followed a non-two-state kinetics.

4 Due to the strong dependence of the NA destabilizing activity of NCp7 on its structure, and the high
5 sensitivity of fluorescence techniques to monitor the NCp7-induced destabilization of stem-loops, a
6 robust and simple high throughput screening assay has been developed to identify small molecules
7 able to inhibit this highly specific activity of NCp7. Using doubly labelled cTAR as a target
8 sequence, five non zinc ejecting NCp7 binders with μM range K_i were identified from a library of
9 4800 compounds. These five compounds constitute a new series of anti-NCp7 molecules that could
10 be used to rationally design molecules with potential anti-viral activities.⁶⁰

11 In addition to these characterizations by fluorescence spectroscopy, the NA destabilizing activity of
12 NCp7 was also investigated using single molecule stretching experiments (Fig. 5).^{40, 61} In these
13 experiments, a λ -DNA is attached by each of its ends to polystyrene beads. One of the two beads is
14 coupled to a glass micropipette tip while the second one is trapped by an optical tweezer. The single
15 DNA molecule can be stretched by moving the pipette relative to the bead in the optical trap. The
16 DNA extension can be retrieved from the positions of the pipette. While stretching the DNA, the
17 bead is slightly moved within the trap due to the external force, so that the trapping laser beam is
18 deflected. Using a dual-beam instrument, the deflection can directly be measured using a position
19 sensitive photodiode detector⁶², so that the force applied on the trapped bead can be measured.⁶³

20 The force-induced melting of a DNA molecule is thermodynamically equivalent to its thermal
21 melting⁶⁴⁻⁶⁹ but can be performed at room temperature. Both the free energy ΔG and the
22 cooperativity parameter σ of the helix to coil transition of a single DNA molecule can be obtained
23 from the force–extension curves. The ΔG value is given by the area between the stretching curves
24 for double- and single-stranded DNA (Fig. 5.B light grey area), while the σ value is deduced from
25 the slope of the transition. Either a decrease in ΔG or in the transition cooperativity (which
26 corresponds to an increase of the σ parameter) indicates a destabilization of the NA secondary
27 structure, since ΔG is also the energy per base pair required to undergo the helix to coil transition,
28 while $\frac{1}{\sqrt{\sigma}}$ indicates approximately the number of base pairs required to initiate a melted domain
29 within a double-stranded NA structure.⁷⁰ In low salt conditions ($\leq 50\text{mM}$), saturating concentrations
30 of NCp7 (1 NCp7 per 8 nucleotides) decreased the cooperativity of the helix-coil transition of the
31 model λ -DNA sequence and reduced by a factor of two its transition free energy at room
32 temperature. As a low cooperativity enables thermal fluctuations to cause the melting and
33 reannealing of DNA domains, the NCp7-induced changes in cooperativity likely allow

1 complementary strands to sample states that differ by only a few base pairs. As a result of this
2 sampling, NCp7 allows the DNA sequence to rearrange in its most stable conformation. In contrast
3 to the wild-type NCp7, the fingerless (SSHS)₂NCp7 increased the helix-coil transition energy in
4 low salt conditions without altering σ , confirming that the ZFs support the NCp7 helix destabilizing
5 function.⁴⁰ A set of mutations within the ZFs was further used to investigate the requirement of the
6 specific ZF architecture in the NA chaperone activity of NCp7.⁶¹ The cooperativity parameter of the
7 DNA helix-coil transition of the NC mutants decreased relative to that of the wild-type NCp7, in a
8 manner that strongly correlates with the chaperone activity measured *in vivo*. Interestingly, the
9 integrity of the proximal ZF appeared more critical than that of the distal one for the NA
10 destabilizing activity of NCp7. The relative deficiencies of the mutants to alter the cooperativity
11 compared to the wild-type NCp7 were likely due to alterations in the specific interactions between
12 the ZFs and the DNA. As a consequence, these data confirmed the requirement of the *bona fide*
13 architecture of the two ZFs for the protein to destabilize the NAs and exhibit optimal chaperone
14 activity.

15 Similar experiments were used to compare the NA destabilizing activity of the sequential
16 proteolytic cleavage products of the structural Pr55^{Gag} polyprotein (namely NCp15, NCp9 and
17 NCp7). The final aim was to explore how the sequential proteolysis of HIV-1 NC affects its NA
18 interaction properties, and whether this can be correlated with the changes of NCp7 functions in the
19 viral life cycle.⁷¹ Addition of NCp7, NCp9 or NCp15 results in a significant DNA destabilization (~
20 1kcal/mol.bp), while the corresponding destabilization was only 0.3 kcal/mol.bp when Gag was
21 added. Moreover, the force at the midpoint of the DNA force-induced melting transition was
22 increased upon Gag binding, suggesting that the ssDNA was torsionally constrained in the presence
23 of Gag. This probably results from strong Gag-Gag interactions that occur within neighbouring Gag
24 molecules bound to the DNA. As a consequence, while the processed forms of NC appear to act as
25 NA chaperones, the NC domain within the context of Gag appears to have mostly a NA binding
26 function. Finally, the effect of selected compounds that target NCp7 was investigated using this
27 technique⁷², which provides the possibility to investigate the mechanism by which these
28 compounds alter the NA chaperone activity of NCp7. Thus, simple competitive inhibitors can be
29 distinguished from specific inhibitors of the NCp7 chaperone activity. Nevertheless, the relative
30 complexity of the method impedes it to substitute for initial high-throughput screening assays.

31 NMR spectroscopy constitutes another biophysical approach to investigate the chaperone properties
32 of NCp7 at the molecular level. Beside its crucial role in solving the structure of both free NCp7^{2-4,}
33⁷³ and NCp7 complexed to various ODNs⁶⁻⁹, NMR spectroscopy was used to determine how NCp7
34 destabilizes NAs. Although technically challenging, NMR analysis of NA-protein complexes is

1 probably one of the few techniques, together with fluorescence spectroscopy, which can provide
2 useful dynamic data and thus, contribute to a deeper understanding of the molecular mechanisms of
3 the NCp7 chaperone functions. The structure of the stable 18-nucleotides (-)PBS DNA stem-loop
4 and its complex with NCp7 were investigated by NMR.⁷⁴ The Watson-Crick imino protons of the (-
5)PBS stem were used to investigate the stability of the stem. Addition of NCp7 to (-)PBS induced a
6 pronounced broadening of the Guanine imino protons involved in the G:C base pairs, as well as a
7 general broadening of the non-exchangeable protons, showing that NCp7 increased the exposure of
8 the stem bases to water. To further characterize the NCp7-induced (-)PBS destabilization, the
9 Watson-Crick imino proton signals were monitored as a function of temperature. The maximal
10 temperature at which the imino protons of the free (-)PBS were visible was 40°C. The signal loss,
11 above the thermal melting transition, results from rapid exchange of the imino protons with water
12 protons when the stem is melted. In the presence of NCp7, added at a 1:1 ratio to (-)PBS, the
13 maximum temperature at which imino proton signals were observable was decreased by a factor of
14 about two. Taken together, these data strongly suggest that NCp7 alters the structure of the (-)PBS
15 stem which results in an increased exposure of the stem bases to water. Both the ZFs and the N_{term}
16 domain of the protein were found to directly participate to this process. Another important feature is
17 that the binding of NCp7 mainly occurs in the fast-exchange regime and involves multiples species,
18 thus precluding high-resolution structure of the complex.

19 Further NMR studies were performed with a $\Delta P(-)$ PBS mutant lacking the 3'-end single strand
20 overhang to decrease the size of the complex. The NC(11–55) peptide was found to bind $\Delta P(-)$ PBS
21 through its hydrophobic plateau to the 5' end of the loop (5-CTG-7) in a slow-exchange equilibrium
22 on the NMR time scale. This binding of NCp7 stretches the loop and exposes the nucleobases 8-
23 TTC-10 to the solvent. This NCp7-induced exposure of the loop bases is likely required to promote
24 the kissing interaction with the complementary (+)PBS sequence.⁴⁵ While interacting with
25 $\Delta P(-)$ PBS, NC(12-55) also weakens the upper C5·G11 base pair of the stem. This destabilization of
26 the stem may favour the conversion of the (-)/(+)PBS kissing loop complex into the final extended
27 duplex. Moreover, NCp7(12–55) binds also $\Delta P(-)$ PBS on an additional site (10-CGG-12), but with
28 a lower affinity and an intermediate-exchange equilibrium on the NMR time scale⁶, further
29 highlighting the NCp7 ability to interact both specifically and non-specifically with NAs. NMR
30 investigation of the NCp7 destabilizing activity was also performed with the much longer ¹⁵N-
31 labelled tRNA^{Lys3} sequence.⁷⁵ NC(12-53) molecules bind preferentially at the inside of the L-shape
32 of tRNA^{Lys3}, namely on the acceptor stem around the G6-U67 base pair and at the level of the
33 tertiary interactions between the D and T- Ψ -C loops. The binding of the NC(12-53) molecules
34 destabilizes these tertiary interactions within the tRNA^{Lys3} structure but does not melt any base-pair,

1 nor unwind the tRNA^{Lys3} helical domains. These results were in full line with circular dichroism
2 spectroscopy experiments showing that NCp7 induces a marginal tRNA^{Lys3} unwinding⁷⁶ and with
3 terbium probing experiments showing that NCp7 disrupts the D-TΨC tertiary interaction of
4 tRNA^{Lys3} at low NCp7/tRNA ratio and slightly destabilizes the acceptor-TΨC helix at saturating
5 NCp7 concentrations.⁷⁷ Moreover, the binding of NC(12-53) on tRNA^{Lys3} results in an unexpected
6 fast exchange mode which suggests that the protein slides between its binding sites, likely
7 explaining the previously reported inability to detect any strong affinity binding site.¹⁶
8 Further insight into the mechanism by which NCp7 destabilizes NA sequences was recently
9 provided by monitoring the interaction of NCp7 with ODNs labeled by 2-deoxy-aminopurine (2-
10 Ap).^{78, 79} 2-Ap is a fluorescent analogue of adenosine that minimally disturbs the NA structures and
11 can site-selectively report on the dynamics of interaction with proteins through time-resolved
12 fluorescence measurements.^{80, 81} Major changes in the structure and dynamics of the 2-Ap-labeled
13 ODNs could be evidenced in response to NCp7 binding. In fact, the effects of NCp7 binding on 2-
14 Ap fluorescence suggest a restriction of the ODN flexibility and 2-Ap local mobility that impedes
15 collisions of 2-Ap with its neighbor bases. The ‘freezing’ of the ODN flexibility and the local
16 dynamics of the bases was clearly attributed to the hydrophobic platform at the top of the folded
17 ZFs that also supports the ability of NCp7 to destabilize the stem of various stem-loops.⁵⁷ In this
18 respect, the restriction of the ps–ns ODN dynamics is likely a mechanistic component of the NA
19 destabilizing activity of NCp7, which may favour the disruption of base pairs in the double strand
20 segments bound to NC.^{34, 36, 37, 50, 54, 56} Moreover, the ‘freezing’ of the loop backbone and bases in
21 stem-loops likely constitutes together with the NCp7-induced changes in the loop structure an initial
22 event that favours the interaction with their complementary sequences, explaining the loop–loop
23 promoted annealing of (–)PBS/(+)PBS⁴⁵ and mini-TAR sequences⁴⁶ as well as the NCp7-promoted
24 formation of PBS^{38, 59} and DIS⁸² kissing loop homodimers. As a consequence, the NA “freezing”
25 activity of NCp7 is likely a key mechanistic component of its chaperone activity.
26
27 Taken together, these biophysical studies suggest that the NA destabilizing activity of NCp7 is
28 mainly driven by its highly dynamic binding through its ZFs and basic residues. The destabilizing
29 activity of NCp7 appears to be weak and strongly dependent of the stability of the NA sequences.
30 Of particular importance is the presence of destabilizing motifs (bulges, internal loops, mismatches
31 and terminal base pairs) scattered along the NA sequence. These motifs are likely of importance to
32 ensure a selective destabilization by NCp7 of a limited number of NA sequences that play a key role
33 in different steps of the viral life cycle. On a mechanistic point of view, one can speculate that the
34 Np7-induced NA destabilization initiates preferentially at single-stranded regions generated by

1 thermal fluctuations in the vicinity of the aforementioned destabilizing motifs. As a consequence,
2 the binding of NCp7, which decreases the enthalpy barrier for the opening of the neighbouring
3 base-pairs, helps the initial destabilization to propagate. This in turn, may allow the sliding of NCp7
4 and/or the binding of additional NCp7 molecules. Since the melting can be initiated from different
5 locations within the NA, the extent of transiently melted domain within a double-stranded sequence
6 likely depends on the number of concomitantly melted regions in close vicinity that could act in
7 synergy. Nevertheless, an extended melting by NCp7 of an imperfect double stranded sequence and
8 a subsequent strand exchange can only be achieved when its fully complementary NA sequence is
9 present.⁵⁴ Thus, the destabilizing activity of NCp7 induces complex secondary structure
10 fluctuations over a wide time range and ultimately leads to the reactive species required for the
11 specific intra- or -intermolecular annealing processes occurring during the viral life cycle.

12

13 **Nucleic acid annealing activity of NCp7**

14

15 HIV viral assembly and replication require multiple annealing reactions involving a large number of
16 NA sequences. For instance, hybridization of the cellular tRNA^{Lys3} to the 18-nucleotide priming
17 binding site (PBS) is required to initiate reverse transcription, while RNA-DNA hybridization
18 between the terminal redundancies of the LTRs involving the TAR sequences and DNA-DNA
19 hybridization of (-)PBS with (+)PBS occur during the first and second strand transfer step of
20 reverse transcription, respectively. A number of simple and robust assays have been developed for
21 examining the NA chaperone properties of NCp7 as well as other RNA-binding proteins.^{28, 83} Gel
22 electrophoresis, using radioactive labelling is still the most popular technique to monitor RNA-,
23 DNA- or RNA/DNA-annealing reactions. Though this technique is simple and robust, it suffers
24 from a limited time resolution and number of time points, poor resolution of intermediate species
25 and deleterious effects of the electric field on the low stable migrating complexes. An alternative
26 approach is to monitor by FRET the annealing of two fluorescently labelled complementary
27 sequences (Fig. 6A).⁸⁴ Indeed, on duplex formation, the fluorescence donor on one sequence and
28 the fluorescence acceptor on its complementary sequence become close enough for FRET to occur.
29 Since the increase in energy transfer is proportional to the fraction of formed duplexes, this
30 technique is well suited to monitor in real time the annealing reaction and to identify intermediates
31 that are characterized by a different FRET efficiency than the final extended duplex.^{49, 85} Both gel
32 electrophoresis and FRET techniques showed that, like a variety of other basic proteins, NCp7
33 promotes the annealing of various complementary NA sequences, in a concentration dependent
34 manner, being optimal at saturating levels of the protein.^{29, 32, 48, 49, 86-88} The largest annealing

1 promotion by NCp7 is observed with folded RNAs.⁸⁹ For instance, the NCp7-assisted rate of a 149
2 bp DNA fragment was 100-200-fold faster than that obtained in conditions (68 °C, 1 M NaCl)
3 optimal for strand renaturation, suggesting that NCp7 strongly facilitates the nucleation of the
4 duplex. Nucleation is a diffusion-limited association, modulated by the electrostatic repulsions
5 between the annealing strands and by the low probability for complementary strands to encounter
6 with an appropriate conformation for annealing.^{90, 91} As a consequence, a large part of the NCp7
7 activity in promoting the annealing of complementary sequences is thought to rely on its ability to
8 screen the NA phosphate charges and promote the NA aggregation, thus increasing the effective
9 concentration of NAs through a molecular crowding effect.^{33, 40, 61, 92} Interestingly, a rate
10 enhancement comparable or even greater to that for NCp7 has been observed for other RNA
11 binding proteins like HIV Tat⁹³, the Hepatitis C virus core protein⁹⁴, p53⁹⁵ and the human prion
12 protein.⁹⁶ However, these annealing assays suffer from a limited specificity in identifying chaperone
13 proteins, since non specific ions or polycations like Mg²⁺⁹⁷, spermidine⁹⁸, polylysine^{77, 99}, and
14 cationic detergents¹⁰⁰ can also induce attractive forces between NAs and increase the NA annealing
15 efficiency.^{30, 91, 100}

16 Strand-exchange assays are thought to rely more stringently than annealing assays on helix-
17 destabilization and thus, are supposed to provide a more specific test to determine the chaperone
18 activity of a given protein (Fig. 6B). These assays utilize three ODNs, namely a plus-strand which
19 is heat-annealed to a mutated minus-strand and a wild-type minus single-stranded sequence. Since
20 the mutated minus strand exhibits less complementarity than the wild-type sequence for the plus
21 strand sequence, addition of the chaperone results in an exchange within the two minus sequences.
22 The mutated strand is removed from the plus strand which can subsequently (or simultaneously)
23 anneal to the perfectly complementary sequence, in line with the ability of NCp7 to chaperone the
24 most stable NA conformations. NCp7 accelerated the overall process, leading to the accumulation
25 of the more stable complex. However, and probably because the strand exchange reaction can occur
26 spontaneously (although slowly), this exchange is also promoted by other probably nonspecific
27 RNA-binding compounds, including for example cationic comb-type copolymers.¹⁰¹⁻¹⁰³ A number
28 of more complex assays, like the hammerhead ribozyme cleavage or the inhibition of self-primed
29 cDNA synthesis, has also been set up to examine the properties of NA chaperone proteins.
30 However, being more integrated assays with multiple partners and reaction steps, the exact effect of
31 the chaperone proteins on each step is difficult to dissect.

32 Strong efforts have notably been performed to unravel the molecular mechanisms of the main
33 annealing reactions mandatory for reverse transcription. A detailed study of the kinetics of the
34 tRNA^{(Lys)3} annealing onto the PBS sequence showed that this reaction is a nucleation-limited

1 bimolecular reaction, enhanced by five orders of magnitude by NCp7. This annealing promotion
2 was attributed to the ability of NCp7 to induce a limited destabilization of the PBS secondary
3 structure and the tRNA^{(Lys)3} tertiary structure, as well as to its ability to strongly facilitate the duplex
4 nucleation.⁴³ Heteronuclear NMR measurements showed that the destabilization of the TΨC stem
5 within the tRNA^{(Lys)3} is necessary to obtain full annealing.¹⁰⁴ This destabilization is specific⁸⁸ and
6 cannot be reproduced using poly-L-lysine.⁷⁵ Moreover, NMR data suggested that the annealing
7 reaction is initiated by the PBS invasion of the four-way junction of the tRNA^{(Lys)3} cloverleaf. In
8 contrast, a mutational analysis of semi-synthetic tRNAs suggested that the annealing reaction
9 initiates at the 3'CCA end of the tRNA^{(Lys)3} acceptor stem. These data were reconciled in a refined
10 model in which both the centre of the cloverleaf and the 3'-end unpaired bases of the primer tRNA
11 are used for PBS annealing.¹⁰⁵ However, it is still unclear whether the same viral RNA molecule
12 can simultaneously invade the tRNA acceptor stem from both ends or whether some proteins
13 initiate the annealing at the tRNA centre while others “zip-in” from the 3'-end, in a statistical
14 manner.

15 During minus-strand DNA transfer, the antisense TAR DNA sequence of the nascent minus-strand
16 viral DNA must anneal with the sense TAR RNA sequence at the genome 3' end for RT to resume
17 the proviral DNA synthesis. In the absence of NCp7, the annealing rate of the TAR RNA/cTAR
18 hybridization reaction is extremely slow.^{42, 46-48} NCp7 chaperones this reaction, by enhancing its
19 rate by about 3000-fold in physiological temperature and salt conditions.^{41, 48, 88, 106-108} The
20 efficiency of the minus strand transfer reaction was initially envisioned as resulting from the extent
21 of homology between the two complementary sequences¹⁰⁹⁻¹¹⁴, although the structure of the
22 reactants was suspected to play a role, too.^{107, 109, 110, 115-117} This assumption was confirmed with the
23 relationship between the NA thermostability and the NC-mediated minus-strand transfer efficiency.
24¹¹⁰ In fact, more than the overall stability of the NA sequence, the stability of local structures
25 appeared to be of particular importance to modulate the strand transfer¹¹⁰, suggesting that a fine
26 thermodynamic balance between the NAs stability and the destabilizing activity of NCp7 is needed
27 for an efficient strand transfer. As a consequence, in complex annealing reactions such as the one
28 involved in the first strand transfer, both the NA destabilizing and annealing properties of NCp7 are
29 required and dissecting the exact contribution of each component in the promotion of this reaction
30 is challenging.

31 Thorough studies of the TAR RNA/cTAR hybridization showed apparent contradictory results in
32 respect with the order of the reaction. A first-order reaction was described for the NCp7-promoted
33 annealing of an 81nt TAR RNA to TAR DNA, suggesting that a slow unfolding of the TAR
34 secondary structure could act as the rate-limiting step of the reaction.⁴⁸ Further studies, using 55nt

1 (Mal Strain) TAR RNA and cTAR hairpins together with the truncated NC(12-55) mutant ⁴², or
2 using a 27nt or 59 nt (Lai Strain) cTAR hairpin ^{46, 47} in the presence of NCp7, showed
3 unambiguously that the protein-promoted annealing of these sequences followed second-order
4 kinetics. In fact, the NA aggregating properties of NCp7 produced large fluctuations of the
5 annealing rate, even inducing a “stalling effect” when the concentration of NCp7 is very high. ^{28, 44}
6 These rate fluctuations and the discrepancies in the order of the reaction kinetics probably resulted
7 from the formation of large multimolecular NA-NCp7 complexes. ^{52, 88, 118} This shows that the
8 chaperone properties of NA binding proteins need to be investigated over a large range of protein-
9 to-nucleotide ratios (typically from 1:100 to 1:1) and that great care should be taken to investigate
10 the annealing kinetics due to the exceedingly complex and heterogeneous nature of these reactions
11 in the presence of NA/protein aggregates. A study of the annealing of a 27-nt TAR RNA and DNA,
12 also called mini-TAR, corresponding to the top hairpin loop region of TAR showed that these two
13 sequences anneal *via* a two-step mechanism. The first bimolecular step leads to the formation of an
14 annealing intermediate which is subsequently converted into the final RNA/DNA duplex. DNA
15 mutagenesis data further showed that the annealing reaction is nucleated via an extended loop-loop
16 kissing interaction involving 17 intermolecular base pairs in the absence of NCp7. ⁴⁶ In the presence
17 of saturating levels of NCp7, an additional annealing pathway nucleated at the 3'/5' ends (zipping
18 pathway) was observed, but contributed only weakly to the reaction.

19 A detailed kinetic analysis using the full length TAR RNA and cTAR hairpins and the truncated
20 NC(12-55) concluded that this peptide facilitates the annealing of these sequences mainly via the
21 stem ends. ⁴² An alternative but minor pathway through TAR RNA and cTAR loops was also
22 observed. Single-molecule spectroscopy studies using full length TAR RNA and cTAR hairpins and
23 the full length NCp7 under non-aggregating conditions were consistent with this zipping pathway.
24 ¹¹⁹ Furthermore, the dependence of the TAR/cTAR annealing reaction on the NCp7 concentration
25 indicated that NCp7 shifts the annealing pathway from a loop-loop “kissing” to a “zipper” pathway
26 as the NCp7 concentration increases. Indeed, in the absence or at low level of NCp7, the TAR
27 DNA/RNA annealing proceeds via an extended loop-loop kissing intermediate, followed by a rate-
28 limiting strand exchange between the terminal stems, ^{46, 47, 120} while at high levels of NCp7, the two
29 sequences anneal through their stem ends. Such a mechanistic switch strongly suggests that the
30 reaction pathway is selected on the basis of the available reactive intermediates; the reaction
31 pathway being governed by the intermediates which require the least bp melting prior annealing. As
32 a consequence, the mechanism of the annealing reaction appears to be both sequence and protein-
33 concentration dependent, with a fine balance between the local stability of the NA sequences
34 involved in the reaction and the destabilization induced by NCp7 in these sequences. This

1 conclusion is fully supported by the NCp7-mediated switch in the annealing pathway of (-)PBS
2 with (+)PBS. NCp7 increases the (-)PBS/(+)PBS annealing kinetics by about six-fold by promoting
3 a loop-loop kissing pathway that poorly contributes to the annealing reaction in the absence of
4 NCp7.⁴⁵ This switch in the mechanistic pathway likely results from the NMR-observed
5 reorganisation and destabilization of the PBS loop structure in response to NCp7 binding.⁶ Thus,
6 multiple annealing pathways appear as a general principle for the NCp7-induced hybridization
7 reactions, which likely allows NCp7 to promote the multiple annealing reactions occurring during
8 reverse transcription as well as during forced or copy choice recombination events.¹²¹
9 Finally, it is worth to notice that if the general action of NCp7 is to strongly activate the annealing
10 reaction of complementary NA sequences^{122, 123}, NCp7 can also promote the reverse reaction.
11 Addition of an excess of NCp7 to small (<20 nt) or imperfect duplexes can result in their
12 destabilization.^{122, 123} NCp7 also dissociates imperfect duplexes of mutated PBS sequences that
13 form spontaneously in the absence of NCp7.⁴⁵ Moreover, NCp7 can promote the dissociation of
14 longer and perfectly complementary sequences, such as TAR DNA and cTAR.¹¹⁹ The NCp7-
15 induced reversion of the annealing reactions is thought to be required in reverse transcription,
16 notably to prevent non-specific self-primed cDNA synthesis.^{107, 124-128}

17

18 **Conclusion**

19

20 The key role of NCp7 for successful virus replication is now widely accepted, but still not fully
21 understood. The critical role of NCp7, notably during the early events of the replication cycle, is
22 thought to largely rely on its NA chaperone properties. Numerous studies have focused on the
23 characterization of the NCp7 chaperone properties and our understanding at the molecular level of
24 these properties has strongly gained from biophysical approaches. The NCp7-chaperoned folding of
25 NAs into their most stable conformations is characterized by both a NA destabilizing activity and a
26 complementary strand annealing activity. The destabilizing activity of NCp7 appears to be weak
27 and strongly dependent on the NA stability, resulting in a transient melting of “weak spots” within
28 the NA. The locally transiently-melted areas allow conformational sampling, resulting in the
29 selection of stable conformations, possibly stable enough to be no more susceptible to NCp7-
30 induced destabilization. NCp7 also promotes of the annealing of complementary NA sequences in a
31 concentration dependent manner. While the destabilization activity of NCp7 generates local single-
32 strand regions prone for nucleation, the ability of NCp7 to efficiently screen repulsive charges and
33 aggregate NAs, mainly through its basic residues, strongly enhances the limiting nucleation step. As
34 a consequence, even if the two activities are deeply interdependent and their exact contribution is

1 difficult to dissect, one can schematically assign the destabilization activity to the folded ZFs and
2 the annealing activity to the unfolded N_{term} domain and the basic linker.

3 Unravelling the mechanistic details at the molecular level of the main annealing reactions involved
4 in reverse transcription showed that the mechanistic pathways are multiple. These reactions involve
5 different nucleation sites but also probably various transition states. This complexity is not
6 completely surprising due to the rugged energy landscape of NA rearrangements. Recent
7 observations showed that NCp7 can “freeze” the DNA. As a consequence, the binding of NCp7
8 results in a loss of entropy, due to the decrease in the number of available NA conformations. In this
9 context, NCp7 can be referred to as a “guide” that funnels NAs toward their most stable
10 conformations by multiple routes within the conformational hyperspace. To prevent these routes
11 from local minima corresponding to intermediates, misfolded states and kinetic traps in the NA
12 rearrangement process, the binding of NCp7 is likely transient and dynamic in nature.
13 Improvements in computational predictions will probably help to provide insights into these
14 processes.

15 Though the biophysical characterization of the NCp7 chaperone activity provided detailed
16 information on its mechanism, a fully integrated understanding of the NCp7 chaperone mechanism
17 is still missing. In addition, studies conducted so far suffer from their reductionist approach.
18 Investigating the chaperone properties in more complex systems and notably in the viral context
19 will further improve our understanding of the chaperone activity and its biological relevance.
20 Again, sophisticated biophysical techniques and notably, fluorescence-based advanced techniques
21 should be able to reach this aim. Finally, the description at the molecular level of the mechanisms of
22 the chaperone activity can provide the basis for designing screening assays to search for compounds
23 able to inhibit the *in vivo* NCp7 activity and lead to anti-NC drugs.

24
25
26
27
28

1 References

- 2
- 3 1. Mély Y, Rocquigny HD, Morellet N, Roques BP, Gérard D. Zinc binding to the
- 4 HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence
- 5 spectroscopy. *Biochemistry* 1996; 35:5175--82.
- 6 2. Morellet N, de Rocquigny H, Mély Y, Jullian N, Déméné H, Ottmann M, Gérard
- 7 D, Darlix JL, Fournie-Zaluski MC, Roques BP. Conformational behaviour of the
- 8 active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR.
- 9 *Journal of molecular biology* 1994; 235:287--301.
- 10 3. Morellet N, Jullian N, Rocquigny HD, Maigret B, Darlix JL, Roques BP.
- 11 Determination of the structure of the nucleocapsid protein NCp7 from the human
- 12 immunodeficiency virus type 1 by 1H NMR. *The EMBO journal* 1992; 11:3059--65.
- 13 4. Summers MF, Henderson LE, Chance MR, Bess JW, South TL, Blake PR, Sagi I,
- 14 Perez-Alvarado G, Sowder RC, Hare DR. Nucleocapsid zinc fingers detected in
- 15 retroviruses: EXAFS studies of intact viruses and the solution-state structure of the
- 16 nucleocapsid protein from HIV-1. *Protein Sci* 1992; 1:563--74.
- 17 5. Amarasinghe GK, Zhou J, Miskimon M, Chancellor KJ, McDonald JA, Matthews
- 18 AG, Miller RR, Rouse MD, Summers MF. Stem-loop SL4 of the HIV-1 psi RNA
- 19 packaging signal exhibits weak affinity for the nucleocapsid protein. structural
- 20 studies and implications for genome recognition. *Journal of molecular biology* 2001;
- 21 314:961--70.
- 22 6. Bourbigot S, Ramalanjaona N, Boudier C, Salgado GFJ, Roques BP, Mély Y,
- 23 Bouaziz S, Morellet N. How the HIV-1 nucleocapsid protein binds and destabilises the
- 24 (-)primer binding site during reverse transcription. *Journal of molecular biology*
- 25 2008; 383:1112--28.
- 26 7. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF.
- 27 Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition
- 28 element. *Science (New York, NY)* 1998; 279:384-8.
- 29 8. Morellet N, Déméné H, Teilleux V, Huynh-Dinh T, de Rocquigny H, Fournié-
- 30 Zaluski MC, Roques BP. Structure of the complex between the HIV-1 nucleocapsid
- 31 protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *Journal of*
- 32 *molecular biology* 1998; 283:419--34.
- 33 9. Spriggs S, Garyu L, Connor R, Summers MF. Potential intra- and intermolecular
- 34 interactions involving the unique-5' region of the HIV-1 5'-UTR. *Biochemistry* 2008;
- 35 47:13064--73.
- 36 10. Bombarda E, Ababou A, Vuilleumier C, Gérard D, Roques BP, Piémont E, Mély
- 37 Y. Time-resolved fluorescence investigation of the human immunodeficiency virus
- 38 type 1 nucleocapsid protein: influence of the binding of nucleic acids. *Biophys J*
- 39 1999; 76:1561--70.
- 40 11. Cimarelli A, Darlix JL. Assembling the human immunodeficiency virus type 1.
- 41 *Cell Mol Life Sci* 2002; 59:1166-84.
- 42 12. Muriaux D, Darlix JL, Cimarelli A. Targeting the assembly of the human
- 43 immunodeficiency virus type 1. *Curr Pharm Des* 2004; 10:3725-39.
- 44 13. Amarasinghe GK, De Guzman RN, Turner RB, Summers MF. NMR structure of
- 45 stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-
- 46 triple platform. *Journal of molecular biology* 2000; 299:145-56.
- 47 14. Shubsda MF, Paoletti AC, Hudson BS, Borer PN. Affinities of packaging
- 48 domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry* 2002;
- 49 41:5276--82.
- 50 15. Fisher RJ, Rein A, Fivash M, Urbaneja MA, Casas-Finet JR, Medaglia M,
- 51 Henderson LE. Sequence-specific binding of human immunodeficiency virus type 1
- 52 nucleocapsid protein to short oligonucleotides. *J Virol* 1998; 72:1902--9.
- 53 16. Mély Y, de Rocquigny H, Sorinas-Jimeno M, Keith G, Roques BP, Marquet R,
- 54 Gérard D. Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3Lys), in
- 55 vitro, is essentially not specific. *The Journal of biological chemistry* 1995; 270:1650-
- 56 -6.
- 57 17. Urbaneja MA, Kane BP, Johnson DG, Gorelick RJ, Henderson LE, Casas-Finet
- 58 JR. Binding properties of the human immunodeficiency virus type 1 nucleocapsid
- 59 protein p7 to a model RNA: elucidation of the structural determinants for function.
- 60 *Journal of molecular biology* 1999; 287:59--75.

- 1 18. Vuilleumier C, Bombarda E, Morellet N, Gérard D, Roques BP, Mély Y. Nucleic
2 acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence
3 study. *Biochemistry* 1999; 38:16816--25.
- 4 19. Goldschmidt V, Miller Jenkins LM, De Rocquigny H, Darlix JL, Mély Y. The
5 nucleocapsid protein of HIV-1 as a promising therapeutic target for antiviral drugs.
6 *HIV Therapy* 2010; 4:179-98.
- 7 20. Musah RA. The HIV-1 nucleocapsid zinc finger protein as a target of
8 antiretroviral therapy. *Curr Top Med Chem* 2004; 4:1605-22.
- 9 21. Aldovini A, Young RA. Mutations of RNA and protein sequences involved in
10 human immunodeficiency virus type 1 packaging result in production of noninfectious
11 virus. *J Virol* 1990; 64:1920--6.
- 12 22. Déméné H, Dong CZ, Ottmann M, Rouyez MC, Jullian N, Morellet N, Mely Y,
13 Darlix JL, Fournié-Zaluski MC, Saragosti S. 1H NMR structure and biological studies
14 of the His23-->Cys mutant nucleocapsid protein of HIV-1 indicate that the
15 conformation of the first zinc finger is critical for virus infectivity. *Biochemistry*
16 1994; 33:11707--16.
- 17 23. Dorfman T, Luban J, Goff SP, Haseltine WA, Gottlinger HG. Mapping of
18 functionally important residues of a cysteine-histidine box in the human
19 immunodeficiency virus type 1 nucleocapsid protein. *J Virol* 1993; 67:6159-69.
- 20 24. Gorelick RJ, Nigida SM, Bess JW, Arthur LO, Henderson LE, Rein A.
21 Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic
22 RNA. *J Virol* 1990; 64:3207--11.
- 23 25. Ottmann M, Gabus C, Darlix JL. The central globular domain of the
24 nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion
25 structure and infectivity. *J Virol* 1995; 69:1778--84.
- 26 26. Turner BG, Summers MF. Structural biology of HIV. *Journal of molecular*
27 *biology* 1999; 285:1--32.
- 28 27. Bampi C, Jacquenet S, Lener D, Decimo D, Darlix JL. The chaperoning and
29 assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and
30 maintenance. *Curr HIV Res* 2004; 2:79-92.
- 31 28. Cristofari G, Darlix JL. The ubiquitous nature of RNA chaperone proteins. *Prog*
32 *Nucleic Acid Res Mol Biol* 2002; 72:223-68.
- 33 29. Darlix JL, Lapadat-Tapolsky M, de Rocquigny H, Roques BP. First glimpses at
34 structure-function relationships of the nucleocapsid protein of retroviruses. *Journal*
35 *of molecular biology* 1995; 254:523-37.
- 36 30. Herschlag D. RNA chaperones and the RNA folding problem. *The Journal of*
37 *biological chemistry* 1995; 270:20871-4.
- 38 31. Rein A, Henderson LE, Levin JG. Nucleic-acid-chaperone activity of retroviral
39 nucleocapsid proteins: significance for viral replication. *Trends Biochem Sci* 1998;
40 23:297-301.
- 41 32. Tsuchihashi Z, Brown PO. DNA strand exchange and selective DNA annealing
42 promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol*
43 1994; 68:5863--70.
- 44 33. Levin JG, Guo J, Rouzina I, Musier-Forsyth K. Nucleic acid chaperone activity
45 of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular
46 mechanism. *Prog Nucleic Acid Res Mol Biol* 2005; 80:217--86.
- 47 34. Beltz H, Azoulay J, Bernacchi S, Clamme J-P, Ficheux D, Roques B, Darlix J-L,
48 Mély Y. Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the
49 destabilizing activity of the nucleocapsid protein NCp7. *Journal of molecular biology*
50 2003; 328:95--108.
- 51 35. Beltz H, Piémont E, Schaub E, Ficheux D, Roques B, Darlix J-L, Mély Y. Role
52 of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing
53 activity of the nucleocapsid protein NCp7. *Journal of molecular biology* 2004;
54 338:711--23.
- 55 36. Bernacchi S, Stoylov S, Piémont E, Ficheux D, Roques BP, Darlix JL, Mély Y.
56 HIV-1 nucleocapsid protein activates transient melting of least stable parts of the
57 secondary structure of TAR and its complementary sequence. *Journal of molecular*
58 *biology* 2002; 317:385--99.
- 59 37. Cosa G, Zeng Y, Liu H-W, Landes CF, Makarov DE, Musier-Forsyth K, Barbara
60 PF. Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned
61 opening of DNA hairpins. *J Phys Chem B* 2006; 110:2419--26.
- 62 38. Egelé C, Schaub E, Ramalanjaona N, Piémont E, Ficheux D, Roques B, Darlix

1 J-L, Mély Y. HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences
2 and chaperones their kissing interactions. *Journal of molecular biology* 2004;
3 342:453--66.
4 39. Liu H-W, Cosa G, Landes CF, Zeng Y, Kovaleski BJ, Mullen DG, Barany G,
5 Musier-Forsyth K, Barbara PF. Single-molecule FRET studies of important
6 intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in
7 HIV-1 reverse transcription. *Biophys J* 2005; 89:3470--9.
8 40. Williams MC, Rouzina I, Wenner JR, Gorelick RJ, Musier-Forsyth K,
9 Bloomfield VA. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid
10 protein revealed by single molecule stretching. *Proc Natl Acad Sci U S A* 2001;
11 98:6121--6.
12 41. Darlix JL, Vincent A, Gabus C, de Rocquigny H, Roques B. Trans-activation of
13 the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse
14 transcription of HIV1 RNA. *C R Acad Sci III* 1993; 316:763-71.
15 42. Godet J, de Rocquigny H, Raja C, Glasser N, Ficheux D, Darlix J-L, Mély Y.
16 During the early phase of HIV-1 DNA synthesis, nucleocapsid protein directs
17 hybridization of the TAR complementary sequences via the ends of their double-
18 stranded stem. *Journal of molecular biology* 2006; 356:1180--92.
19 43. Hargittai MRS, Gorelick RJ, Rouzina I, Musier-Forsyth K. Mechanistic insights
20 into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the
21 primer binding site. *Journal of molecular biology* 2004; 337:951--68.
22 44. Liu H-W, Zeng Y, Landes CF, Kim YJ, Zhu Y, Ma X, Vo M-N, Musier-Forsyth K,
23 Barbara PF. Insights on the role of nucleic acid/protein interactions in chaperoned
24 nucleic acid rearrangements of HIV-1 reverse transcription. *Proc Natl Acad Sci U S A*
25 2007; 104:5261--7.
26 45. Ramalanjaona N, de Rocquigny H, Millet A, Ficheux D, Darlix J-L, Mély Y.
27 Investigating the mechanism of the nucleocapsid protein chaperoning of the second
28 strand transfer during HIV-1 DNA synthesis. *Journal of molecular biology* 2007;
29 374:1041--53.
30 46. Vo M-N, Barany G, Rouzina I, Musier-Forsyth K. Mechanistic studies of mini-
31 TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein.
32 *Journal of molecular biology* 2006; 363:244--61.
33 47. Vo M-N, Barany G, Rouzina I, Musier-Forsyth K. HIV-1 nucleocapsid protein
34 switches the pathway of transactivation response element RNA/DNA annealing from
35 loop-loop "kissing" to "zipper". *Journal of molecular biology* 2009; 386:789--801.
36 48. You JC, McHenry CS. Human immunodeficiency virus nucleocapsid protein
37 accelerates strand transfer of the terminally redundant sequences involved in reverse
38 transcription. *The Journal of biological chemistry* 1994; 269:31491--5.
39 49. Urbaneja MA, Wu M, Casas-Finet JR, Karpel RL. HIV-1 nucleocapsid protein as
40 a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties,
41 structural binding specificity, and annealing activity. *Journal of molecular biology*
42 2002; 318:749--64.
43 50. Azoulay J, Clamme JP, Darlix JL, Roques BP, Mély Y. Destabilization of the
44 HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation
45 of conformational fluctuations. *Journal of molecular biology* 2003; 326:691--700.
46 51. Bernacchi S, Mely Y. Excitonic interaction in molecular beacons: a sensitive
47 sensor for short range modifications of the nucleic acid structure. *Nucleic acids*
48 *research* 2001; 29:E62-2.
49 52. Stoylov SP, Vuilleumier C, Stoylova E, De Rocquigny H, Roques BP, Gerard D,
50 Mely Y. Ordered aggregation of ribonucleic acids by the human immunodeficiency
51 virus type 1 nucleocapsid protein. *Biopolymers* 1997; 41:301-12.
52 53. Livesey AK, Brochon JC. Analyzing the Distribution of Decay Constants in
53 Pulse-Fluorimetry Using the Maximum Entropy Method. *Biophys J* 1987; 52:693--706.
54 54. Hong MK, Harbron EJ, O'Connor DB, Guo J, Barbara PF, Levin JG, Musier-
55 Forsyth K. Nucleic acid conformational changes essential for HIV-1 nucleocapsid
56 protein-mediated inhibition of self-priming in minus-strand transfer. *Journal of*
57 *molecular biology* 2003; 325:1--10.
58 55. Bernacchi S, Piémont E, Potier N, van Dorsselaer A, Mély Y. Excitonic
59 heterodimer formation in an HIV-1 oligonucleotide labeled with a donor-acceptor pair
60 used for fluorescence resonance energy transfer. *Biophys J* 2003; 84:643--54.
61 56. Cosa G, Harbron EJ, Zeng Y, Liu H-W, O'Connor DB, Eta-Hosokawa C, Musier-
62 Forsyth K, Barbara PF. Secondary structure and secondary structure dynamics of DNA

1 hairpins complexed with HIV-1 NC protein. *Biophys J* 2004; 87:2759--67.

2 57. Beltz H, Clauss C, Piémont E, Ficheux D, Gorelick RJ, Roques B, Gabus C,
3 Darlix J-L, de Rocquigny H, Mély Y. Structural determinants of HIV-1 nucleocapsid
4 protein for cTAR DNA binding and destabilization, and correlation with inhibition of
5 self-primed DNA synthesis. *Journal of molecular biology* 2005; 348:1113--26.

6 58. Hausteine E, Schwille P. Fluorescence correlation spectroscopy: novel variations
7 of an established technique. *Annu Rev Biophys Biomol Struct* 2007; 36:151-69.

8 59. Egelé C, Schaub E, Piémont E, de Rocquigny H, Mély Y. Investigation by
9 fluorescence correlation spectroscopy of the chaperoning interactions of HIV-1
10 nucleocapsid protein with the viral DNA initiation sequences. *C R Biol* 2005;
11 328:1041--51.

12 60. Shvadchak V, Sanglier S, Rocle S, Villa P, Haiech J, Hibert M, Dorsselaer AV,
13 Mély Y, de Rocquigny H. Identification by high throughput screening of small
14 compounds inhibiting the nucleic acid destabilization activity of the HIV-1
15 nucleocapsid protein. *Biochimie* 2009.

16 61. Williams MC, Gorelick RJ, Musier-Forsyth K. Specific zinc-finger architecture
17 required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc Natl*
18 *Acad Sci U S A* 2002; 99:8614--9.

19 62. Smith SB, Cui Y, Bustamante C. Optical-trap force transducer that operates by
20 direct measurement of light momentum. *Methods Enzymol* 2003; 361:134-62.

21 63. Wang MD, Yin H, Landick R, Gelles J, Block SM. Stretching DNA with optical
22 tweezers. *Biophys J* 1997; 72:1335-46.

23 64. Rouzina I, Bloomfield VA. Force-induced melting of the DNA double helix. 2.
24 Effect of solution conditions. *Biophys J* 2001; 80:894-900.

25 65. Wenner JR, Williams MC, Rouzina I, Bloomfield VA. Salt dependence of the
26 elasticity and overstretching transition of single DNA molecules. *Biophys J* 2002;
27 82:3160--9.

28 66. Williams MC, Rouzina I. Force spectroscopy of single DNA and RNA
29 molecules. *Curr Opin Struct Biol* 2002; 12:330--6.

30 67. Williams MC, Rouzina I, Bloomfield VA. Thermodynamics of DNA interactions
31 from single molecule stretching experiments. *Acc Chem Res* 2002; 35:159--66.

32 68. Williams MC, Wenner JR, Rouzina I, Bloomfield VA. Entropy and heat capacity
33 of DNA melting from temperature dependence of single molecule stretching. *Biophys*
34 *J* 2001; 80:1932--9.

35 69. Williams MC, Wenner JR, Rouzina I, Bloomfield VA. Effect of pH on the
36 overstretching transition of double-stranded DNA: evidence of force-induced DNA
37 melting. *Biophys J* 2001; 80:874--81.

38 70. Cantor CR, Schimmel PR, eds. *Biophysical Chemistry*: Freeman, New York,
39 1980.

40 71. Cruceanu M, Urbaneja MA, Hixson CV, Johnson DG, Datta SA, Fivash MJ,
41 Stephen AG, Fisher RJ, Gorelick RJ, Casas-Finet JR, Rein A, Rouzina I, Williams MC.
42 Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid
43 proteins. *Nucleic acids research* 2006; 34:593--605.

44 72. Cruceanu M, Stephen AG, Beuning PJ, Gorelick RJ, Fisher RJ, Williams MC.
45 Single DNA molecule stretching measures the activity of chemicals that target the
46 HIV-1 nucleocapsid protein. *Anal Biochem* 2006; 358:159--70.

47 73. Omichinski JG, Clore GM, Sakaguchi K, Appella E, Gronenborn AM. Structural
48 characterization of a 39-residue synthetic peptide containing the two zinc binding
49 domains from the HIV-1 p7 nucleocapsid protein by CD and NMR spectroscopy. *FEBS*
50 *Lett* 1991; 292:25--30.

51 74. Johnson PE, Turner RB, Wu ZR, Hairston L, Guo J, Levin JG, Summers MF. A
52 mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein
53 during reverse transcription. *Biochemistry* 2000; 39:9084--91.

54 75. Tisne C, Roques BP, Dardel F. Heteronuclear NMR studies of the interaction of
55 tRNA(Lys)₃ with HIV-1 nucleocapsid protein. *Journal of molecular biology* 2001;
56 306:443-54.

57 76. Gregoire CJ, Gautheret D, Loret EP. No tRNA³Lys unwinding in a complex with
58 HIV NCp7. *The Journal of biological chemistry* 1997; 272:25143--8.

59 77. Hargittai MR, Mangla AT, Gorelick RJ, Musier-Forsyth K. HIV-1 nucleocapsid
60 protein zinc finger structures induce tRNA(Lys,3) structural changes but are not
61 critical for primer/template annealing. *Journal of molecular biology* 2001; 312:985--
62 97.

- 1 78. Avilov S, Godet J, Piemont E, Mély Y. Site-specific characterization of HIV-1
2 nucleocapsid protein binding to oligonucleotides with two binding sites. *Biochemistry*
3 2009.
- 4 79. Avilov SV, Piemont E, Shvadchak V, de Rocquigny H, Mely Y. Probing
5 dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-
6 aminopurine. *Nucleic acids research* 2008; 36:885-96.
- 7 80. Nordlund TM, Andersson S, Nilsson L, Rigler R, Graslund A, McLaughlin LW.
8 Structure and dynamics of a fluorescent DNA oligomer containing the EcoRI
9 recognition sequence: fluorescence, molecular dynamics, and NMR studies.
10 *Biochemistry* 1989; 28:9095-103.
- 11 81. Sowers LC, Fazakerley GV, Eritja R, Kaplan BE, Goodman MF. Base pairing
12 and mutagenesis: observation of a protonated base pair between 2-aminopurine and
13 cytosine in an oligonucleotide by proton NMR. *Proc Natl Acad Sci U S A* 1986;
14 83:5434-8.
- 15 82. Rist M, Marino J. Association of an RNA kissing complex analyzed using 2-
16 aminopurine fluorescence. *Nucleic acids research* 2001; 29:2401--8.
- 17 83. Rajkowitsch L, Semrad K, Mayer O, Schroeder R. Assays for the RNA
18 chaperone activity of proteins. *Biochem Soc Trans* 2005; 33:450--6.
- 19 84. Cardullo RA, Agrawal S, Flores C, Zamecnik PC, Wolf DE. Detection of nucleic
20 acid hybridization by nonradiative fluorescence resonance energy transfer. *Proc Natl*
21 *Acad Sci U S A* 1988; 85:8790--4.
- 22 85. Rajkowitsch L, Schroeder R. Dissecting RNA chaperone activity. *RNA* 2007;
23 13:2053--60.
- 24 86. Barat C, Lullien V, Schatz O, Keith G, Nugeyre MT, Gruninger-Leitch F, Barre-
25 Sinoussi F, LeGrice SF, Darlix JL. HIV-1 reverse transcriptase specifically interacts
26 with the anticodon domain of its cognate primer tRNA. *The EMBO journal* 1989;
27 8:3279-85.
- 28 87. Dib-Hajj F, Khan R, Giedroc DP. Retroviral nucleocapsid proteins possess
29 potent nucleic acid strand renaturation activity. *Protein Sci* 1993; 2:231--43.
- 30 88. Lapadat-Tapolsky M, Pernelle C, Borie C, Darlix JL. Analysis of the nucleic
31 acid annealing activities of nucleocapsid protein from HIV-1. *Nucleic acids research*
32 1995; 23:2434-41.
- 33 89. Heath MJ, Derebail SS, Gorelick RJ, DeStefano JJ. Differing roles of the N-
34 and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-
35 enhanced nucleic acid annealing. *The Journal of biological chemistry* 2003;
36 278:30755--63.
- 37 90. Pontius BW, Berg P. Renaturation of complementary DNA strands mediated by
38 purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: implications
39 for a mechanism for rapid molecular assembly. *Proc Natl Acad Sci U S A* 1990;
40 87:8403--7.
- 41 91. Pontius BW, Berg P. Rapid renaturation of complementary DNA strands
42 mediated by cationic detergents: a role for high-probability binding domains in
43 enhancing the kinetics of molecular assembly processes. *Proc Natl Acad Sci U S A*
44 1991; 88:8237--41.
- 45 92. Echols H. Nucleoprotein structures initiating DNA replication, transcription,
46 and site-specific recombination. *The Journal of biological chemistry* 1990;
47 265:14697-700.
- 48 93. Boudier C, Storchak R, Sharma KK, Didier P, Follenius-Wund A, Muller S,
49 Darlix JL, Mély Y. The mechanism of HIV-1 Tat-directed nucleic acid annealing
50 supports its role in reverse transcription. *Journal of molecular biology*.
- 51 94. Sharma K, Didier P, Darlix JL, de Rocquigny H, Bensikaddour H, Lavergne JP,
52 Penin F, Lessinger JM, Mely Y. Kinetic analysis of the nucleic acid chaperone
53 activity of the hepatitis C virus core protein. *Nucleic acids research* 2010; 38:3632-
54 42.
- 55 95. Nedbal W, Frey M, Willemann B, Zentgraf H, Sczakiel G. Mechanistic insights
56 into p53-promoted RNA-RNA annealing. *Journal of molecular biology* 1997; 266:677-
57 87.
- 58 96. Gabus C, Derrington E, Leblanc P, Chnaiderman J, Dormont D, Swietnicki W,
59 Morillas M, Surewicz WK, Marc D, Nandi P, Darlix JL. The prion protein has RNA
60 binding and chaperoning properties characteristic of nucleocapsid protein NCP7 of
61 HIV-1. *The Journal of biological chemistry* 2001; 276:19301--9.
- 62 97. Vo M-N, Barany G, Rouzina I, Musier-Forsyth K. Effect of Mg(2+) and Na(+)

1 on the nucleic acid chaperone activity of HIV-1 nucleocapsid protein: implications for
2 reverse transcription. *Journal of molecular biology* 2009; 386:773--88.
3 98. Sikorav JL, Church GM. Complementary recognition in condensed DNA:
4 accelerated DNA renaturation. *Journal of molecular biology* 1991; 222:1085--108.
5 99. Brulé F, Marquet R, Rong L, Wainberg MA, Roques BP, Grice SFJL, Ehresmann
6 B, Ehresmann C. Structural and functional properties of the HIV-1 RNA-tRNA(Lys)₃
7 primer complex annealed by the nucleocapsid protein: comparison with the heat-
8 annealed complex. *RNA* 2002; 8:8--15.
9 100. Nedbal W, Homann M, Sczakiel G. The association of complementary
10 ribonucleic acids can be strongly increased without lowering Arrhenius activation
11 energies or significantly altering structures. *Biochemistry* 1997; 36:13552--7.
12 101. Kim WJ, Akaike T, Maruyama A. DNA strand exchange stimulated by
13 spontaneous complex formation with cationic comb-type copolymer. *J Am Chem Soc*
14 2002; 124:12676--7.
15 102. Kim WJ, Ishihara T, Akaike T, Maruyama A. Polycation graft copolymers
16 accelerating DNA strand exchange: involvement of ionic interaction. *Nucleic Acids*
17 *Res Suppl* 2001:151--2.
18 103. Kim WJ, Ishihara T, Akaike T, Maruyama A. Comb-type cationic copolymer
19 expedites DNA strand exchange while stabilizing DNA duplex. *Chemistry* 2001;
20 7:176--80.
21 104. Tisne C, Roques BP, Dardel F. The annealing mechanism of HIV-1 reverse
22 transcription primer onto the viral genome. *The Journal of biological chemistry* 2004;
23 279:3588-95.
24 105. Barraud P, Gaudin C, Dardel F, Tisné C. New insights into the formation of
25 HIV-1 reverse transcription initiation complex. *Biochimie* 2007; 89:1204--10.
26 106. Allain B, Lapadat-Tapolsky M, Berlioz C, Darlix JL. Transactivation of the
27 minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the
28 retroviral genome. *The EMBO journal* 1994; 13:973-81.
29 107. Guo J, Henderson LE, Bess J, Kane B, Levin JG. Human immunodeficiency
30 virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral
31 DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-
32 stop DNA. *J Virol* 1997; 71:5178--88.
33 108. Tanchou V, Gabus C, Rogemond V, Darlix JL. Formation of stable and
34 functional HIV-1 nucleoprotein complexes in vitro. *Journal of molecular biology*
35 1995; 252:563-71.
36 109. Berkhout B, van Wamel J, Klaver B. Requirements for DNA strand transfer
37 during reverse transcription in mutant HIV-1 virions. *Journal of molecular biology*
38 1995; 252:59--69.
39 110. Heilman-Miller SL, Wu T, Levin JG. Alteration of nucleic acid structure and
40 stability modulates the efficiency of minus-strand transfer mediated by the HIV-1
41 nucleocapsid protein. *The Journal of biological chemistry* 2004; 279:44154--65.
42 111. Klaver B, Berkhout B. Premature strand transfer by the HIV-1 reverse
43 transcriptase during strong-stop DNA synthesis. *Nucleic acids research* 1994; 22:137-
44 -44.
45 112. Lobel LI, Goff SP. Reverse transcription of retroviral genomes: mutations in
46 the terminal repeat sequences. *J Virol* 1985; 53:447--55.
47 113. Luo GX, Chao M, Hsieh SY, Sureau C, Nishikura K, Taylor J. A specific base
48 transition occurs on replicating hepatitis delta virus RNA. *J Virol* 1990; 64:1021-7.
49 114. Ramsey CA, Panganiban AT. Replication of the retroviral terminal repeat
50 sequence during in vivo reverse transcription. *J Virol* 1993; 67:4114--21.
51 115. van Wamel JL, Berkhout B. The first strand transfer during HIV-1 reverse
52 transcription can occur either intramolecularly or intermolecularly. *Virology* 1998;
53 244:245--51.
54 116. Chen Y, Balakrishnan M, Roques BP, Bambara RA. Steps of the acceptor
55 invasion mechanism for HIV-1 minus strand strong stop transfer. *The Journal of*
56 *biological chemistry* 2003; 278:38368-75.
57 117. Chen Y, Balakrishnan M, Roques BP, Fay PJ, Bambara RA. Mechanism of minus
58 strand strong stop transfer in HIV-1 reverse transcription. *The Journal of biological*
59 *chemistry* 2003; 278:8006-17.
60 118. Le Cam E, Coulaud D, Delain E, Petitjean P, Roques BP, Gerard D, Stoylova E,
61 Vuilleumier C, Stoylov SP, Mely Y. Properties and growth mechanism of the ordered
62 aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron

1 microscopy investigation. *Biopolymers* 1998; 45:217-29.
2 119. Zeng Y, Liu HW, Landes CF, Kim YJ, Ma X, Zhu Y, Musier-Forsyth K, Barbara
3 PF. Single-molecule Chemistry and Biology Special Feature: Probing nucleation,
4 reverse annealing, and chaperone function along the reaction path of HIV-1 single-
5 strand transfer. *Proc Natl Acad Sci U S A* 2007; 104:12651-6.
6 120. Kanevsky I, Chaminade F, Ficheux D, Moumen A, Gorelick R, Negroni M,
7 Darlix J-L, Fossé P. Specific interactions between HIV-1 nucleocapsid protein and the
8 TAR element. *Journal of molecular biology* 2005; 348:1059--77.
9 121. Basu VP, Song M, Gao L, Rigby ST, Hanson MN, Bambara RA. Strand transfer
10 events during HIV-1 reverse transcription. *Virus Res* 2008; 134:19-38.
11 122. Khan R, Giedroc DP. Recombinant human immunodeficiency virus type 1
12 nucleocapsid (NCp7) protein unwinds tRNA. *The Journal of biological chemistry*
13 1992; 267:6689--95.
14 123. Li X, Quan Y, Arts EJ, Li Z, Preston BD, de Rocquigny H, Roques BP, Darlix
15 JL, Kleiman L, Parniak MA, Wainberg MA. Human immunodeficiency virus Type 1
16 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis
17 primed by human tRNA(Lys3) in vitro: studies of viral RNA molecules mutated in
18 regions that flank the primer binding site. *J Virol* 1996; 70:4996-5004.
19 124. Arts EJ, Wainberg MA. Human immunodeficiency virus type 1 reverse
20 transcriptase and early events in reverse transcription. *Adv Virus Res* 1996; 46:97-
21 163.
22 125. Golinelli M-P, Hughes SH. Secondary structure in the nucleic acid affects the
23 rate of HIV-1 nucleocapsid-mediated strand annealing. *Biochemistry* 2003; 42:8153--
24 62.
25 126. Guo J, Wu T, Anderson J, Kane BF, Johnson DG, Gorelick RJ, Henderson LE,
26 Levin JG. Zinc finger structures in the human immunodeficiency virus type 1
27 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J Virol* 2000;
28 74:8980-8.
29 127. Lapadat-Tapolsky M, Gabus C, Rau M, Darlix JL. Possible roles of HIV-1
30 nucleocapsid protein in the specificity of proviral DNA synthesis and in its
31 variability. *Journal of molecular biology* 1997; 268:250-60.
32 128. Driscoll MD, Hughes SH. Human immunodeficiency virus type 1 nucleocapsid
33 protein can prevent self-priming of minus-strand strong stop DNA by promoting the
34 annealing of short oligonucleotides to hairpin sequences. *J Virol* 2000; 74:8785-92.
35
36

1

2 **Figure legends**

3

4 **Figure 1. Sequence (A) and 3D structure (B) of NCp7.**

5 NCp7 is a basic protein of 55 aminoacids with two zinc fingers, ZF₁ and ZF₂ (green) separated by a
6 highly basic linker (orange), and flanked by poorly folded N_{terminal} (grey) and C_{terminal} (deep purple)
7 domains. The 3D structure is derived from NMR data (**PDB 1AAF**⁴)

8 **Figure 2. Investigation of cTAR conformations by time-resolved fluorescence measurements.**

9 A. TAR RNA and cTAR secondary structures (sequence from the Mal strain).

10 B. Time-resolved fluorescence decays of the singly-labelled cTAR-3'-FAM (green) and the
11 doubly-labelled TMR-5'-cTAR-3'-FAM (blue), together with the instrumental response
12 function (FWHM ~40ps).

13 C. TCSPC principle. The sample is excited with a pulsed laser at low intensity, so that the
14 sample emits no more than one photon at each pulse, as illustrated by the output signal of
15 the detector. For each detected photon, the time between the excitation pulse and the photon
16 detection is measured. These individual times are channelled to reconstitute the original
17 fluorescence decay by counting thousands of photons.

18 D. Distribution of lifetimes retrieved from the fluorescence decay of the doubly-labelled cTAR
19 in the presence of NCp7.

20 E. Amplitudes associated to the different lifetimes in the absence (black) and in the presence
21 (grey) of NCp7 added at a ratio of one NCp7 molecule per 5 nt. The dark species correspond
22 to conformations characterized by a fluorescence lifetime shorter than the time-resolution of
23 the set-up (20 ps). NCp7 shifts the existing conformations towards the longer lived ones.

24 F. Scheme of the cTAR conformations associated to each lifetime (adapted from ³⁶).

25 **Figure 3. Characterization by fluorescence correlation spectroscopy (FCS) of the dynamics of** 26 **cTAR fraying.**

27 A. FCS is based on the analysis by a temporal autocorrelation function of the fluorescence intensity
28 fluctuations observed within a sub-femtoL focal volume. In solution, fluorescence fluctuations
29 mainly result from the Brownian diffusion of labelled molecules through the focal volume. The
30 apparent diffusion time (τ_D) and the average number of fluorescent molecules (N) within the
31 excitation volume can be extracted from the autocorrelation curve.

32 B. Autocorrelation curves of singly- and doubly-labelled cTAR molecules.

33 C. Ratio of the two autocorrelation curves shown in B. The G*(tau) curve describes the
34 fluorescence fluctuations due to cTAR fraying. The black line is the fit that allows recovering
35 the opening and closing rate constants of the fraying mechanism.

36

1

2 **Figure 4. Characterization by single-molecule FRET of cTAR destabilization by NCp7.**

3 The conformational distribution and fraying dynamics of the cTAR hairpin is explored by recording
4 the FRET trace between the donor and the acceptor of single immobilized doubly-labelled cTAR
5 hairpins.

6 A. Scheme of the immobilization process. The DNA hairpins are immobilized onto the glass slide
7 using the biotin/streptavidin chemistry. In presence of NCp7, cTAR hairpins coexist under a
8 closed conformation (state 1) and open (state 2) conformations.

9 B. Transitions between the closed ($E_{app} \sim 1$) and open ($E_{app} = 0.77$) conformations can be observed
10 through the time trajectory of the FRET efficiency (adapted from ⁵⁶). The interval defined
11 between the two horizontal blue lines corresponds to $E_{app}(\text{state 2}) \pm 2.SD$.

12 **Figure 5. Characterization of the NA destabilizing activity of NCp7 by single DNA molecule**
13 **stretching experiments.**

14 A. Schematic drawing of an optical tweezer experiment. A λ -DNA is stretched between two
15 polystyrene beads. One bead is held in an optical trap while the other is held on the end of a
16 micropipette. Displacement of the pipette to move the beads apart induces a helix to coil
17 transition of the DNA. The extension of the DNA and the force applied on the trapped bead are
18 simultaneously recorded.

19 B. Force-extension curves for single molecules of dsDNA (red triangles) and ssDNA (black line).
20 The theoretical curve for dsDNA is shown (green line). The overstretching transition appears to
21 be a transition from dsDNA to ssDNA. The ΔG value of the transition is given by the area
22 between the stretching curves for double- and single-stranded DNA (grey area) (adapted from
23 ⁴⁰).

24 **Figure 6. FRET-based annealing (A) and strand-displacement (B) assay to investigate the**
25 **NCp7 annealing activity.**

26 A. The two complementary sequences are fluorescently labelled. On NCp7-promoted duplex
27 formation, the fluorescence donor and acceptor become close enough for FRET to occur. The
28 increase in energy transfer is proportional to the fraction of formed duplexes.

29 B. Two non fully-complementary sequences are heat-annealed to form a doubly-labelled imperfect
30 duplex. The competitor (i.e. the perfectly complementary sequence to the ssDNA labelled with
31 the donor) is added in the presence of NCp7. The restoration of the donor fluorescence allows
32 following the strand-exchange reaction in real time.

33 **Figure 7. Schematic model of the tRNA^{Lys3} annealing reaction to the PBS (adapted from ¹⁰⁴).**

34 A. tRNA^{Lys3} secondary and tertiary structure.

35 B. NCp7-facilitated tRNA^{Lys3} nucleation to the PBS by its 3'CCA end and/or by the centre of
36 the cloverleaf.

1 C. Destabilization of the tertiary interactions within the tRNA^{Lys3}, catalysed by NCp7.

2 D. Formation of the fully annealed duplex.

3 **Figure 8. Schematic model of the TAR/cTAR annealing.**

4 In the absence of NCp7, the TAR/cTAR duplex formation is initiated through a loop-loop
5 interaction, resulting in the fast formation of a kissing complex intermediate which is subsequently
6 converted into the final extended duplex. NCp7 destabilizes cTAR and shifts the conformation
7 equilibrium towards species where the lower half of the stem is melted. Then, the partly melted
8 cTAR species readily nucleate with TAR through their stem termini. Thus, NCp7 strongly enhances
9 the annealing reaction by promoting the TAR/cTAR duplex formation mainly through the zipping
10 pathway.

11

12

13

14

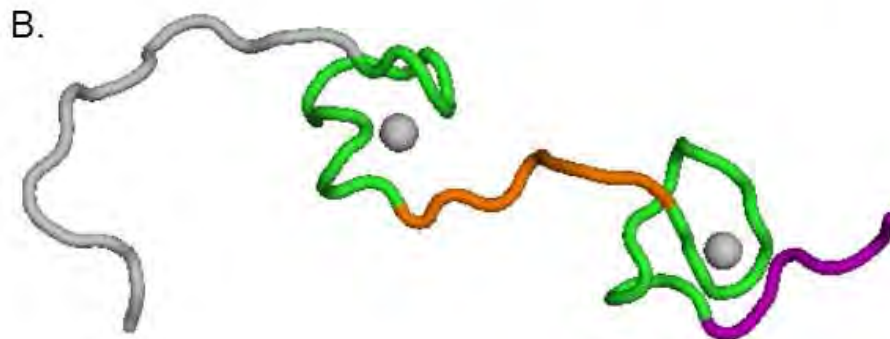
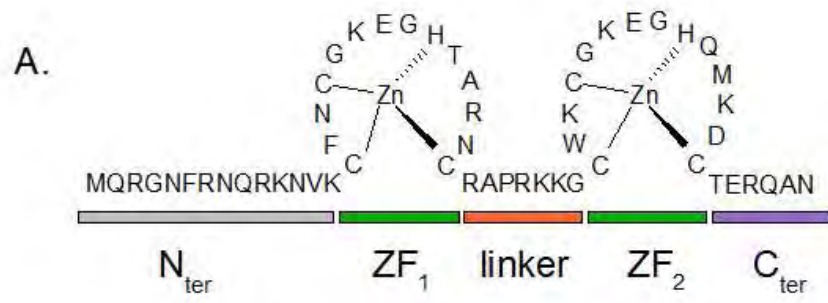


Figure 1

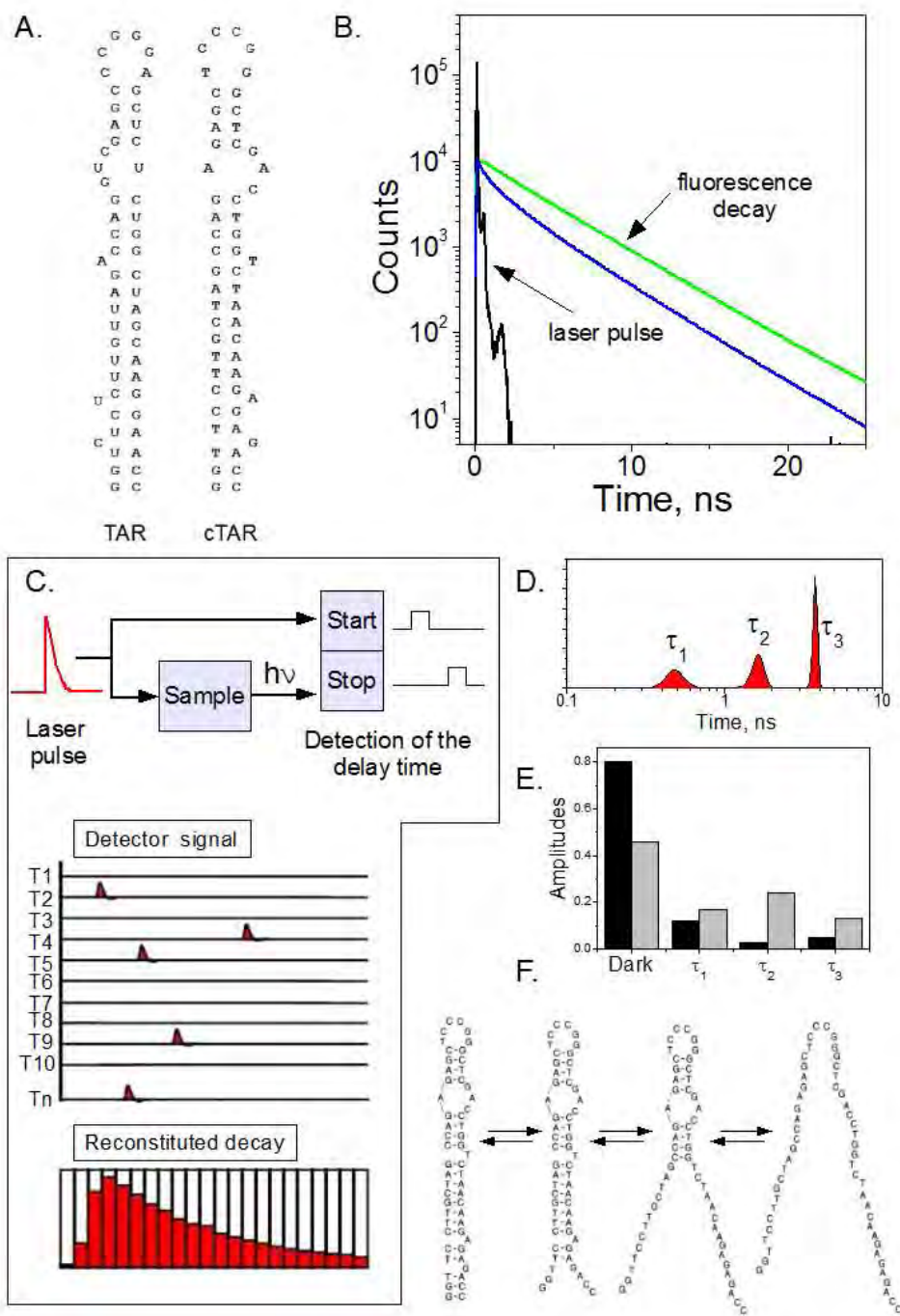


Figure 2

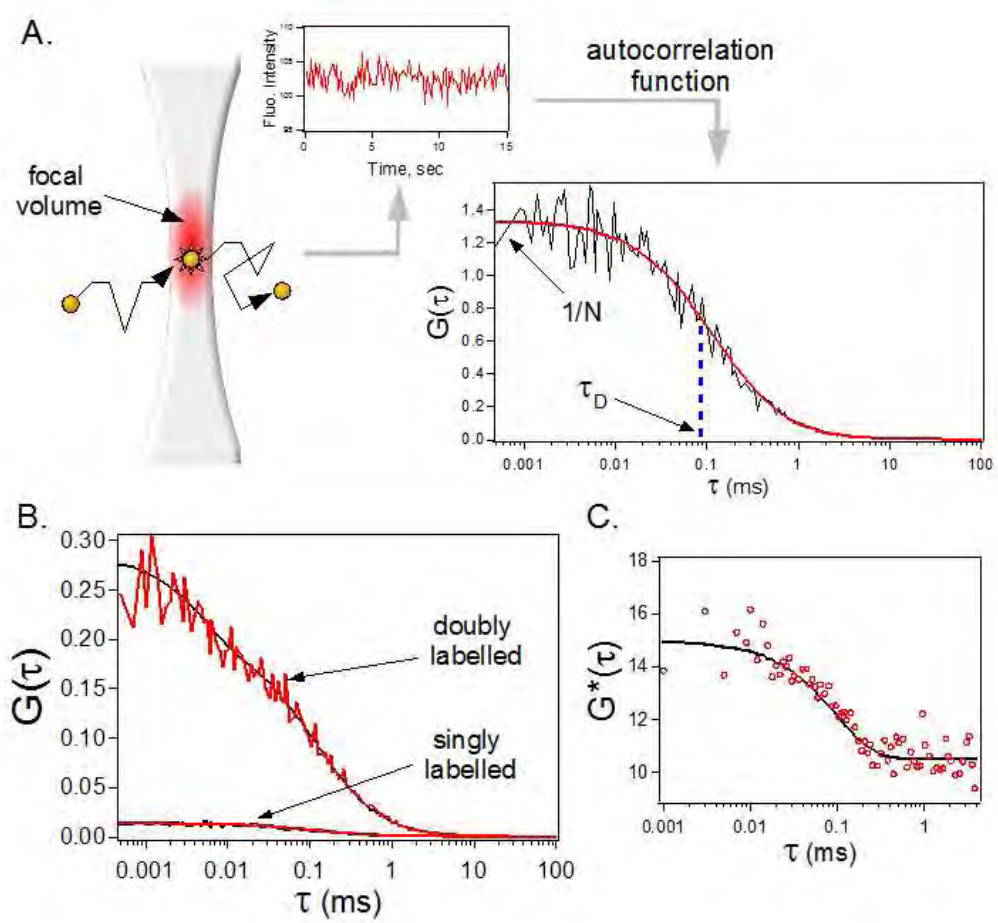


Figure 3

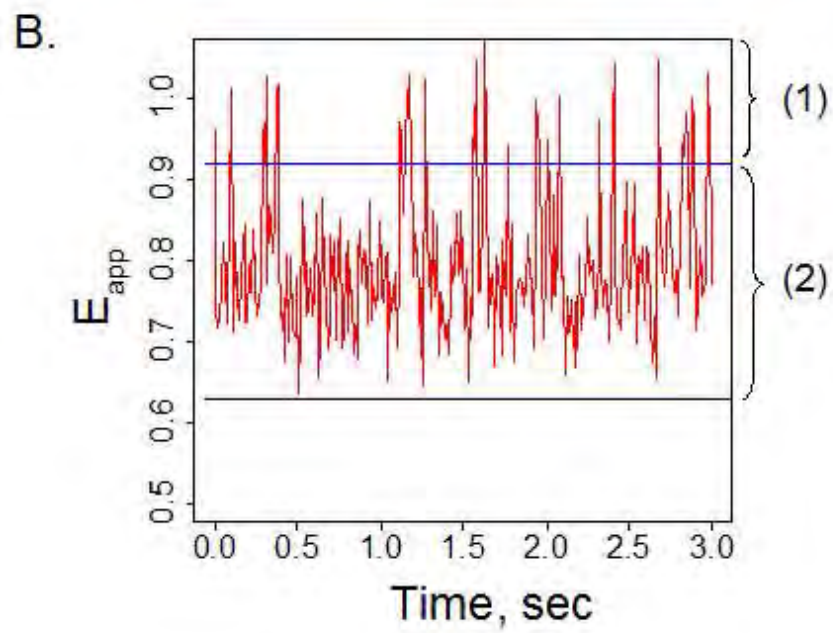
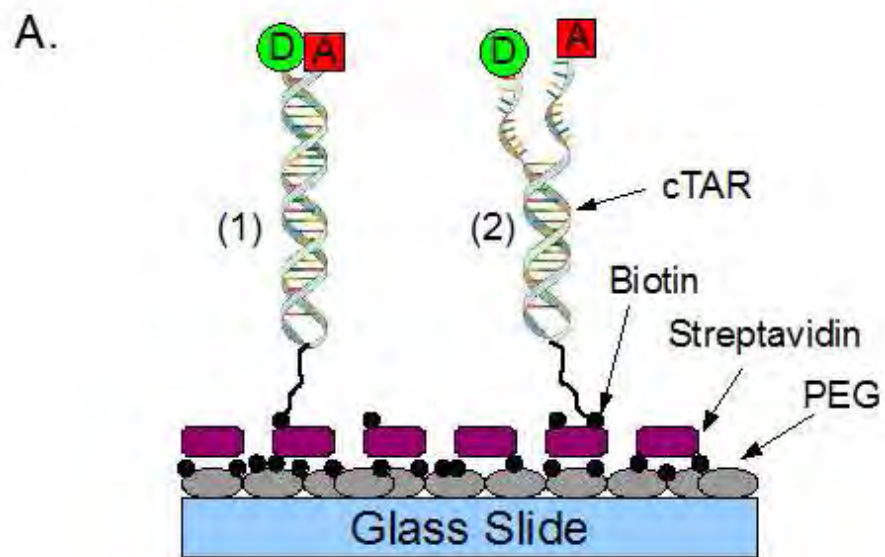


Figure 4

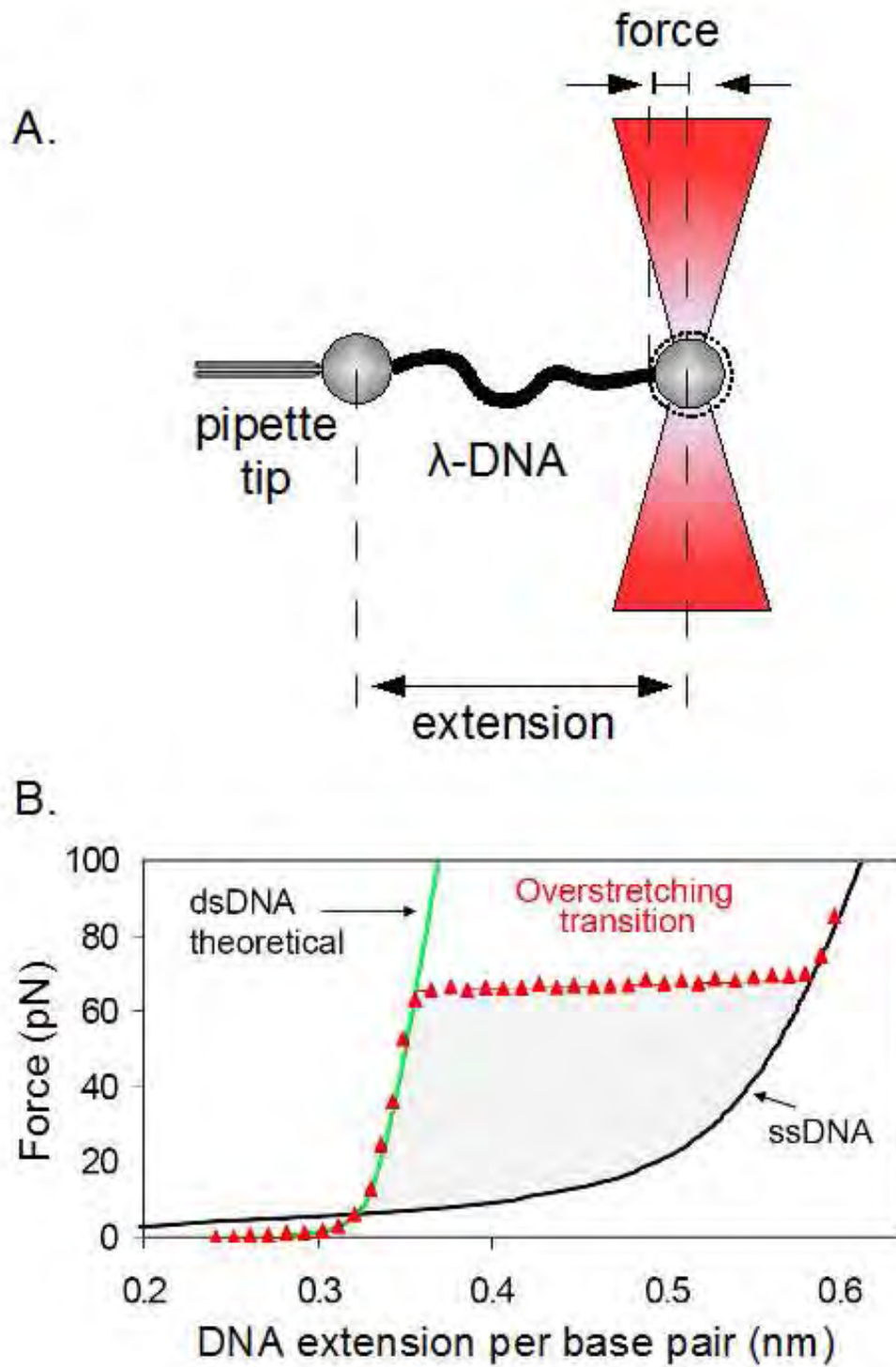


Figure 5

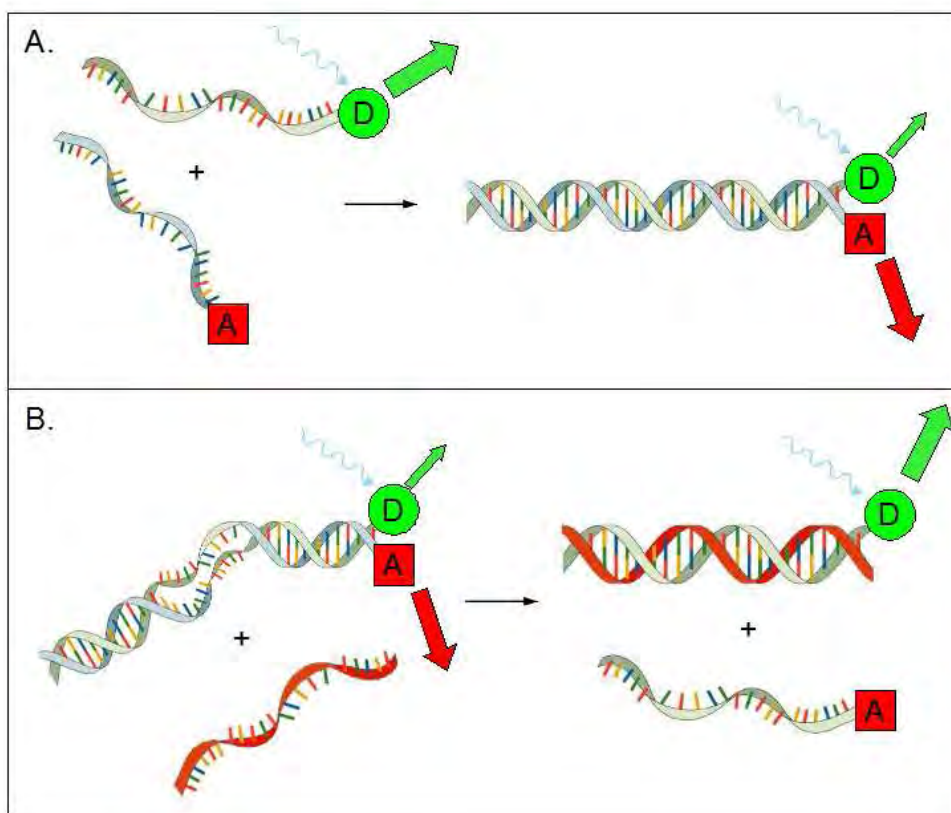


Figure 6

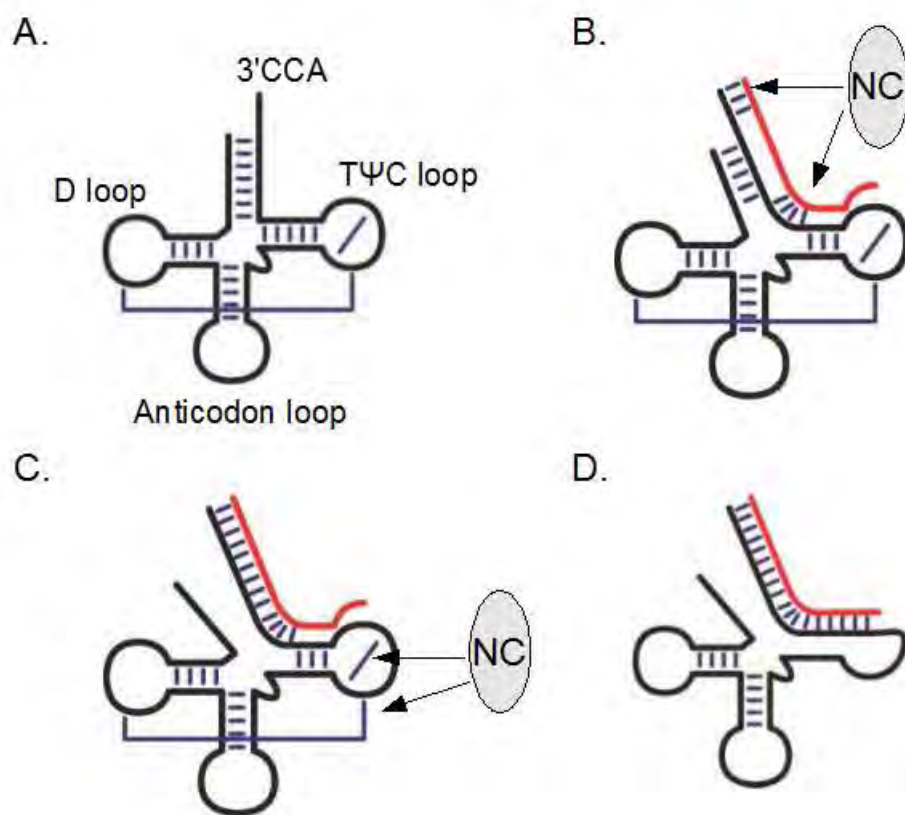


Figure 7

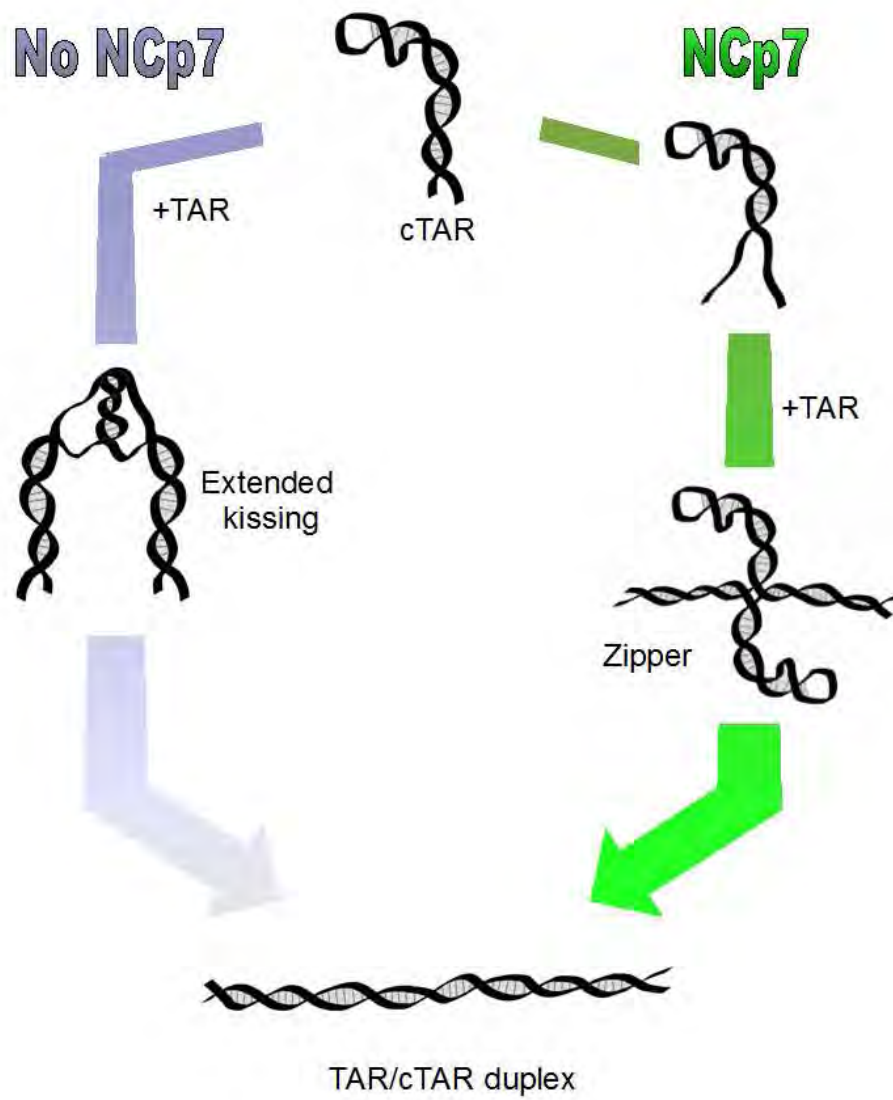


Figure 8

Matériels et méthodes

2.1 Fluorescence résolue en temps

Les principales techniques de mesure de déclin de fluorescence sont (*i*) la technique de phase et de modulation [374, 375] et (*ii*) la technique impulsionnelle selon la méthode du comptage de photon unique [376]. La première technique repose sur une illumination continue de l'échantillon à l'aide d'un faisceau dont l'intensité est modulée sinusoidalement en fonction du temps. La seconde technique repose sur une excitation pulsée et une détection directe du temps de l'état excité. Ces deux méthodes possèdent des résolutions picoseconde-nanoseconde, selon les sources et les détecteurs utilisés. Les mesures que nous avons réalisées dans le cadre de ces travaux ont été faites par la technique du photoélectron unique (PEU).

2.1.1 Technique du photoélectron unique

La technique du PEU consiste à exciter des fluorophores tout en ne détectant que quelques unités de photons de fluorescence (moins de 5 photons pour 100 impulsions d'excitation). L'intervalle de temps entre l'impulsion initiale et la détection du photon, qui correspond au temps passé par une molécule à l'état excité, est mesuré. En répétant un grand nombre de fois cette mesure, il est possible de reconstituer un déclin de fluorescence. La source d'excitation est générée par un laser Titane-Saphir (Tsunami Spectra Physics) fonctionnant en mode *locké*. Ce laser est pompé par un laser continu solide Nd :YO₄ (Millenia X Spectra-Physics) émettant à 532 nm (7-9W en sortie). Le laser Ti :Saphir émet dans un intervalle de longueur d'onde allant de 720 à 1080 nm, avec des impulsions de largeur temporelle de 1 ps. La fréquence de répétition naturelle du laser est de 82 MHz. Cette fréquence, trop élevée pour la plupart des mesures de temps de vie de fluorescence des fluorophores usuels, est ramenée à 4 Mhz à l'aide d'un sélecteur d'impulsions (pulse-picker). À la sortie de ce sélecteur, le faisceau est dirigé sur un système de cristaux non linéaires en accord de phase angulaire (GST). Le faisceau est tout d'abord focalisé sur un cristal LBO (lithium borate) permettant la génération de l'harmonique secondaire ($\lambda/2$). Par la suite, après passage par une ligne de retard optique, le faisceau fondamental et le faisceau doublé sont focalisés sur un cristal BBO (β barium borate) permettant la génération de la troisième harmonique ($\lambda/3$). Le doublage et le triplage de fréquence du faisceau fondamental permet

2.1 Fluorescence résolue en temps

ainsi de pouvoir exciter un grand nombre de fluorophores usuels, respectivement dans le bleu pour le doublé (fluoresceine, rhodamine, ...) et dans l'UV pour le triplé (tryptophane, aminopurine, ...).

Une partie du faisceau doublé est envoyé vers une photodiode rapide (Hamamatsu) connecté à une voie du discriminateur rapide (Tennelec TC 454), générant le signal de départ (*Start*) pour le convertisseur temps-amplitude (TAC) (ORTEC) et le début de la charge du condensateur. Les photons de fluorescence émis par l'échantillon sont détectés selon une direction perpendiculaire à l'excitation par un photomultiplicateur (PMT) à galette de micro-canaux (Hamamatsu R3809U) couplé à un préamplificateur (Philips). Les longueurs d'ondes d'émission sont sélectionnées par un monochromateur (Jobin-Yvon) dont la bande passante est ajustable (4 à 16nm). Un filtre passe-haut est placé devant le PMT afin d'éliminer les photons d'excitation diffusés résiduels. Le signal correspondant au premier photon détecté sert de signal *Stop* pour le TAC, ce qui provoque l'arrêt du chargement du condensateur. La charge accumulée dans le TAC est donc proportionnelle à l'intervalle de temps entre les impulsions *Start* et *Stop*. Cette charge est convertie en une tension comprise entre 0 et 10V, qui est analysée par un analyseur multicanaux (Ortec 921). 2048 canaux sont utilisés pour l'analyse de l'amplitude, ce qui, combiné à une échelle de temps de 50 ns du TAC, donne une linéarité de 25.5 ps/canal. L'histogramme de la probabilité d'émission de fluorescence est reconstitué par la mesure d'un grand nombre de photons, typiquement de 5×10^5 photons.

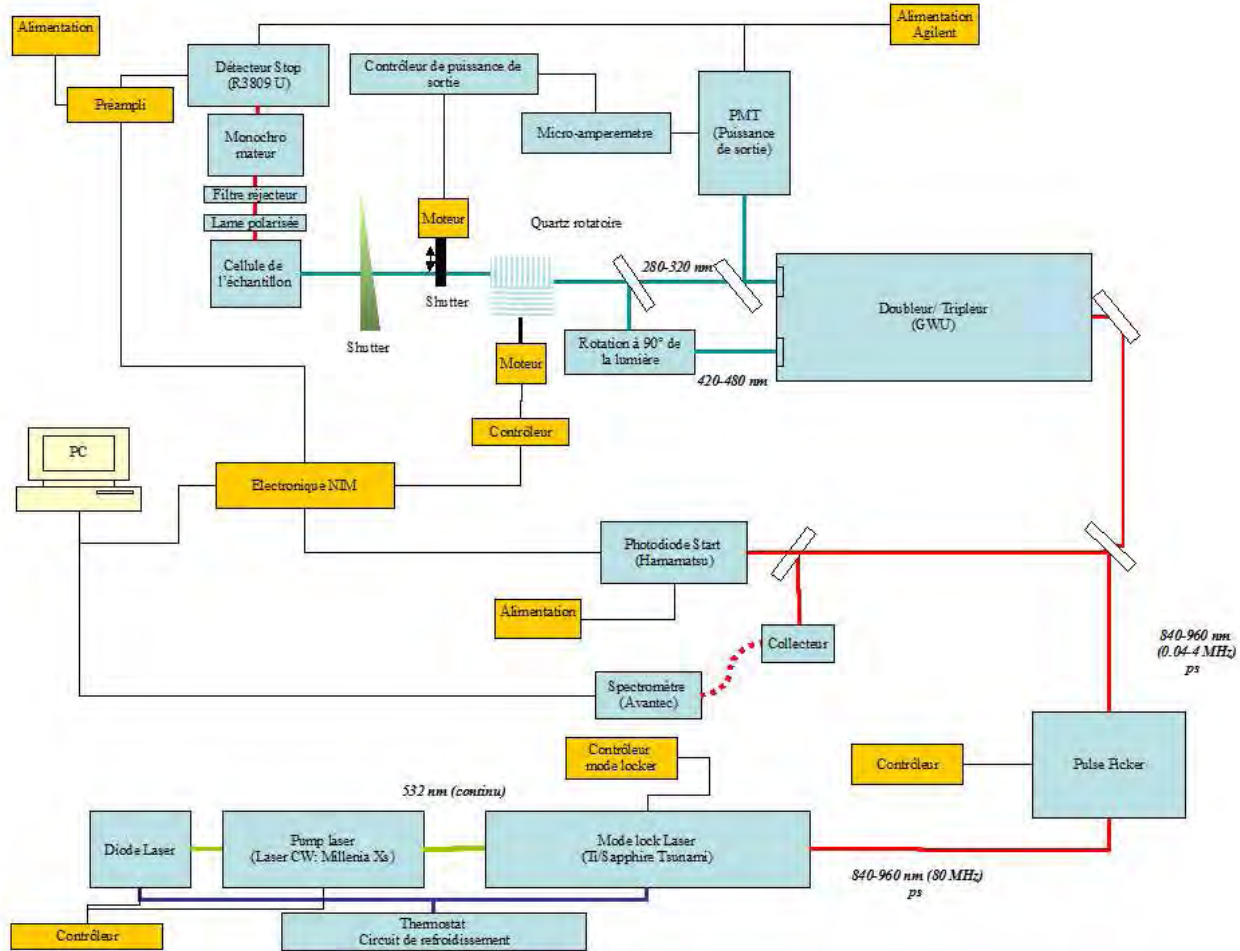


FIGURE 2.1 – Schéma du montage permettant les mesures de fluorescence résolue en temps.

2.1 Fluorescence résolue en temps

2.1.2 Déclins de fluorescence

En excitant l'échantillon avec un flash de lumière polarisée verticalement, le déclin de fluorescence mesuré $I(t)$, qui est le produit de convolution de la loi théorique d'émission de fluorescence de l'échantillon $i(t)$ et de la fonction de réponse instrumentale $E(t)$ se décompose en deux composantes parallèles $I_{vv}(t)$ et perpendiculaires $I_{vh}(t)$, respectivement, selon :

$$I_{vv}(t) = \frac{1}{3}E_{\lambda}(t) \otimes \left\{ \int_0^{\infty} \int_0^{\infty} \int_{-0.2}^{0.4} \gamma(\tau, \theta, r_0) e^{-\frac{t}{\tau}} \left(1 + 2r_0 e^{-\frac{t}{\theta}} \right) d\tau d\theta dr_0 \right\} \quad (2.1)$$

$$I_{vh}(t) = \frac{1}{3}E_{\lambda}(t) \otimes \left\{ \int_0^{\infty} \int_0^{\infty} \int_{-0.2}^{0.4} \gamma(\tau, \theta, r_0) e^{-\frac{t}{\tau}} \left(1 - r_0 e^{-\frac{t}{\theta}} \right) d\tau d\theta dr_0 \right\} \quad (2.2)$$

dans lesquelles $\gamma(\tau, \theta, r_0)$ correspond au nombre de fluorophores caractérisés par un déclin de fluorescence τ , un temps de corrélation de rotation θ et une anisotropie initiale r_0 et où la fonction de réponse instrumentale $IRF_{\lambda}(t)$ est elle-même constituée par le produit de convolution de la forme temporelle de l'impulsion lumineuse incidente $N(t)$ et de la réponse de l'instrument $H(t)$ selon :

$$IRF_{\lambda}(t) = N(t) \otimes H(t) = \int_0^t N(t') H(t-t') dt' \quad (2.3)$$

La fonction de réponse instrumentale est déterminée expérimentalement à l'aide d'un réflecteur d'aluminium poli et est mesurée dans des conditions identiques à celles utilisées pour réaliser les mesures. La largeur à mi-hauteur du maximum (Full Width at Half-Maximum) de la fonction de réponse dans ces conditions est typiquement d'environ 40ps.

La fluorescence totale est obtenue en sommant ces deux composantes selon :

$$I(t) = I_{vv}(t) + 2g I_{vh}(t) = IRF_{\lambda}(t) \otimes i(t) = IRF_{\lambda}(t) \otimes \int_0^{\infty} h(\tau) e^{-\frac{t}{\tau}} d\tau \quad (2.4)$$

où $h(\tau)$ est la distribution des temps de vie de fluorescence, donné par :

$$h(\tau) = \int_0^{\infty} \int_{-0.2}^{0.4} r_0 \gamma(\tau, \theta, r_0) d\theta dr_0$$

2.1 Fluorescence résolue en temps

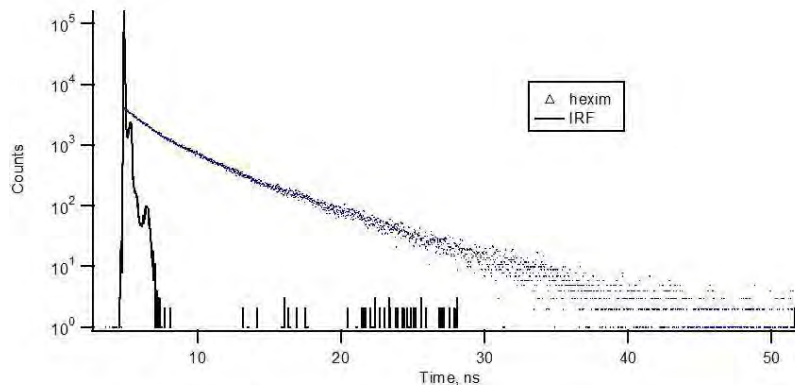


FIGURE 2.2 – Déclin de temps de vie de fluorescence de Hexim, une protéine contenant 4 Trp ($\lambda_{exc} = 295nm$, $\lambda_{em} = 350nm$). La fonction de réponse de l'appareil est également représentée.

et g est un facteur de correction pour la réponse des optiques et de la détection des émissions de fluorescence polarisées verticalement et horizontalement ($1.00 \leq g \leq 1.02$ dans nos mesures).

Si l'on ne s'intéresse qu'au déclin de fluorescence, l'utilisation d'un polariseur réglé à l'angle magique (54.7°) permet de rendre le signal de fluorescence indépendant des mouvements de rotation des molécules excitées. Le déclin mesuré est alors directement :

$$i(t) = IRF_\lambda(t) \otimes \int_0^\infty h(\tau) e^{-\frac{t}{\tau}} d\tau \quad (2.5)$$

2.1.3 Analyse des données

Parmi les méthodes d'analyse des déclins de fluorescence, la méthode de maximisation de l'entropie (MEM) [377] présente l'avantage d'une démarche sans *a priori* (prior non-informative), ce qui la singularise des méthodes utilisant la statistique des moindres carrés, pour lesquelles les résultats peuvent dépendre du choix des conditions initiales. Le maximum d'entropie est en effet un critère optimal pour reconstruire la distribution des temps de vie les plus probables $h(\tau)$ à partir d'un déclin expérimental $I(t)$. La formulation de la fonction d'entropie généralisée de Livesey et Skilling, uniquement dépendante de $h(\tau)$, s'exprime

2.1 Fluorescence résolue en temps

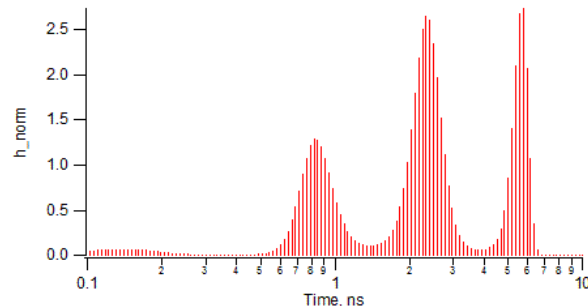


FIGURE 2.3 – Distribution des temps de vie de fluorescence des Trp de la protéine Hexim (obtenus par MEM) correspondant au déclin de la figure 2.2.

comme :

$$S = \int_0^{\infty} h(\tau) - m(\tau) - h(\tau) \log \frac{h(\tau)}{m(\tau)} d\tau \quad (2.6)$$

dans laquelle $m(\tau)$ est une mesure qui quantifie l'importance relative des différents temps de vie τ . S mesure la déviation de la distribution $h(\tau)$ par rapport à $m(\tau)$ (le maximum de S est atteint pour $h(\tau) = m(\tau)$). En pratique, on suppose au départ que la distribution de $h(\tau)$ est constituée par une fonction $m(\tau)$ de 200 valeurs discrètes τ_j régulièrement espacées sur une échelle logarithmique dans un intervalle de temps donné (en général 0.01 – 12ns). Puisque S est maximisée par une *prior* uniforme et non informative en espace logarithmique (tout les temps de vie sont équiprobables), la méthode du maximum d'entropie donne la reconstruction de la distribution des temps de vie la plus probable, et c'est l'information contenue dans les données qui donne la structure à la distribution postérieure. De manière simultanée, un ajustement est réalisé selon le test statistique du χ^2 selon

$$\chi_r^2 = \frac{1}{N} \sum_{n=1}^N \frac{(i(t_n) - I(t_n))^2}{\sigma_n^2} \quad (2.7)$$

où N est le nombre de canaux et σ_n^2 la variance du n^{ieme} canal de la courbe du déclin de fluorescence estimé selon une statistique de Poisson.

2.2 Anisotropie de fluorescence

2.2.1 Principe

Lorsque l'on illumine une population de fluorophores en solution par une lumière polarisée linéairement, ceux dont les moments dipolaires de transition d'absorption $\vec{\mu}_A$ sont orientés parallèlement au vecteur champ électrique du faisceau incident sont préférentiellement excités (photosélection) puisque la probabilité d'absorption est proportionnelle au cosinus carré de l'angle θ entre le vecteur du moment d'absorption $\vec{\mu}_A$ et le vecteur du champ électrique \vec{E} de la lumière d'excitation. La fluorescence est émise selon la direction du moment dipolaire de transition d'émission, $\vec{\mu}_E$. Les moments de transition d'absorption et d'émission ont une orientation fixe au sein d'un fluorophore. L'angle β entre ces deux moments détermine l'anisotropie fondamentale r_0 (c'est à dire l'anisotropie théorique en absence de tout mouvement) selon :

$$r_0 = \frac{2}{5} \left(\frac{3\cos^2\beta - 1}{2} \right) \quad (2.8)$$

Ainsi, r_0 atteint une valeur limite égale à 0.4 lorsque les moments de transitions sont parallèles. La photosélection induit donc une polarisation de la fluorescence selon l'orientation moyenne des molécules excitées au moment de l'émission. Les mouvements Browniens induisent des changements de direction du moment d'émission $\vec{\mu}_E$, d'un angle $\omega(t)$, pendant la durée de vie de l'état excité. Dans ce cas, $r(t)$ est égal à :

$$r(t) = \frac{(3\cos^2(\beta + \omega(t)) - 1)}{5} \quad (2.9)$$

La diffusion rotationnelle induit donc une dépolarisation partielle ou totale de la fluorescence, dépendant de l'amplitude du mouvement Brownien. Avec un temps de rotation de l'ordre de 100 ps par exemple, l'orientation de fluorophores libres dans l'eau devient aléatoire pendant l'état excité et la fluorescence ne présente plus de polarisation (anisotropie proche de 0). Si par contre, le temps de rotation est plus long que le temps de vie de l'état excité (dans la limite d'environ dix fois le temps de vie), comme c'est souvent le cas pour des fluorophores liés à des biomolécules de masse importante, les mesures d'anisotropie fournissent

2.2 Anisotropie de fluorescence

alors des informations sur la mobilité, la taille, la forme et la flexibilité de ces molécules. La mesure d'anisotropie trouve donc une large gamme d'applications dans la caractérisation des interactions entre biomolécules. L'anisotropie de fluorescence, r , peut être calculée en mesurant les intensités de fluorescence résultant d'une excitation polarisée (généralement verticale) avec un polariseur d'émission orienté parallèlement ($I_{//}$) et perpendiculairement (I_{\perp}) au vecteur électrique de la lumière d'excitation. r peut alors être déterminée par :

$$r = \frac{I_{//}/I_{\perp} - 1}{I_{//}/I_{\perp} + 2} \quad (2.10)$$

qui est plus souvent présentée sous la forme :

$$r = \frac{I_{//} - I_{\perp}}{I_{//} + 2 \times I_{\perp}} \quad (2.11)$$

L'anisotropie est sans dimension puisque la différence ($I_{//} - I_{\perp}$) est normalisée par l'intensité totale de fluorescence ($I_{//} + 2 \cdot I_{\perp}$). Il convient ici également de rappeler que la dépolarisation peut aussi être induite par du transfert d'énergie d'une molécule "donneur" à une molécule "accepteur" orientée différemment.

La loi d'additivité de l'anisotropie impose en outre que l'anisotropie d'un mélange corresponde à la somme des anisotropies individuelles des composés du mélange, pondérées par leur fluorescence relative f tel que :

$$r = \sum_i r_i f_i$$

2.2.2 Mesure d'anisotropie à l'état stationnaire

Pour les mesures d'anisotropie de fluorescence à l'état stationnaire, nous avons utilisé un spectrofluorimètre Fluorolog (Jobin Yvon Horiba) possédant un seul canal d'émission (format dit en L), ainsi qu'un spectromètre SLM 8000 (format en T).

2.2 Anisotropie de fluorescence

Montage en L

Dans le cas du montage en L, l'émission de fluorescence mesurée passe par un monochromateur possédant une efficacité de transmission différente pour les lumières polarisées verticalement ou horizontalement. En conséquence, les intensités des émissions parallèles et perpendiculaires à la polarisation du faisceau d'excitation, mesurées grâce à la rotation du polariseur d'émission, sont biaisées. En effet, les intensités mesurées I_{VV} (polarisation d'excitation et d'émission verticales) et I_{VH} (polarisation d'excitation verticale et d'émission horizontale) peuvent s'exprimer selon :

$$I_{VV} = k \cdot S_V \cdot I_{//}$$

$$I_{VH} = k \cdot S_H \cdot I_{\perp}$$

où $S_{V,H}$ traduit la sensibilité de la voie de détection pour les composantes polarisées verticalement ou horizontalement, respectivement, k est un facteur de proportionnalité qui tient compte du rendement quantique des fluorophores et autres facteurs instrumentaux et $I_{//}$, I_{\perp} sont les intensités non biaisées des composantes parallèles et perpendiculaires.

Donc

$$\frac{I_{VV}}{I_{VH}} = G \cdot \frac{I_{//}}{I_{\perp}} \quad \text{avec } G = \frac{S_V}{S_H}$$

La mesure de l'anisotropie ne peut donc être faite qu'après la détermination de ce facteur G . Il faut noter que la valeur de G dépendant de la longueur d'onde d'émission, et d'une moindre manière, de la bande passante du monochromateur. En utilisant une polarisation horizontale pour la lumière d'excitation, les intensités mesurées s'expriment cette fois-ci selon :

$$I_{HV} = k \cdot S_V \cdot I_{\perp}$$

$$I_{HH} = k \cdot S_H \cdot I_{\perp}$$

2.2 Anisotropie de fluorescence

dont le ratio permet d'expliciter directement G .

Ce qui donne, en reprenant les expressions précédentes :

$$\frac{I_{//}}{I_{\perp}} = \frac{I_{VV}}{I_{VH}} \times \frac{1}{G}$$

Comme l'anisotropie de fluorescence, r , est déterminée par l'équation 2.10 ou 2.11, on peut l'exprimer en fonction des intensités mesurées selon :

$$r = \frac{I_{//} - G \times I_{\perp}}{I_{//} + 2G \times I_{\perp}} \quad (2.12)$$

Montage en T

Dans un montage en T (Figure 2.4), les intensités des composantes parallèles et perpendiculaires sont mesurées simultanément en utilisant deux voies séparées de détection. Puisque les positions des polariseurs d'émission sont fixes, la mesure de la sensibilité aux polarisations n'est plus nécessaire. Seule la sensibilité de chacune des deux voies de détection doit être connue, ce qui est réalisé en changeant la polarisation de l'excitation. Le polariseur d'excitation est tout d'abord placé verticalement, et le ratio des signaux parallèle et perpendiculaire est donné par

$$R_V = \frac{G_{//} \cdot I_{//}}{G_{\perp} \cdot I_{\perp}}$$

$G_{//}$ et G_{\perp} sont les gains des canaux parallèle et perpendiculaire respectivement.

Le polariseur d'excitation est ensuite placé horizontalement et le ratio est cette fois-ci exprimé comme

$$R_H = \frac{G_{//} \cdot I_{\perp}}{G_{\perp} \cdot I_{//}}$$

ou directement

$$R_H = \frac{G_{//}}{G_{\perp}}$$

Le ratio de ces dernières égalités donne directement

$$\frac{R_V}{R_H} = \frac{I_{//}}{I_{\perp}}$$

2.2 Anisotropie de fluorescence

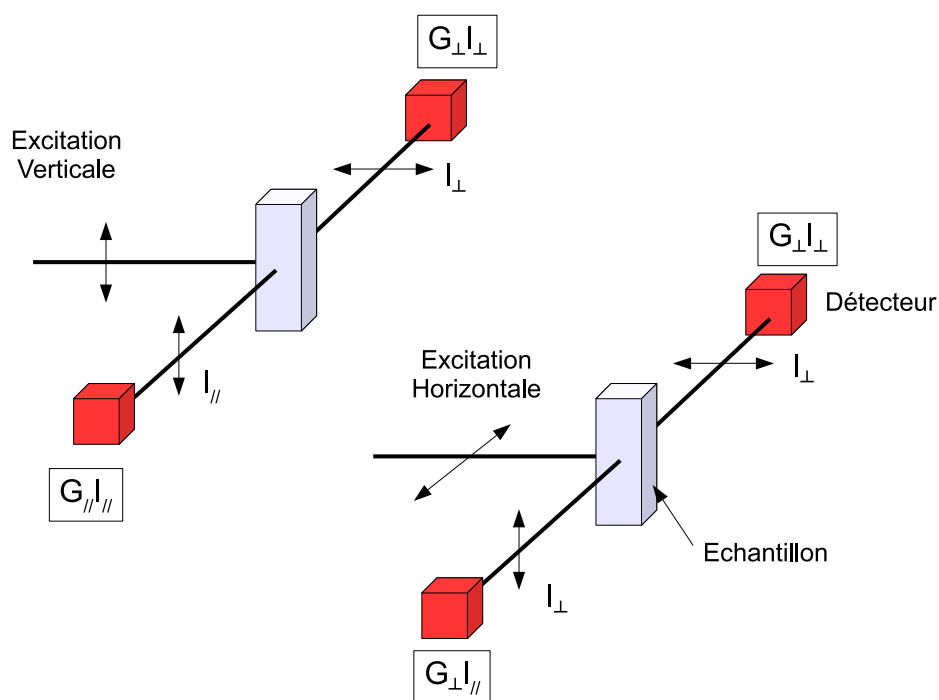


FIGURE 2.4 – Schéma d'un format en T pour des mesures d'anisotropie de fluorescence.

ce qui permet de calculer l'anisotropie selon l'équation 2.10.

2.2.3 Mesure d'anisotropie résolue en temps

Si l'on substitue l'excitation continue par une source pulsée polarisée, alors il est possible de reconstituer un déclin d'anisotropie résolu en temps (voir équation 2.1 et 2.2). L'information obtenue est plus riche que la valeur d'anisotropie moyenne obtenue à l'état stationnaire. En effet, la forme que prend le déclin d'anisotropie résolue en temps dépend de la taille, de la forme et de la flexibilité des molécules marquées, ce qui se traduit généralement par un déclin multi-exponentiel. Pour une molécule dont la forme dévie fortement d'une sphère par exemple, les temps de corrélations observés sont déterminés par les vitesses de rotation de la molécule selon ses différents axes, ce qui se traduit par un déclin multi-exponentiel. De même, les mouvements locaux (comme le mouvement de la sonde vis à vis de la molécule

2.2 Anisotropie de fluorescence

à laquelle elle est liée) ou segmentaux, induisent des composantes exponentielles supplémentaires. Enfin, il existe des cas de figure où la rotation des molécules est dite gênée (la molécule ne diffuse pas librement). Ce dernier cas de figure induit des déclin d'anisotropie présentant une anisotropie résiduelle non-nulle.// Les déclin d'anisotropie résolue en temps sont analysés selon les équations :

$$I_{//}(t) = I(t) \frac{[1 + 2r(t)]}{3} \quad (2.13)$$

$$I_{\perp}(t) = \frac{I(t) [1 - r(t)]}{3} \quad (2.14)$$

$$r(t) = \frac{I_{//}(t) - G \times I_{\perp}(t)}{I_{//}(t) + 2G \times I_{\perp}(t)} = r_0 \sum_i \beta_i e^{\frac{-t}{\tau_i}} \quad (2.15)$$

2.3 Transfert d'énergie non radiatif de fluorescence

Le transfert d'énergie est un phénomène observé lorsqu'une molécule luminescente (donneur) cède, à l'état excité, son énergie à une seconde molécule (accepteur). Le transfert énergétique peut être radiatif ou non. Le FRET (*Fluorescence Resonance Energy Transfer*) correspond à un transfert non radiatif, c'est à dire sans émission de photons, par couplage dipôle-dipôle. Une des raisons pour lesquelles cette méthode est si répandue dans les études biophysiques est que le FRET sonde efficacement des distances comprises entre 10 et 100 Å environ (soit de 1 à 10 nm), ce qui correspond à l'ordre de grandeur de la taille d'une protéine ou de l'épaisseur d'une membrane cellulaire (Figure 2.5). De plus, et puisqu'il s'agit d'une interaction dipôle-dipôle à distance, le transfert d'énergie est très peu affecté par les biomolécules de l'échantillon et dépend donc presque uniquement des propriétés spectrales des fluorophores utilisés, ce qui le rend très prédictible.

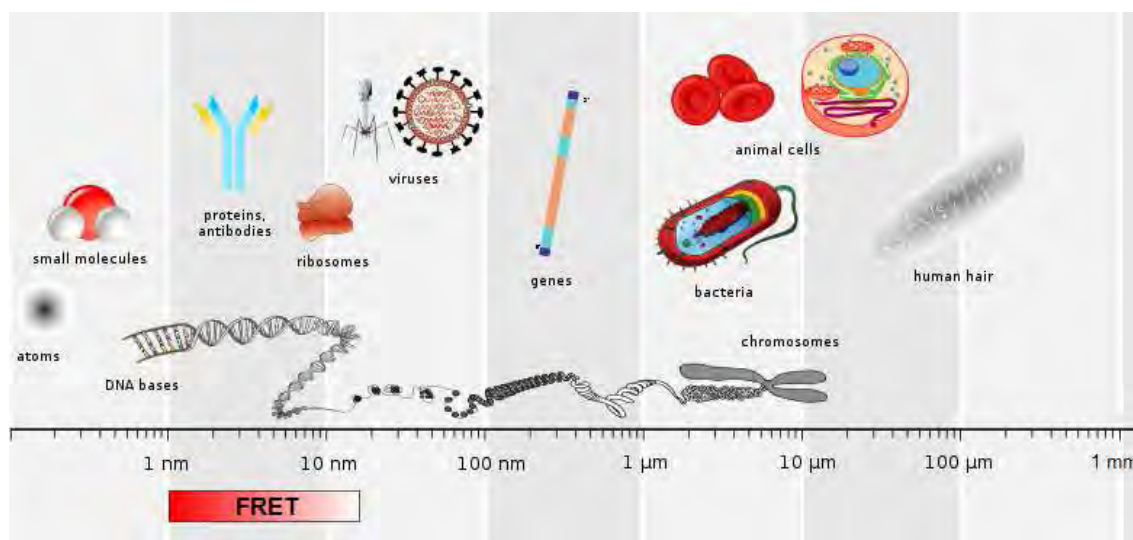


FIGURE 2.5 – Représentation de la gamme de distances sondées par la technique de FRET au sein d'une échelle biologique.

Théorie du transfert d'énergie pour une paire donneur-accepteur

Les équations de transfert d'énergie dérivent à la fois de la mécanique classique et de la mécanique quantique. La formulation par Förster [378] de la vitesse du transfert d'énergie

2.3 Tranfert d'énergie non radiatif de fluorescence

s'exprime selon :

$$k_T(r) = \frac{QY_D \kappa^2}{\tau_D r^6} \left(\frac{9000(\ln 10)}{128 \Pi^5 N n^4} \right) J_{AD} \quad (2.16)$$

QY_D et τ_D représentent respectivement le rendement quantique et la durée de vie de fluorescence du donneur en absence de l'accepteur

n est l'indice de réfraction du milieu (usuellement fixé à entre 1.33 et 1.4 pour les biomolécules en solution)

r la distance entre les deux sondes

J_{AD} traduit le recouvrement spectral entre les spectres d'émission du donneur et d'absorption de l'accepteur. C'est un paramètre qui est calculé selon :

$$J_{AD} = \int_0^\infty F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda \quad (2.17)$$

où $F_D(\lambda)$ est l'intensité de fluorescence corrigée et normalisée à l'unité du donneur dans la gamme de longueur d'onde de λ à $\lambda + \Delta\lambda$ et où $\epsilon_A(\lambda)$ est le coefficient d'absorption de l'accepteur à la longueur d'onde λ .

La distance de Förster est la distance particulière pour laquelle la moitié des molécules se désexcitent par transfert. En exprimant λ en cm et J_{AD} en $M^{-1}cm^3$, R_0 vaut

$$R_0 = 9.78 \times 10^3 (\kappa^2 n^{-4} QY_D J_{AD})^{1/6} \text{ (en } \text{Å)} \quad (2.18)$$

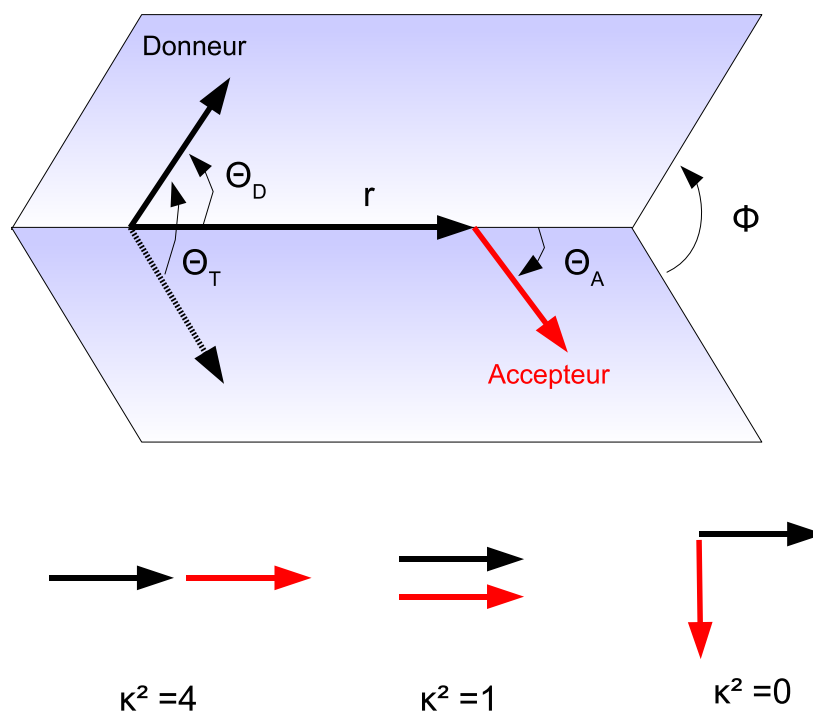
dans laquelle κ^2 est un facteur qui tient compte de l'orientation relative des deux sondes.

$$\kappa^2 = (\cos(\theta_T) - 3 \cos(\theta_D) \cos(\theta_A))^2 \quad (2.19)$$

θ_T est l'angle entre la transition d'émission du dipôle du donneur et la transition d'absorption du dipôle de l'accepteur, θ_A et θ_D sont les angles entre les dipôles et le vecteur joignant le donneur à l'accepteur (Figure 2.6).

Selon cette définition, κ^2 peut varier de 0 à 4. La valeur de $\kappa^2 = 2/3$ est habituellement utilisée, ce qui correspond à une valeur moyenne obtenue si les sondes sont mobiles. Des mesures d'anisotropie de fluorescence résolue en temps du donneur (et de l'accepteur) permettent de s'assurer de la validité de cette approximation.

2.3 Transfert d'énergie non radiatif de fluorescence



Orientations particulières des dipôles

FIGURE 2.6 – Représentation des orientations des dipôles du donneur et de l'accepteur, ainsi que les valeurs de κ^2 pour trois orientations particulières

Ainsi, si l'on reprend l'équation 2.16 et 2.18, en fixant $\kappa^2 = 2/3$, $k_T(r)$ se simplifie en une fonction de r et de τ_D^{-1} selon

$$k_T(r) = \frac{1}{\tau_D} \left(\frac{R_0}{r} \right)^6 \quad (2.20)$$

L'efficacité de transfert (E) est la fraction de photons absorbés par le donneur qui sont transférés à l'accepteur. Le transfert se traduit par une diminution de l'intensité de fluorescence et de la durée de vie de l'état excité du donneur. En quantifiant ces variations, il est possible de déterminer l'efficacité de transfert E .

$$E = \frac{k_T(r)}{\tau_D^{-1} + k_T(r)} \quad (2.21)$$

L'équation 2.21 met en évidence que si $k_T(r) \gg \tau_D^{-1}$, alors E est proche de 1.

Si l'on rappelle que $k_T(r) = \tau_D^{-1} (R_0/r)^6$, il vient que

2.3 Transfert d'énergie non radiatif de fluorescence

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (2.22)$$

L'équation 2.22 montre que l'efficacité de transfert dépend fortement de la distance inter-sondes (selon une puissance 6).

Ceci amène 2 remarques : la première est que pour des distances proches de R_0 , une faible variation de distance entre les sondes induit une forte variation de E , et deuxièmement, en-deça de $R_0/2$ ou au-delà de $2 \times R_0$, les variations de E deviennent négligeables, même pour des variations de distances significatives à l'échelle de biomolécules (Figure 2.7). Les mesures de FRET se montrent donc particulièrement appropriées pour mettre en évidence des fluctuations de distances inter-sondes au sein de biomolécules.

Du point de vue expérimental, E peut être calculé par la fluorescence relative du donneur en présence de l'accepteur selon :

$$E = 1 - \frac{I_{DA}}{I_D} \quad (2.23)$$

ou encore à l'aide des temps de vie de fluorescence du donneur

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (2.24)$$

Les indices D et DA dans les deux équations précédentes précisent que le paramètre indicé correspond à celui du donneur de fluorescence en absence et en présence de l'accepteur, respectivement. Pour le calcul des distances inter-sondes r , la distance de Förster est estimée à partir des spectres d'émission du donneur et d'absorption de l'accepteur selon le code Matlab de la section annexe chapitre 8.

Les équations 2.23 et 2.24 ne sont valides que pour un modèle dans lequel il existe une distance fixe entre le donneur et l'accepteur. Si plusieurs distances coexistent, alors le FRET observé correspondra à la valeur moyenne des efficacités de transfert pondérées par leurs populations relatives.

Ainsi, le FRET est un outil intéressant pour sonder des distances au sein de biomolécules

2.3 Tranfert d'énergie non radiatif de fluorescence

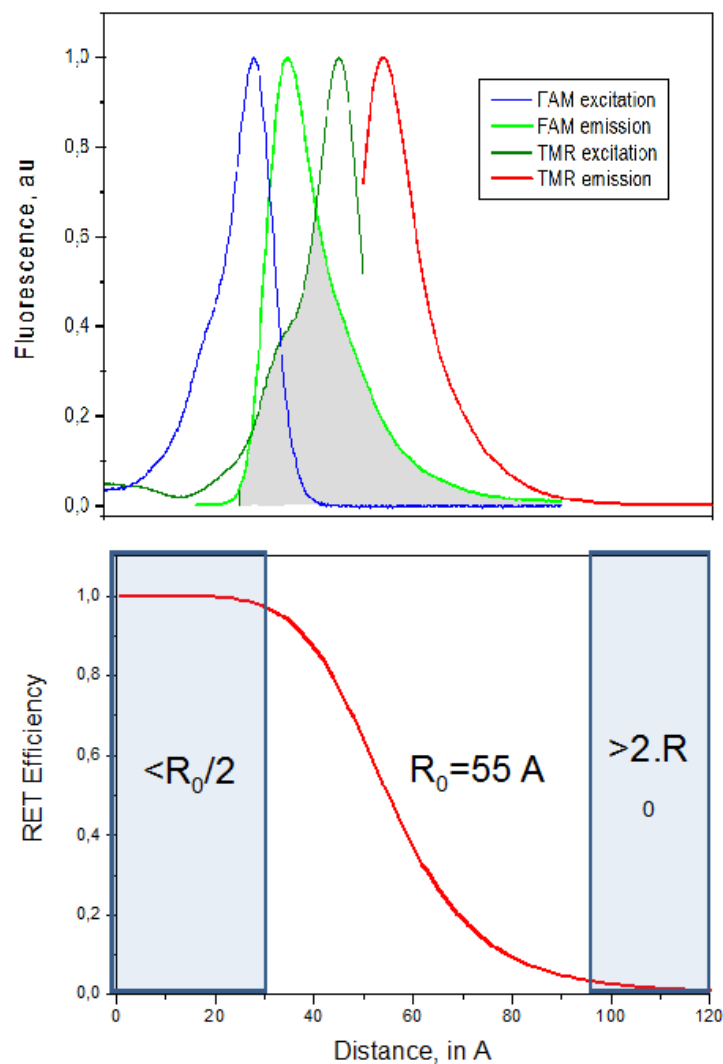


FIGURE 2.7 – Haut : Spectre d'excitation et d'émission de fluorescence de la fluorescéine et de la tétraméthylrhodamine (TMR), un couple de fluorophores largement utilisés pour les expériences de FRET.

L'overlap J_{AD} entre les spectres d'émission du donneur et d'absorption de l'accepteur est la zone grisée.

Bas : Variation de l'efficacité de transfert en fonction de la distance inter-fluorophores. La gamme des distances mesurables est comprise entre $0.5 \times R_0$ et $2 \times R_0$

2.3 Transfert d'énergie non radiatif de fluorescence

ou de complexes biomoléculaires doublement marqués. Dans le cadre des travaux effectués dans cette thèse, nous avons utilisé des mesures de FRET afin de déterminer des distances au sein d'un complexe trimoléculaire ADN-protéine-peptide. En multipliant les positions de marquages, et donc la connaissance de différentes distances au sein de ces complexes, il nous a été possible de confirmer la validité d'un modèle structural en solution obtenu par SAXS pour les récepteurs RAR-RXR et VDR-RXR liés à un ADN double brin (Publication IX : Rochel & al., chapitre 9). Un des intérêts du FRET est de permettre un suivi cinétique des variations de distances, et ce sur une large gamme d'échelle de temps (allant de la μsec à plusieurs heures), et de concentrations (de la molécule unique à quelques μM). Lorsqu'il est possible de synchroniser un processus ou une réaction, des mesures d'ensemble de FRET (FRET en solution) peuvent être exploitées pour suivre la cinétique de ces phénomènes (Publication III : Godet & al. 2006 [280]). Dans le cas contraire, des mesures de FRET à l'échelle de la molécule unique permettent d'accéder à l'ensemble des caractéristiques des populations qui composent le mélange.

Spectroscopie à corrélation de fluorescence

Spectroscopie à corrélation de fluorescence (FCS)

La spectroscopie à corrélation de fluorescence (FCS) est une méthode permettant l'analyse temporelle et spatiale de biomolécules faiblement concentrées. Alors que la plupart des techniques de fluorescence s'intéressent à l'intensité du signal de fluorescence, la FCS est basée sur l'analyse des fluctuations temporelles de l'intensité de fluorescence mesurées dans un volume fini (volume focal). La FCS repose sur l'idée que les variations minimales d'un système puissent être reflétées par les fluctuations du signal de fluorescence. Là même où généralement on considère ces fluctuations comme du bruit de mesure, la FCS cherche à exploiter ces fluctuations pour en tirer de l'information. Le traitement des fluctuations par une fonction d'autocorrélation (donnant son nom à cette méthode) permet une mesure de l'auto-similarité de mesures répétées temporellement. Par voie de conséquence, elle décrit les informations persistantes contenues dans ces fluctuations. Parmi ces informations, on retrouve en particulier la concentration locale en biomolécules fluorescentes ou la constante de diffusion de ces molécules. L'ensemble des phénomènes capables d'engendrer des fluctuations du signal de fluorescence peuvent en théorie être appréhendés par FCS.

3.1 Autocorrélation : principe et modélisation des données

L'origine principale des fluctuations est le mouvement Brownien des particules entrant et sortant du volume d'excitation. Le nombre de particules présentes dans le volume focal au moment t suit une loi de Poisson. La racine de la variation du nombre moyen de molécules au sein du volume focal peut être donné par :

$$\frac{\sqrt{\langle(\delta N)^2\rangle}}{\langle N\rangle} = \frac{1}{\sqrt{\langle N\rangle}} \quad (3.1)$$

ce qui implique que les fluctuations relatives deviennent plus importantes lorsque la concentration en espèces marquées diminue (limité néanmoins par le fait que le signal de fluores-

3.1 Principe

cence doit être plus grand que le bruit résiduel). En pratique, le nombre de molécules au sein du volume focal peut aller de 0.1 à ~ 1000 , ce qui, pour un volume focal de 0.3 fL, donne des concentrations allant du sub-nanomolaire au micromolaire. Ces gammes de concentrations sont largement compatibles avec les concentrations "physiologiques" que l'on peut observer au sein de cellules.

Les fluctuations de fluorescence sont définies comme les déviations par rapport au signal moyen :

$$\delta F(t) = F(t) - \langle F(t) \rangle \quad (3.2)$$

$$\langle F(t) \rangle = \frac{1}{T} \int_0^T F(t) dt \quad (3.3)$$

Si on suppose que les fluctuations de fluorescence ne résultent que de variations de la concentration locale δC au sein du volume focal V_{eff} , on peut écrire :

$$\delta F(t) = \kappa \int_V I_{ex}(r) \cdot S(r) \cdot \delta(\sigma \cdot QY \cdot C(r, t)) dV \quad (3.4)$$

ce qui décrit la probabilité d'exciter un fluorophore au sein du volume focal et de détecter les photons résultants. En effet, κ caractérise l'efficacité de détection, $I_{ex}(r)$ traduit la distribution de la lumière d'excitation, $S(r)$ détermine l'efficacité de collection des photons du montage et $\delta(\sigma qC(r, t))$ traduit la dynamique du fluorophore au niveau moléculaire ($\delta\sigma$ pour les fluctuations de la cross-section d'absorption, δQY pour les variations du rendement quantique de fluorescence et $\delta C(r, t)$ pour les fluctuations de la concentration locale en fonction du temps).

Les paramètres $I_{ex}(r)$ et $S(r)$ sont généralement combinés en une fonction $W(r)$ qui décrit la distribution spatiale des photons émis. En général $W(r)$ est approximé par une Gaussienne 3D. De même, les paramètres κ , σ et QY sont combinés pour donner un paramètre Cr qui détermine le flux de photons par molécule et par seconde (*count rate*). Ceci permet de transformer l'équation 3.4 en :

$$\delta F(t) = \int_V W(r) \cdot \delta(Cr \cdot C(r, t)) dV \quad (3.5)$$

3.1 Principe

Puisque la fonction d'autocorrelation normalisée est définie par :

$$G(\tau) = \frac{\langle \delta F(t) \cdot \delta F(t + \tau) \rangle}{\langle F(t) \rangle^2} \quad (3.6)$$

En substituant l'équation 3.5 dans l'équation 3.6, on obtient

$$G(\tau) = \frac{\iint W(r) \cdot W(r') \langle \delta(Cr.C(r, t)) \delta(Cr.C(r', t + \tau)) \rangle dV dV'}{(\int W(r) \langle \delta(Cr.C(r, t)) \rangle dV)^2} \quad (3.7)$$

Si l'on considère que δCr est constant pour un système donné, $\delta(Cr.C(r, t)) = \delta Cr.C(r, t) + Cr.\delta C(r, t)$ se simplifie.

De même, si l'on considère uniquement les particules qui diffusent librement dans les 3 dimensions de l'espace avec un coefficient D,

$$\langle \delta(Cr.C(r, 0)) \cdot \delta(Cr.C(r', \tau)) \rangle = \langle C \rangle \frac{1}{(4\pi D\tau)^{-3/2}} \cdot \exp\left(-\frac{(r-r')^2}{4D\tau}\right) \quad (3.8)$$

L'équation 3.7 devient :

$$G(\tau) = \frac{1}{(4\pi D\tau)^{-3/2}} \frac{\iint W(r)W(r') \cdot \exp\left(-\frac{(r-r')^2}{4D\tau}\right) dV dV'}{(\int W(r)dV)^2} \quad (3.9)$$

Enfin, si l'on insère d'une part la relation entre le coefficient de diffusion D et le temps de diffusion latérale τ_D

$$\tau_D = \frac{r_0^2}{4D}$$

et d'autre part, la définition du volume focal effectif

$$V_{eff} = \frac{(\int W(r)dV)^2}{\int W^2(r)dV} = \pi^{3/2} \cdot r_0^2 \cdot z_0 \quad (3.10)$$

on obtient

$$G(\tau) = \frac{1}{V_{eff}\langle C \rangle} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{\left(\frac{r_0}{z_0}\right)^2 \tau_d} \tau\right)^{-0.5} \quad (3.11)$$

3.2 Instrumentation

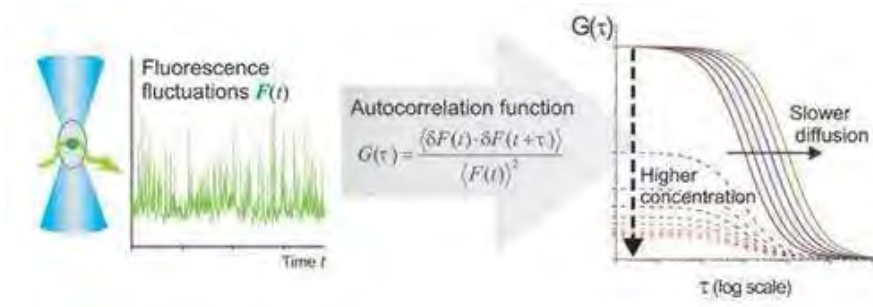


FIGURE 3.1 – Schéma illustrant le principe de la FCS

que l'on formule en général selon

$$G(\tau) = \frac{1}{N} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{S^2} \frac{\tau}{\tau_d}\right)^{-0.5} \quad (3.12)$$

3.2 Instrumentation

La génération du volume de mesure est obtenue à l'aide d'une excitation biphotonique. Le faisceau d'excitation est un faisceau impulsionnel femtoseconde produit par laser pulsé Ti-Sapphire (Tsunami, Spectra Physics) fonctionnant en mode locké, pompé par un laser continu solide Nd :YO₄ (Millenia X Spectra-Physics). Le volume focal résulte de la focalisation de ce faisceau à l'aide d'un objectif 60x (NA 1.2) monté sur un microscope inversé (Olympus IX70). La fluorescence est collectée par ce même objectif puis dirigée à l'aide d'un miroir dichroïque (COWL 750 nm ; Coherent) vers la voie de détection. Le faisceau de fluorescence est filtré (E700SP, Chroma) et focalisé (lentille achromatique 200mm) sur une fibre optique multimode couplée à une photodiode à avalanche (APD) (SPCM-AQR-14-FC, EG&G), elle même couplée à un corrélateur ALV-5000E (ALV GmbH, Allemagne) permettant le calcul en temps réel des courbes de corrélation. Dans le cadre des mesures de corrélation croisée à deux couleurs, les signaux de fluorescence sont séparés à l'aide d'un miroir dichroïque puis filtrés avant d'être détectés chacun par une APD différente. Pour les mesures de FCS à deux *foci* (2fFCS), un interféromètre de Michelson est placé sur la voie d'excitation. En désalignant le beamsplitter, il est possible de générer deux faisceaux non-

3.3 Applications

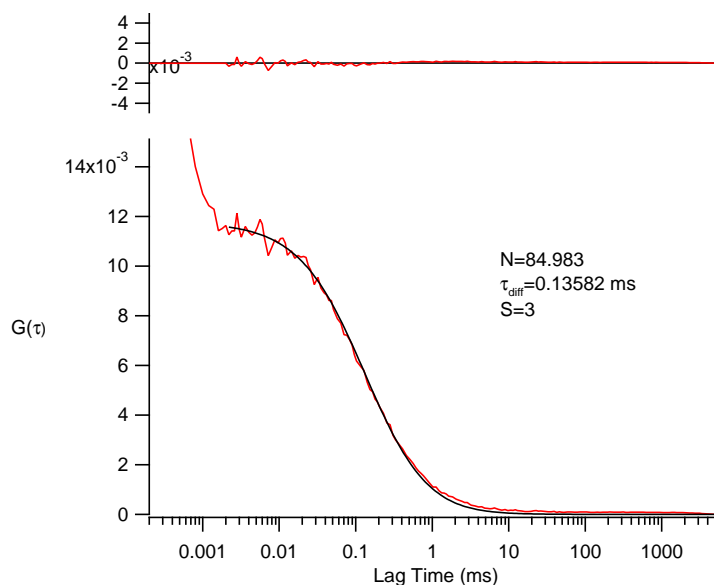


FIGURE 3.2 – Exemple de courbe d'autocorrélation correspondant à une solution de NCp7 marquée par une tétraméthylrhodamine (TMR)

collinéaires décalés latéralement dans le plan focal du microscope. Les signaux provenant de chacun de ces volumes focaux sont détectés et corrélés de manière croisée.

3.3 Application à l'étude des interactions de la NCp7 avec des oligonucléotides

Selon l'équation d'autocorrélation 3.12, il vient immédiatement que

$$G(0) = \frac{1}{N}$$

La courbe d'autocorrélation permet donc de calculer des concentrations locales au niveau du volume focal. En suivant précisément les variations du nombre apparent de molécules dans le volume focal, il devrait être possible de mettre en évidence la formation de dimères, ou de complexes d'ordre supérieur par une simple diminution du nombre de molécules diffusant dans le milieu de mesure.

Suite aux travaux de Andersen & al. [296] montrant que TAR pouvait jouer un rôle dans

3.3 Applications

la dimérisation de l'ARN génomique de VIH-1, nous avons cherché à caractériser la promotion par la NCp7 de la formation d'homodimères de TAR, cTAR (l'ADN complémentaire de TAR), ainsi que de différents mutants de ces séquences. Un mutant de NCp7, la (11-55)NCp7, a également été utilisé pour évaluer l'importance, dans ce processus, de la composante d'agrégation des acides nucléiques de NCp7.

Evidence by Fluorescence Correlation Spectroscopy of NCp7-induced promotion of TAR kissing complexes

Abstract

TAR is known to be a multifunctional region involved in transcriptional activation, packaging and reverse transcription of the HIV-1 genome. Based on the appearance of two dimeric TAR complexes with different mobilities in a native gel, TAR hairpin has been shown to harbor a dimerization signal. The dimerization of TAR hairpin induced by NCp7 involves the exposure by NCp7 chaperone properties of a partially hidden palindromic sequence within the top half of TAR. In order to further characterize the formation of the TAR homodimers and to elucidate the role of the NCp7 in this process, we performed fluorescence correlation spectroscopy on rhodamine (Rh6G)-labelled TAR sequence in the presence and in the absence of NCp7. We could to evidence by one color FCS the NCp7-induced promotion of of TAR and cTAR homodimerization . The promotion of the TAR and cTAR dimer formation was not achieved with (11-55)NCp7, a N_{term} -truncated mutant that lacks the aggregating properties of NCp7, showing that the non-specific interaction of the N_{term} domain of NCp7 plays a critical role in this process. Nevertheless, the (11-55)NCp7 was able to induce dimerization of short hairpin DNA sequences in a manner that strongly depends of the stability of the nucleic acids. The NCp7-induced homodimerization of HIV sequences could be of importance in HIV-1 recombination events occurring during reverse-transcription. All these observations also provide evidence that FCS constitutes a versatile approach to study dimerization of macromolecules.

Introduction

The nucleocapsid protein NCp7 of human immunodeficiency virus type 1 (HIV-1) is a small, basic protein tightly associated with genomic RNA in the mature infectious virus [275]. NCp7 is characterized by two copies of a highly conserved retroviral-type, zinc-finger motif, also called CCHC motif, that binds zinc ions strongly [212, 219] in a constrained structure [206, 224, 379]. Besides its histone-like activity [342], NCp7 either as the mature

3.3 Applications

protein or in the Pr55 polyprotein precursor, plays several critical roles in the viral life-cycle. NCp7 is thought to interact specifically with the genomic RNA, enabling its selective encapsidation into budding particles [380, 251]. NCp7 also stimulates dramatically the two obligatory HIV-1 strand transfer during reverse transcription. In addition, NCp7 possesses a nucleic acid chaperone activity [244, 277, 381] and thus, directs the rearrangement of a nucleic acid molecule into the most stable conformation. This appears to be linked to the ability of NCp7 to lower the energy barrier for breakage and reformation of base-pairs. The chaperone activity of NCp7 is thought to be critical for efficient and complete proviral DNA synthesis. The HIV-1 contains two-identical (or nearly identical) copies of the RNA genome within the virion. These RNA genomes are non-covalently joined together as a dimer. Dimerization involves a specific sequence within the HIV-1 genome called the Dimerization Initiation Site (DIS), located in the 5'LTR of the genome that is critical for viral replication (for a review see [14]). DIS forms an homodimer through a kissing-loop interaction, further stabilized under an extended duplex form. The maturation of the DIS dimer is thought to be regulated by NCp7. Andersen & al.[296] showed that a phylogenetically conserved palindrome in the TAR RNA forms stable dimers independently of the DIS sequence when exposed to the NCp7 protein. The interaction between two TAR hairpins, which they characterized using a combination of bioinformatics, biochemistry, and atomic force microscopy (AFM) techniques, provides a molecular explanation for the observed role of TAR in packaging and reverse transcription of the viral RNA. Using fluorescence correlation spectroscopy, we were able to evidence the NCp7-induced formation of TAR RNA homodimers. Our objective was then to extend their structure-activity relationship by using DNA structures (cTAR and mutants thereof) relevant of the copied RNA in the RTion. In this study, we used TMR-labeled TAR, cTAR sequence, the DNA complementary sequence of TAR, as well as different mutants designed, on the basis of the proposed mechanism for TAR dimerization, to alter dimer formation. We observed that NCp7 induced TAR dimerization but not (11-55)NCp7, a mutant that lacks the aggregating properties of NC. We also found that (11-55)NCp7 favors the formation of stable dimers of short DNA sequences in the absence of internal bulges. The NCp7-induced homodimerization of HIV sequences could be of importance in the understanding of the recombination process that occurs du-

3.3 Applications

ring the RTion step of the viral replication.

Materials

NCp7 or NC(11-55) peptide were synthesized as described [208] and stored lyophilized in its zinc-bound form. Their purity were greater than 98%. An extinction coefficient of $5700M^{-1}cm^{-1}$ at 280 nm was used to determine its concentration. Singly labeled DNA oligonucleotides were synthesized at a 0.2 mmol scale by IBA GmbH Nucleic Acids Product Supply (Göttingen, Germany). The 5' terminus of the oligonucleotides was labeled by Rh6G via an amino-linker with a six carbon spacer arm. Oligonucleotides were purified by the manufacturer by double reverse-phase HPLC. Experiments were performed in 25 mM Tris-HCl pH 7.5, 30 mM NaCl, 0.2 mM $MgCl_2$ unless specified otherwise.

FCS setup and data analysis

FCS measurements were performed on a two-photon platform, as previously described [267, 382]. Two photon excitation is provided by a mode-locked Tsunami Ti :Sapphire laser pumped by a Millennia V solid state laser (Spectra Physics). Pulses of about 100 fs were produced at a wavelength of 850 nm . FCS measurements were performed on an Olympus IX70 inverted microscope equipped with a Märzhäuser motorized microscope stage. After a beam expander, the parallel infrared laser beam was focused into the sample by a water-immersion objective ($60\times$ 1.2-NA Olympus). The back aperture of the objective lens was slightly overfilled, creating a diffraction-limited focal spot of about $0.6\text{ }\mu\text{m}$ in diameter. The measurements were carried out in a 8 Well Lab-Tek II coverglass system, using a $400\text{ }\mu\text{L}$ volume per well. The focal spot was set about $20\text{ }\mu\text{m}$ above the coverslip. The resulting fluorescence was collected through the same objective and directed by a dichroic mirror (COWL 750 nm; Coherent) to the lateral output of the microscope. After rejection of residual IR light by a short pass filter (E700SP, Chroma), the emitted photons were focused by a 200 mm achromatic lens on the $64\text{ }\mu\text{m}$ core of a multimode optical fiber coupled to an avalanche photodiode (SPCM-AQR-14-FC, EG&G). The detector signal was correlated online by an ALV-5000E correlator (ALV, Germany). At this step, the normalized autocorrelation

3.3 Applications

function, $G(\tau)$ was calculated from the fluorescence fluctuations, $F(t)$, by :

$$G(\tau) = \frac{\langle \delta F(t) \cdot \delta F(t + \tau) \rangle}{\langle F(t) \rangle^2} \quad (3.13)$$

where τ is the lag time and $\langle F(t) \rangle$ is the mean fluorescence intensity. The analysis of $G(\tau)$ can provide information about the underlying mechanisms responsible for the intensity fluctuations such as diffusion of the particles, electronic transition within the molecules and transitions between states of different brightness. For an ideal case of freely diffusing monodisperse fluorescent particles undergoing a triplet-like blinking process in a Gaussian excitation volume, the correlation function, $G(\tau)$, calculated from the fluorescence fluctuations can be fitted according to :

$$G(\tau) = \frac{1}{N} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{S^2} \frac{\tau}{\tau_d}\right)^{-0.5} \left(1 + \frac{f_t}{1 - f_t} e^{-\frac{\tau}{\tau_t}}\right) \quad (3.14)$$

where τ_d is the apparent diffusion time (a parameter that is inversely related to the diffusion constant of the molecule), N is the mean number of molecules within the excitation volume, S is the ratio between the axial and lateral radii of the sample volume, f_t is the mean fraction of fluorophores in their dark state and τ_t is the triplet state lifetime. The excitation volume is about $0.3 \mu\text{m}^3$ and S is about 3. Using carboxytetramethylrhodamine (TMR) in water as a reference ($D_{TMR} = 421 \mu\text{m}^2 \cdot \text{s}^{-1}$), the diffusion coefficient, D_{exp} , of the labeled oligonucleotides was calculated by :

$$D_{exp} = D_{TMR} \times \frac{\tau_{da}(TMR)}{\tau_{da}(oligo)}$$

where $\tau_{da}(TMR)$ and $\tau_{da}(oligo)$ are the measured correlation times for TMR and the oligonucleotide, respectively. As a compromise between photobleaching and a good signal to noise ratio, we select a power of 5 mW . At this power, the photon count rate per molecule is about 5 kHz . Typical data recording times were 10 min.

3.3 Applications

Results

NCp7 induces TAR homodimers at saturation concentrations

We performed fluorescence correlation spectroscopy on rhodamine labeled TAR sequence in presence and in absence of NCp7 in order to characterize the formation of the TAR homodimers and the role of the NCp7 in this process. FCS method analyzes the fluctuations of fluorescence intensity in the very small volume (about 0.25 fl) provided by a two photon excitation. This analysis provides information on the phenomena that generate these fluctuations but also on the average number of fluorescent molecules present in the excited volume. For the singly-labeled derivatives, the fluorescence fluctuations are thought to mainly occur from diffusion in and out the excited volume and from a conversion between a fluorescent state and a nonfluorescent state of the Rh6G resulting from the dynamic of the probe in respect to the oligonucleotide. In line with this hypothesis, the correlation function of Rh6G-5'-TAR could be adequately fitted with Equation 3.14, providing an average number, N , of molecules in the excited volume. The diffusion constant, D_{exp} , calculated from the correlation time (Table 3.1) was in agreement with the theoretical diffusion constant, D_{th} , calculated by assuming that TAR may be modeled as a rod-like double-stranded DNA of 26 bp :

$$D_{th} = \frac{k_b T}{3\nu\pi} \frac{1}{L} \ln \left(\frac{L}{l} + \gamma \right) \quad (3.15)$$

where k_b is the Boltzmann constant, T is the absolute temperature and ν is the viscosity of the solution. γ designates the end-correction parameter and is about 0.39. Assuming a rise per base of 3.4Å, the length, L , of the rodlike RNA, was calculated to be about 95Å. Finally, the hydrodynamic diameter, l , was assumed to be comprised between 20.5 and 28Å. Addition of NCp7 to TAR induces a significant change in the correlation curve (Fig. 3.3). The number of diffusing species measured in FCS throughout the focal volume were strongly reduced by a factor of slightly more than two at saturating concentrations of NCp7. Moreover, an increase by a factor of about two of the count-rate per molecule showed that the newly-formed species were doubly-labelled.

While homodimers formed with high concentrations of NCp7 (a ratio nucleotide/protein of 8 was needed), a decrease of the NCp7 concentration strongly impeded the TAR homo-

3.3 Applications

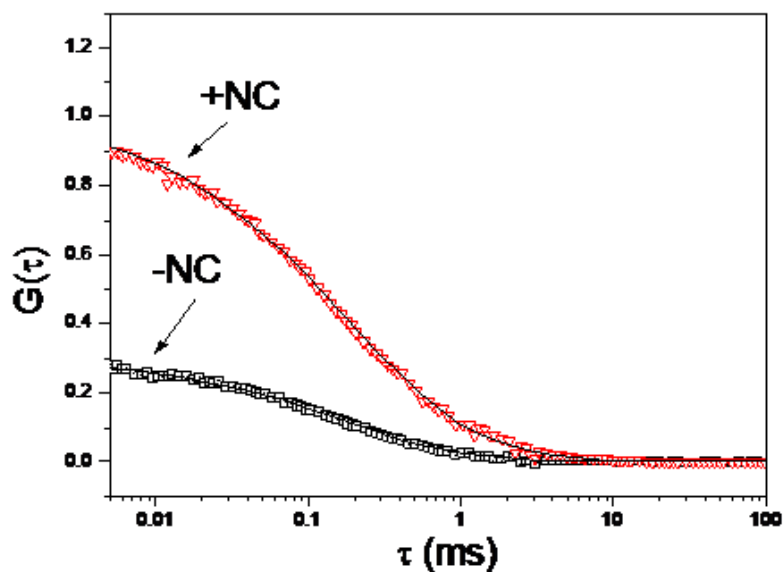


FIGURE 3.3 – Effect of NCp7 on TAR-Rh6G autocorrelation curve

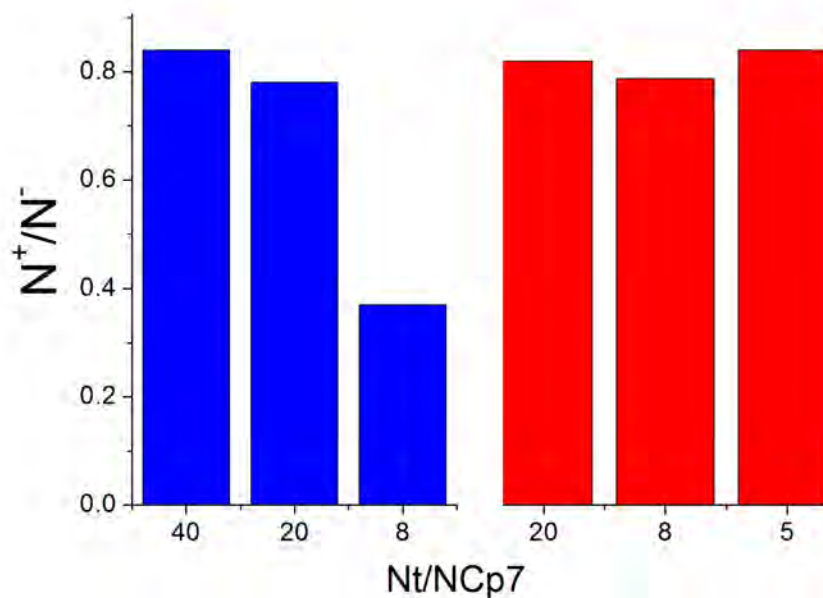


FIGURE 3.4 – Effect of NCp7 (blue) or (11-55)NCp7 (red) on TAR-Rh6G homodimer formation.

dimer formation (Figure 3.4). The TAR homodimers thus mainly formed when the concentration of NCp7 allowed a sufficient level of NCp7 coating, close to the one reported in

3.3 Applications

the viral particle. These levels correspond to concentrations that allowed the formation of aggregates. This strongly suggested that the aggregating activity of NCp7, which requires a sufficient protein coating of the nucleic-acid, is needed to allow the formation of TAR homodimers. This was confirmed by the inability of NC(11-55) to form TAR homodimers. Indeed, the addition of (11-55)NCp7 protein, a mutant of NCp7 which lacks N_{term} domain of NCp7 and thus less aggregates nucleic acids, to a diluted Rh6G-TAR solution also induced changes in the autocorrelation curve, but while the diffusion coefficient decreased significantly, confirming the binding of the peptide to the oligonucleotide, the number of molecules was not significantly affected.

Taken together, our data confirmed the suggested NCp7-induced TAR homodimer formation in a manner that strongly depend on the aggregating properties of NCp7.

NCp7 but not (11-55)NCp7 promotes cTAR homodimerization

Based on the results obtained with TAR, we investigated the ability of NCp7 to promote cTAR homodimer formation. As for TAR, addition of the NCp7 protein to a diluted Rh6G-cTAR solution induces a significant change in the autocorrelation curve. The average number of apparent diffusing molecules was similarly decreased by a factor of approximately two. The diffusion rate was also altered. Using different concentrations of NCp7 and/or cTAR, we also observed that the cTAR concentration did not significantly modify the dimer formation unless the quantity of protein was kept sufficient in respect to the concentration of cTAR. As for TAR, NC(11-55) was not able to induce cTAR homodimers (Table 3.1).

(11-55)NCp7 promotes the homodimerization of stable short hairpin-loop DNA

We further characterized the interaction of NC(11-55) with the various cTAR derivatives by FCS experiments.

As for cTAR, addition of the (11-55)NCp7 peptide to (14-39)cTAR induced a significant change in the correlation curve but the number of molecules was not significantly affected, although the diffusion coefficient decreased significantly, confirming the binding of the peptide to the oligonucleotide (Figure 3.5 and table 3.2).

3.3 Applications

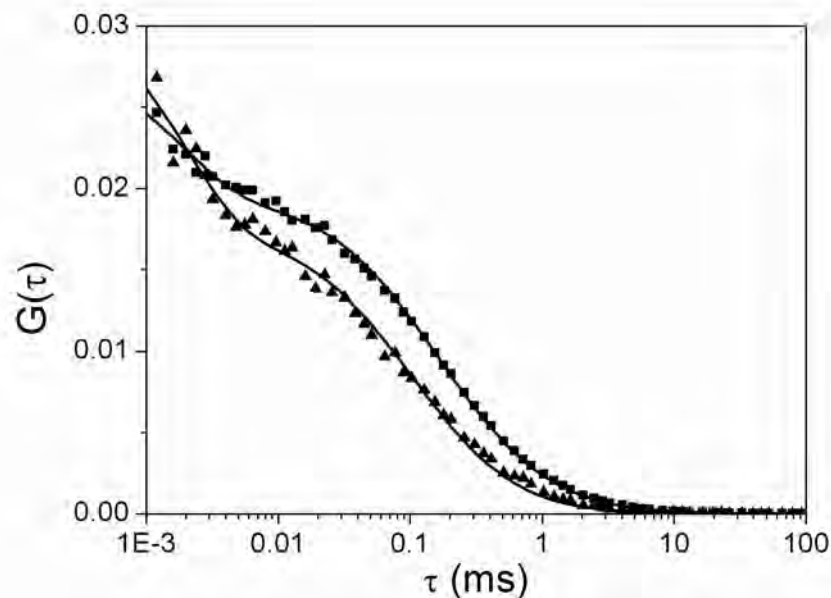


FIGURE 3.5 – Interaction of (11-55)NCp7 with Rh6G-5'-(14-39)cTAR as monitored by two photon FCS.

The autocorrelation curves of 400nM of Rh6G-5'-(14-39)cTAR were measured either in the absence (triangles) or in the presence (squares) of (11-55)NCp7. The solid lines correspond to fits of the experimental points with Equation 3.14. The triplet-like lifetime, τ_t , is about 2 to 3 μs

To calculate D_{th} , the peptide was assumed to saturate the oligonucleotide. However, since the dimensions of NC(12-55) are between 10Å and 25Å, it readily appears that the width of the oligonucleotide/peptide complex becomes comparable to its length. As a consequence, the rod-like model was substituted by the spherical model and D_{th} was calculated by the Stokes-Einstein equation : $D_{th} = k_b T / 6\pi \nu r$. An excellent agreement was found between D_{exp} and D_{th} , suggesting that the oligonucleotide may be largely coated by the peptide. Similar conclusions were obtained when (14-39)cTAR was substituted by A_{33} (14-39)cTAR, A_{35} (14-39)cTAR or ΔL (14-39)cTAR (Table 3.2).

In sharp contrast, addition of (11-55)NCp7 to ΔB (14-39)cTAR or (19-32)cTAR decreased the number of molecules by a factor of two, suggesting that (11-55)NCp7 promotes the dimerisation of these sequences. Interestingly, the NC-induced formation of the homodimers appeared for much lower protein concentration. Indeed, as far as there is in average one NCp7 per DNA sequence, the number of diffusing molecules appeared to be divided by about two. These dimers may correspond either to kissing complexes (where the sequences anneal

3.3 Applications

mainly through the partially autocomplementary sequences of their top loops) or extended duplexes (where the annealing extends over the whole sequence). To discriminate between these two types of dimers, we incubated equimolecular amounts of doubly labeled and non labeled $\Delta B(14-39)cTAR$ or $(19-32)cTAR$ derivatives in the presence of $(11-55)NCp7$. Statistically, formation of dimers should lead to 25% of dimers with non labeled monomers, 25% of dimers with doubly labeled monomers and 50% of dimers with one non labeled and one doubly labeled monomers (Figure 3.6 (a)). In extended duplexes, the Rh6G dye is moved far away from its quencher in the latter species, inducing a large fluorescence increase. As a consequence, the final fluorescence should be about half that obtained by mixing equimolecular amounts of singly Rh6G-labeled with non labeled oligonucleotides. In sharp contrast, since the annealing in kissing complexes involves regions that are far from the terminus of $cTAR$ derivatives, only limited fluorescence changes are expected and thus the final fluorescence should be far less than that of the corresponding dimers with singly labeled derivatives. The very limited NC-induced fluorescence changes observed with both $\Delta B(14-39)cTAR$ and $(19-32)cTAR$ derivatives clearly suggest that $(11-55)NCp7$ promotes the formation of kissing complexes.

Mechanism insights into the NCp7-induced cTAR hairpin homodimerization

To identify the nature of the $(11-55)NCp7$ -induced homodimers formed with $cTAR$, we performed a mutational analysis of the the NC-induced homodimer formation of $\Delta B(14-39)cTAR$. Mutants were designed according to a model adapted from Andersen's model [296]. In this model, $cTAR$ dimerization induced by $NCp7$ gives rise to a kissing-loop complex further stabilized in an extended duplex. The initial kissing-loop interaction involves 10 base-pairs of a highly conserved palindromic sequence in the top half of $cTAR$. Then, the extend of this interaction on 4 bases on each side of the initial kissing complex followed by the formation of a second parallel duplex, allows the formation of a more stable duplex and result in an extended kissing complex (Figure 3.7).

The model of Andersen can also be transposed to $\Delta B(14-39)cTAR$. Although steady-state experiments using doubly-labelled $\Delta B(14-39)cTAR$ showed that the extended duplex did not formed, two types of kissing intermediates can exist (Figure 3.8) that we will be

3.3 Applications

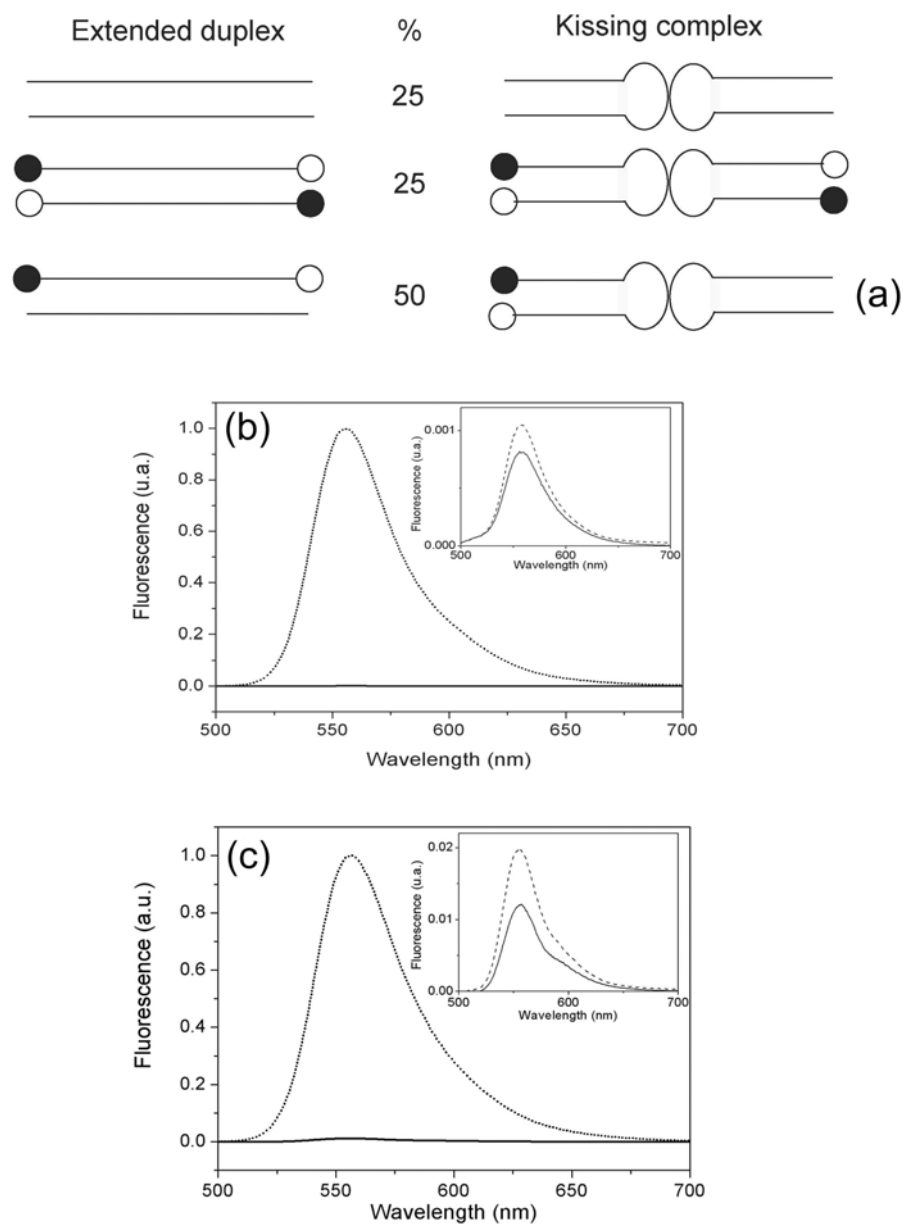


FIGURE 3.6 – Dimerization of $\Delta B(14-39)cTAR$ and $(19-32)cTAR$ induced by $(11-55)NCp7$. Equimolecular concentrations of non labeled and doubly labeled oligonucleotides were mixed with $(11-55)NCp7$.

(a) Theoretical composition of the mixtures when either extended duplexes or kissing complexes are formed. The empty and closed circles correspond to Rh6G and DABCYL dyes, respectively.

(b) Fluorescence spectra of the mixture of Rh6G-5'- $\Delta B(14-39)cTAR$ -3'-DABCYL with non labeled $\Delta B(14-39)cTAR$ either in the absence (solid line) and in the presence (dashed line) of NC(12-55). For comparison, the spectrum of Rh6G-5'- $\Delta B(14-39)cTAR$ with non labeled $\Delta B(14-39)cTAR$ in the presence of $(11-55)NCp7$ (dotted line) was recorded. Inset, magnification of the spectra with the doubly labeled oligonucleotides.

(c) Fluorescence spectra with $(19-32)cTAR$. Experiments were performed as in b.

3.3 Applications

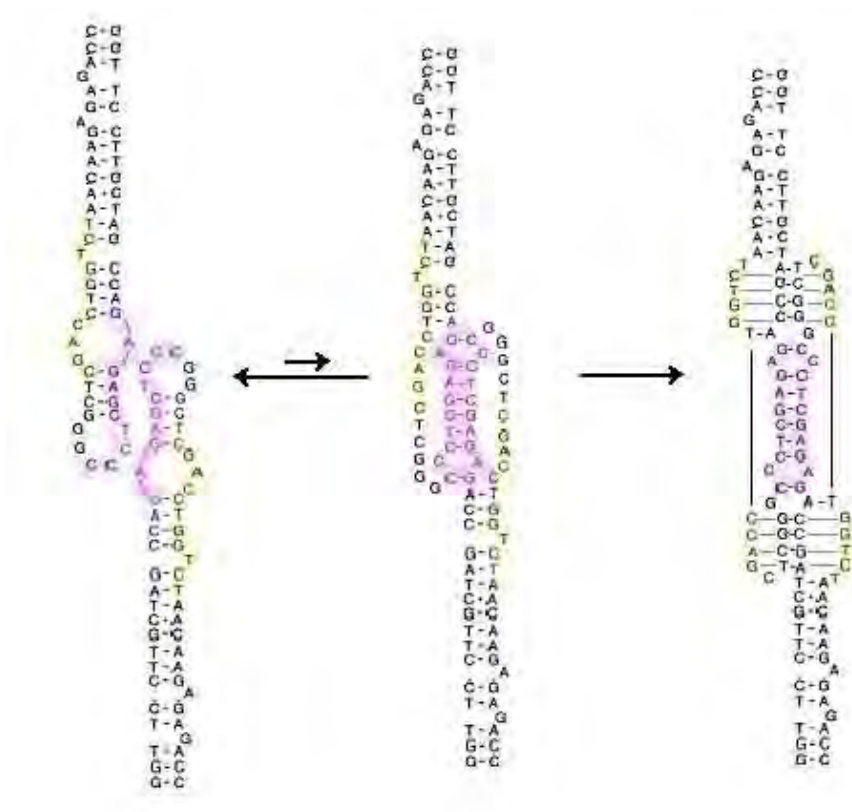


FIGURE 3.7 – Mechanistic model proposed to explain NCp7-induced cTAR homodimer formation (adapted from [296])

further referred as KC (for kissing complex) or EKC for (extended kissing complex).

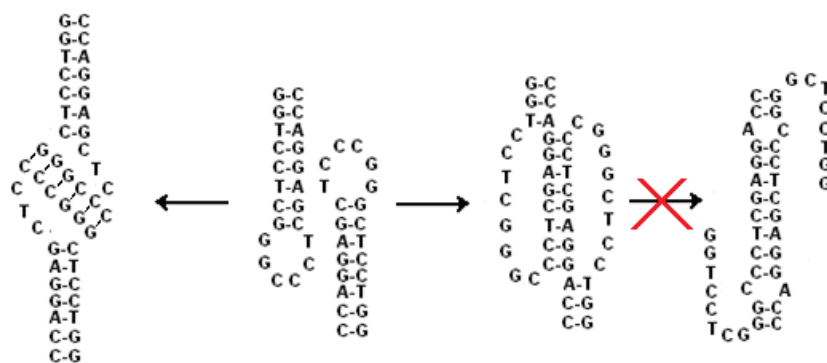


FIGURE 3.8 – Model of the two types of kissing complexes that can be implied in the NCp7-induced $\Delta B(14-39)cTAR$ homodimer formation (KC on the left, EKC on the right)

3.3 Applications

Three mutants have been designed. The $\Delta B(14-39)cTAR$ iL with an inverted loop disrupts the palindromic sequence within the (14-39)cTAR and thus impede the formation of the EKC. It also slightly modifies the formation of the kissing complex (Figure 5.7a). The $\Delta B(14-39)cTAR$ T that is expected to destabilize the KC and on the contrary, stabilize the EKC by the formation of one supplementary base-pair (Figure 5.8a). At least, the $\Delta B(14-39)cTAR$ TGM that destabilizes the stem of the $\Delta B(14-39)cTAR$ should provide information about the role of the stem stability in this process (Figure 3.9c).

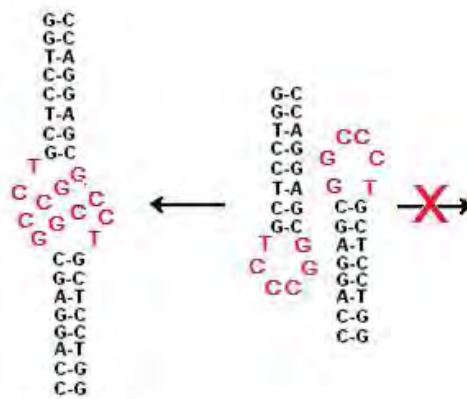
Interestingly, none of these three mutants was able to form homodimer to the same extent as the $\Delta B(14-39)cTAR$ (Table 3.3). As a consequence, it is likely that the dimerization within $\Delta B(14-39)cTAR$ occurs through different possible pathways, and neither the KC nor the EKC constitute the homogeneous way for the $\Delta B(14-39)cTAR$ rearrangements. In addition, and since NC is reported to rearrange nucleic acids into their most stable conformation, the kissing complexes may constitute the most stable forms. Finally, the stability of the stem (i.e. the absence of internal bulges) appeared to be critical for the homodimer formation.

Conclusions

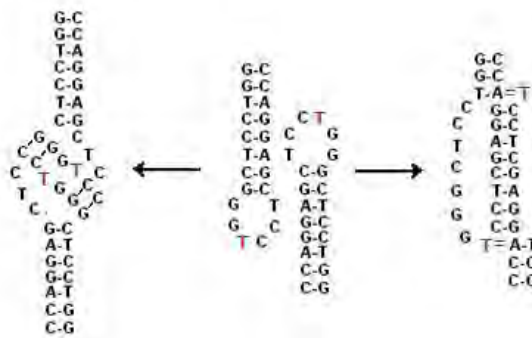
All these observations allow drawing a picture where NCp7 favors the formation of nucleic-acid homodimer. The aggregation properties of NCp7, although not sufficient, appeared to play a critical role in this process, probably by promoting the contact between the different nucleic-acids sequences. The activation of homodimer formation probably translates the well reported annealing component of NC chaperone activity. During reverse transcription, kissing complexes are thought to occur between TAR RNA and DNA. In this respect, it may be deduced that NCp7 not only stimulates the transition between kissing complexes and extended duplexes, but also promotes the formation of the kissing complexes, themselves. We also evidenced that the homodimer formation was not limited to native HIV-1 sequences. The phenomenon thus appeared general, provided the stability of the final duplex is higher than that of the monomers.

The amount of NCp7 needed for the homodimer to form appeared to strongly differ bet-

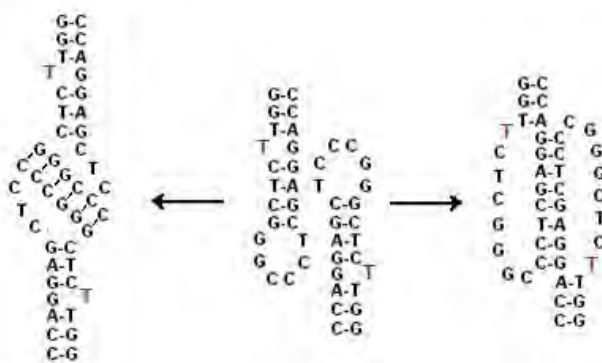
3.3 Applications



(a)



(b)



(c)

FIGURE 3.9 – Mutation-induced alterations of the KC and EKC formation

3.3 Applications

ween the NA sequences. The homodimers of sequence as TAR or cTAR clearly required saturating concentrations of NCp7, or at least concentrations allowing aggregates to form. On the other hand, the formation of homodimers involving very short and stable DNA hairpin loops was efficient when only one (11-55)NCp7 was bound per sequence. Thus, the mechanisms governing the formation of homodimers likely differed in the two types of NA sequences. In addition, the dimer formation appeared NA sequence specific. Indeed, we observed homodimers for $\Delta B(14-39)cTAR$ and $(19-32)cTAR$ but not for $(14-39)cTAR$. If, the stability of the stem appeared to be important, the reasons for such differences of behavior remained elusive. Since all the sequences appeared coated with NCp7, one explanation could come from the location of the NC-binding sites within these sequences. Indeed, in line with a conformational capture process, the binding of NCp7 is thought to preferentially occur in proximity of bulges, internal loops, mismatches within imperfect double stranded NA. In $(14-39)cTAR$ for example, NCp7 is preferentially bound at the 3' end of the sequence, downstream the internal loop (Mauffret O., personal communication). As far as NCp7 is able to destabilize the stem, NCp7 can bind almost all over the sequence. On the opposite, NCp7 is "concentrated" at the apical loop in the presence of perfect and stable stems. This is for example the case for the NMR described binding of NCp7 on SL2, SL3 or PBS. In this context, the binding of NCp7 at on the loop would explain the NCp7-induced formation of homodimer of stable hairpin sequences.

Taken together, these observations provide an insight into the understanding of the activity of NCp7 during the reverse transcription, notably its ability to enhance homologous recombination during synthesis. On a technical point of view, our data confirm that FCS constitutes a versatile approach to study dimerization of macromolecules.

| | | R (nt/NC) | N | τ_{diff} | $Diffusion$ $\mu m^2.s^{-1}$ | $Brilliance$ $Hz/molecule$ | $\frac{N^+}{N^-}$ | $\frac{CR^+}{CR^-}$ | $\frac{D^+}{D^-}$ |
|-------------|--------------------|----------------|-------|---------------|---------------------------------|-------------------------------|-------------------|---------------------|-------------------|
| TAR | (1-55)NCp7 | - | 18.60 | 0.13 | 106 | 2.35 | | | |
| | | 40 | 15.7 | 0.13 | 106 | 3.72 | 0.84 | 1.58 | 1.10 |
| | | 20 | 14.5 | 0.12 | 100 | 4.64 | 0.78 | 1.98 | 1.27 |
| | | 8 | 6.8 | 0.11 | 93 | 5.05 | 0.37 | 2.15 | 0.94 |
| TAR | (11-55)NCp7 | - | 17.90 | 0.13 | 106 | 2.27 | | | |
| | | 20 | 14.0 | 0.12 | 101 | 2.69 | 0.78 | 1.27 | 0.82 |
| | | 10 | 13.4 | 0.12 | 101 | 3.62 | 0.75 | 1.60 | 1.04 |
| | | 5 | 16.4 | 0.12 | 100 | 5.98 | 0.92 | 2.64 | 1.05 |
| cTAR | (1-55)NCp7 | - | 38.50 | 0.09 | 156 | 2.82 | | | |
| | | 40 | 31.0 | 0.11 | 131 | 3.61 | 0.80 | 1.28 | 1.19 |
| | | 20 | 29.3 | 0.11 | 130 | 4.89 | 0.76 | 1.32 | 1.22 |
| | | 8 | 16.5 | 0.12 | 122 | 5.46 | 0.43 | 1.94 | 1.28 |
| cTAR | (11-55)NCp7 | | 48.90 | 0.12 | 152 | 1.54 | | | |
| | | 20 | 41.1 | 0.14 | 133 | 2.21 | 0.84 | 1.44 | 1.14 |
| | | 10 | 40.0 | 0.15 | 132 | 3.13 | 0.82 | 2.04 | 1.15 |
| | | 5 | 34.6 | 0.17 | 85 | 3.31 | 0.95 | 2.02 | 1.33 |
| | | 1 | 33.7 | 0.17 | 83 | 3.45 | 0.93 | 2.11 | 1.36 |

TABLE 3.1 – Effect of NCp7 on the FCS parameters of TAR or cTAR.

3.3 Applications

| | r | D_{exp} ($cm^2.s^{-1}$) $\times 10^7$ | D_{th} ($cm^2.s^{-1}$) $\times 10^7$ | $\frac{N_{+NC}}{N_{-NC}}$ |
|-----------------------|----------|--|---|---------------------------|
| (14-39)cTAR | - | 9.0 ± 0.2 | $8.2 - 11^b$ | |
| | 5 | 6.3 ± 0.2 | $4.6 - 6.7^c$ | 1.2 ± 0.4 |
| $\Delta B(14-39)cTAR$ | - | 6.9 ± 0.2 | $7.8 - 11^b$ | |
| | 5 | 7.1 ± 0.1 | $7.8 - 9.7^b$ | 0.6 ± 0.1 |
| $\Delta L(14-39)cTAR$ | - | 9.2 ± 0.5 | $7.4 - 11^b$ | |
| | 5 | 7.1 ± 0.1 | $5.1 - 7.9^c$ | 1.0 ± 0.1 |
| (19-32)cTAR | - | 9.5 ± 0.2 | $4.1 - 9.7^b$ | |
| | 5 | 8.1 ± 0.4 | $8.3 - 11^b$ | 0.5 ± 0.2 |
| $A_{35}(14-39)cTAR$ | - | 7.7 ± 0.4 | $8.2 - 11^b$ | |
| | 5 | 6.0 ± 0.1 | $4.6 - 6.7^c$ | 1.0 ± 0.1 |
| $A_{33}(14-39)cTAR$ | - | 7.9 ± 0.1 | $8.2 - 11^b$ | |
| | 5 | 5.6 ± 0.4 | $4.6 - 6.7^c$ | 1.1 ± 0.3 |

TABLE 3.2 – FCS parameters of the interaction of (11-55)NCp7 with cTAR derivatives^a

^a The experimental diffusion coefficient D_{exp} was calculated from the apparent diffusion time as described in Material and Methods. Both the apparent diffusion time and the average number of molecules in the illuminated volume are obtained by fitting the data to equation 3.14. The $\frac{N_{+NC}}{N_{-NC}}$ parameter describes the ratio of the average number of molecules in the presence of (11-55)NCp7 to that in the absence.

^b D_{th} calculated from equation 3.15 using the rod-like model

^c D_{th} calculated from the Stokes Einstein equation using the spherical model

| | r | D_{exp} ($cm^2.s^{-1}$) $\times 10^7$ | $\frac{N_{+NC}}{N_{-NC}}$ |
|---------------------------|----------|--|---------------------------|
| $\Delta B(14-39)cTAR$ | - | 6.9 ± 0.2 | |
| | 20 | 6.8 ± 0.2 | 0.6 ± 0.3 |
| | 10 | 7.2 ± 0.4 | 0.5 ± 0.2 |
| | 5 | 7.2 ± 0.3 | 0.6 ± 0.1 |
| $\Delta B(14-39)cTAR$ iL | - | 7.6 ± 0.3 | |
| | 20 | 7.5 ± 0.4 | 0.75 ± 0.2 |
| | 10 | 7.1 ± 0.3 | 0.70 ± 0.1 |
| | 5 | 6.9 ± 0.3 | 0.68 ± 0.2 |
| $\Delta B(14-39)cTAR$ T | - | 7.7 ± 0.3 | |
| | 20 | 7.5 ± 0.2 | 0.80 ± 0.2 |
| | 10 | 7.2 ± 0.3 | 0.85 ± 0.5 |
| $\Delta B(14-39)cTAR$ TGM | - | 10.1 ± 0.3 | |
| | 20 | 9.8 ± 0.2 | 0.82 ± 0.5 |
| | 10 | 8.7 ± 0.3 | 0.92 ± 0.1 |

TABLE 3.3 – FCS parameters of the interaction of (11-55)NCp7 with $\Delta B(14-39)cTAR$ derivatives

3.4 Quand la fluorescence n'est plus constante...

L'hypothèse que nous avons effectuée jusqu'à maintenant considère que la fluorescence du fluorophore est invariante durant le temps de diffusion dans le volume focal. Cette hypothèse n'est pas toujours vérifiée. En effet, des réactions intra ou intermoléculaires peuvent donner lieu à des fluctuations de fluorescence dans une échelle de temps plus rapide que celle du temps de diffusion. Plutôt que de redériver toutes les équations précédentes, on peut modifier simplement l'équation résultante 3.12 selon

$$G_{obs}(\tau) = G_{diff}(\tau).X(\tau)$$

où $X(\tau)$ traduit les fluctuations supplémentaires. Il va de soit que cette approximation n'est valide que pour les cas où le coefficient de diffusion n'est pas altéré par ces réactions [383, 384].

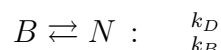
Dans le cas de l'existence d'un état triplet par exemple, le scintillement (*blinking*) induit peut être décrit par un déclin mono-exponentiel selon

$$X_{triplet}(\tau) = 1 - f_t + f_t.exp\left(-\frac{\tau}{\tau_{triplet}}\right)$$

qui est en général normalisé par $1 - f_t$ pour donner l'équation d'autocorrélation suivante :

$$G(\tau) = \frac{1}{N} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{S^2} \frac{\tau}{\tau_d}\right)^{-0.5} \left(1 + \frac{f_t}{1 - f_t} e^{-\frac{\tau}{\tau_t}}\right) \quad (3.16)$$

Bien évidemment cette démarche est applicable à toute réaction induisant des transitions entre un état brillant B et un état noir N selon une réaction d'équilibre du type :



pour laquelle

$$X_F(\tau) = 1 - F + F.exp\left(-\frac{\tau}{\tau_f}\right) \quad (3.17)$$

3.4 Quand la fluorescence n'est plus constante...

ce qui donne

$$G(\tau) = \frac{1}{N} \left(1 + \frac{\tau}{\tau_d}\right)^{-1} \left(1 + \frac{1}{S^2} \frac{\tau}{\tau_d}\right)^{-0.5} \left(1 + \frac{F}{1-F} e^{-\frac{\tau}{\tau_f}}\right) \quad (3.18)$$

dans laquelle $\tau_f = \frac{1}{k_D+k_B}$ est le temps de relaxation de la réaction, $F = \frac{k_D}{k_D+k_B}$ est le temps moyen pendant lequel la molécule est dans l'état noir. Enfin, si l'état noir N n'est pas complètement noir, l'expression de F devient :

$$F = \frac{k_D k_B (C R_B - C R_D)^2}{(k_D + k_B)(k_D C R_D^2 + k_B C R_B^2)}$$

Dans le cas où la diffusion est affectée, l'équation 3.11 doit être généralisée afin de prendre en compte toutes les diffusions pouvant exister dans le système. Celles-ci sont pondérées par leur *count rate* relatif

$$G_{motion}(\tau) = \frac{1}{V_{eff}} \frac{\sum C R_i \langle C_i \rangle M_i(\tau)}{(\sum C R_i \langle C_i \rangle)^2} \quad (3.19)$$

où

$$M_i(\tau) = \left(1 + \frac{\tau}{\tau_{d_i}}\right)^{-1} \left(1 + \frac{1}{\left(\frac{r_0}{z_0}\right)^2} \frac{\tau}{\tau_{d_i}}\right)^{-0.5}$$

dans le cas d'une diffusion libre à 3 dimensions.

Caractérisation de la dynamique conformationnelle de la Rhodamine-6G liée de manière covalente à un ADN

Objectifs :

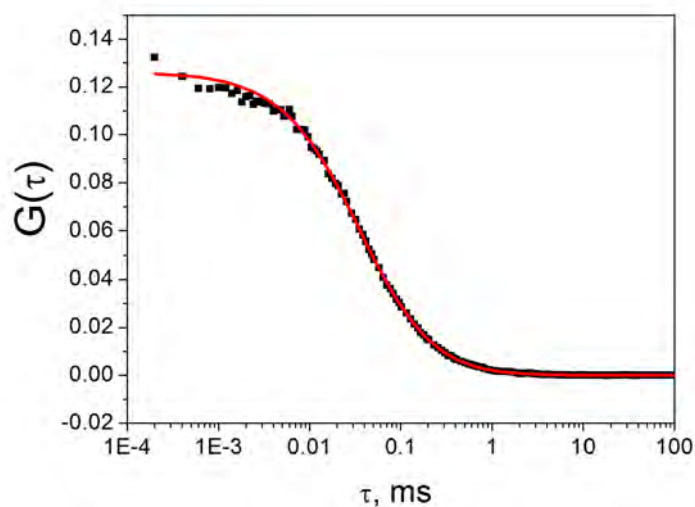
L'objectif de ces expériences était d'évaluer par FCS la dynamique de la Rhodamine 6G (Rh6G) attachée de manière covalente par un linker flexible à un oligonucléotide (ODN).

Contexte :

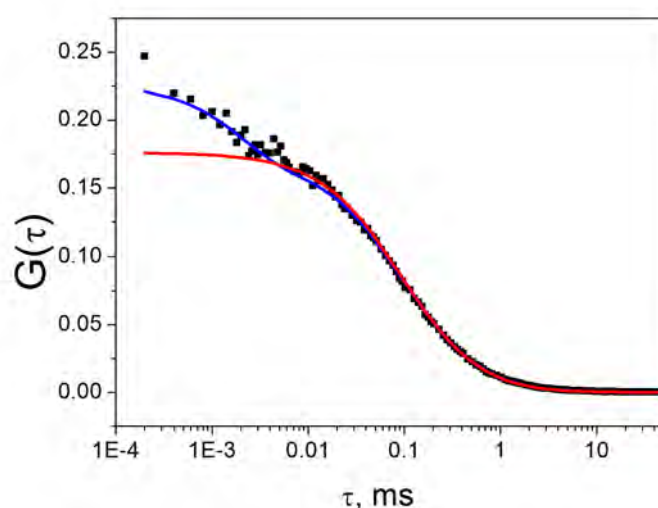
Alors que les courbes de FCS de Rh6G libre en solution sont décrites correctement par un modèle de diffusion 3D selon l'équation 3.12 (figure 3.10a), les courbes d'autocorrélation correspondant à des séquences ADN (cTAR ou (14-39)cTAR) marquées par la Rh6G divergent aux temps courts de ce modèle, suggérant que la fluorescence de la sonde n'est pas constante pendant la traversée du volume focal. Ces courbes nécessitent donc l'ajout d'une composante supplémentaire (figure 3.10b).

Étant donné que les fluctuations de fluorescence se produisent sur une gamme de temps plus rapide que celles causées par la diffusion des particules marquées, une séparation des dynamiques est possible et les courbes d'autocorrélation des séquences ADN marquées par la Rh6G ont pu être ajustées avec un modèle de diffusion simple associée à l'existence d'un état triplet. Les courbes d'autocorrélation des ODN marqués à la Rh6G ont donc été modélisées à l'aide de l'équation 3.16 considérant ainsi implicitement l'existence d'un état triplet [267, 268, 269, 227, 385, 386, 348]. Pour autant, la preuve de l'existence de cet état triplet n'a pas été apportée. D'autres phénomènes peuvent engendrer ce genre de fluctuations supplémentaires, comme l'existence d'un *quenching* transitoire de la sonde, résultant de l'interaction de la sonde avec les nucléosides situés dans son voisinage. Les bases de l'ADN, et en particulier les G, peuvent éteindre la fluorescence de la Rh6G. La dynamique de la sonde à l'égard de l'oligonucléotide avec lequel elle est liée peut donc être responsable de fluctuations additionnelles de fluorescence. Plusieurs éléments laissent à penser que cette dernière hypothèse est probable. Les différences entre le nombre apparent de molécules diffusant dans le volume focal et le nombre théorique attendu, calculé par spectroscopie d'absorption, suggèrent qu'il existe une proportion significative de molécules dans un état

3.4 Quand la fluorescence n'est plus constante...



(a)



(b)

FIGURE 3.10 – Exemples de courbes d'autocorrélations de Rhodamine-6G (a) et de (14-39)cTAR-Rh6G (b).

La courbe d'autocorrélation de cTAR ne s'ajuste pas correctement à l'aide d'un modèle de diffusion simple (représenté par la courbe rouge) et nécessite une composante supplémentaire pour être correctement décrit.

non-fluorescent pendant la traversée du volume focal. Ces différences sont constantes quelle que soit la concentration d'ODN marqué utilisée, ce qui permet d'écartier les problèmes d'adsorption des ODN sur les parois des chambres de mesure. D'autre part, l'addition de NCp7 sur ces ODN est responsable d'une diminution de l'amplitude de cette composante courte, indiquant que la NCp7 pourrait avoir une influence sur la dynamique de la sonde.

3.4 Quand la fluorescence n'est plus constante...

Afin de caractériser l'origine des fluctuations additionnelles de fluorescence, nous avons mesuré les courbes d'autocorrélation de fluorescence de (14-39)cTAR marqué à son extrémité 5' par de la Rh6G en absence et en présence de NCp7. Afin d'identifier l'origine dynamique et/ou photophysique de ces fluctuations, nous avons fait varier la viscosité du milieu ainsi que la puissance de la source d'excitation laser.

Résultats :

Les paramètres des courbes de FCS de 5'Rh6G-(14-39)cTAR enregistrées à différentes puissances lasers, différentes viscosités, en absence ou en présence de NCp7 sont résumés dans le tableau 3.4. Ces paramètres sont obtenus en ajustant les courbes expérimentales avec l'équation 3.17. Le temps de diffusion de la séquence 5'Rh6G-(14-39)cTAR, proche de $100 \mu s$ ($\tau_{Diff}(TMR) = 38 \mu s$) à $1 cP$, est en accord avec la taille et la forme d'une séquence en tige-boucle de 26 nucléotides. Néanmoins, le paramètre de fit du temps de diffusion de la composante rapide (quelques microsecondes) s'avère instable et largement dépendant de la fenêtre de fit. Ceci résulte principalement de la présence de bruit au niveau des premiers points du spectre d'autocorrélation.

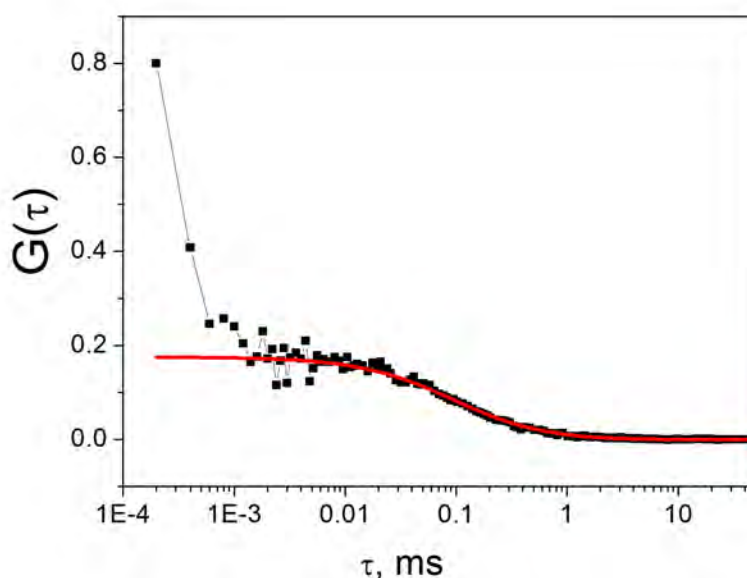


FIGURE 3.11 – Illustration d'une courbe d'autocorrélation bruitée aux temps courts en raison de la présence d'*afterpulses* de l'APD.

3.4 Quand la fluorescence n'est plus constante...

Afin de déterminer avec plus de précision les événements survenant aux temps courts de la courbe d'autocorrélation, l'émission de fluorescence a été séparée en deux faisceaux perpendiculaires qui sont chacun détectés sur une photodiode à avalanche (*Avalanche Photodiode* ou APD). Les réponses des deux détecteurs sont ensuite corrélées de manière croisée. Cette amélioration du montage expérimental permet de réduire largement la détection artefactuelle d'impulsions résultant de charges parasites au sein de l'APD (*afterpulses*). La probabilité de détecter simultanément deux *afterpulses* sur deux détecteurs est extrêmement faible. La courbe de cross-correlation résultante est donc exempte d'*afterpulses* et la résolution aux temps courts s'en trouve très largement améliorée. Le prix à payer pour cette meilleure résolution est la diminution par un facteur deux du signal de fluorescence.

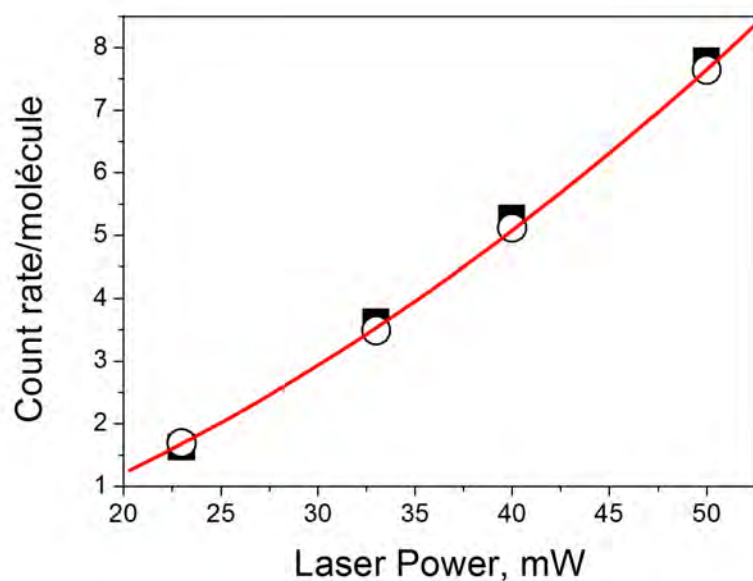
Afin de discriminer l'origine des fluctuations supplémentaires aux temps courts, les courbes ont été enregistrées avec différentes puissances laser. Ces expériences amènent plusieurs remarques :

1. la *count-rate* par molécule de Rh6G libre ou lié avec cTAR augmente de manière quadratique avec la puissance laser lorsque les puissances utilisées restent suffisamment faibles pour ne pas induire de blanchiment de la sonde (figure(3.12a)).
2. la composante courte n'apparaît pas avec la Rh6G libre lorsque l'intensité laser augmente.
3. la fraction de la composante rapide dans les courbes de 5'Rh6G-(14-39)cTAR ne semble pas affectée (figure 3.12b) .

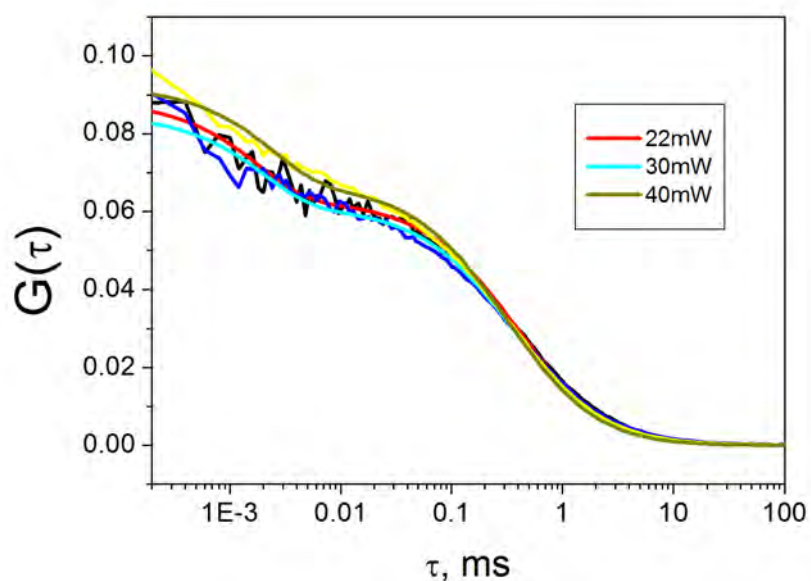
Ces observations suggèrent que le clignotement de la sonde liée à cTAR ne résulte pas de l'existence d'un état triplet.

Afin de déterminer si la composante additionnelle résulte d'un effet de quenching dynamique (collisionnel avec les bases du bas de la tige de (14-39)cTAR), les mêmes mesures ont été effectuées en présence de sucrose à une viscosité de 4cP ($\sim 35\%$ (w/w) sucrose) (tableau 3.4). On observe que le τ_{Diff} est augmenté d'un facteur 3 à 4, en accord avec l'augmentation de viscosité. Le τ_{fast} est également généralement augmenté, mais ces augmentations varient largement (entre 1.1 et 10 fois), probablement en grande partie due à l'imprécision des

3.4 Quand la fluorescence n'est plus constante...



(a)



(b)

FIGURE 3.12 – (a) Dépendance de la brillance de la Rh6G en fonction de la puissance laser à une viscosité de 1 cP (carrés) ou de 4 cP (ronds)

(b) La composante courte de la courbe d'autocorrélation de Rh6G-(14-39)cTAR dépend peu de la puissance laser, suggérant que cette composante ne résulte pas de la présence d'un état triplet

3.4 Quand la fluorescence n'est plus constante...

mesures. En effet, les valeurs obtenues pour le τ_{fast} à 1 cP sont à la limite de ce que l'on peut mesurer avec notre système.

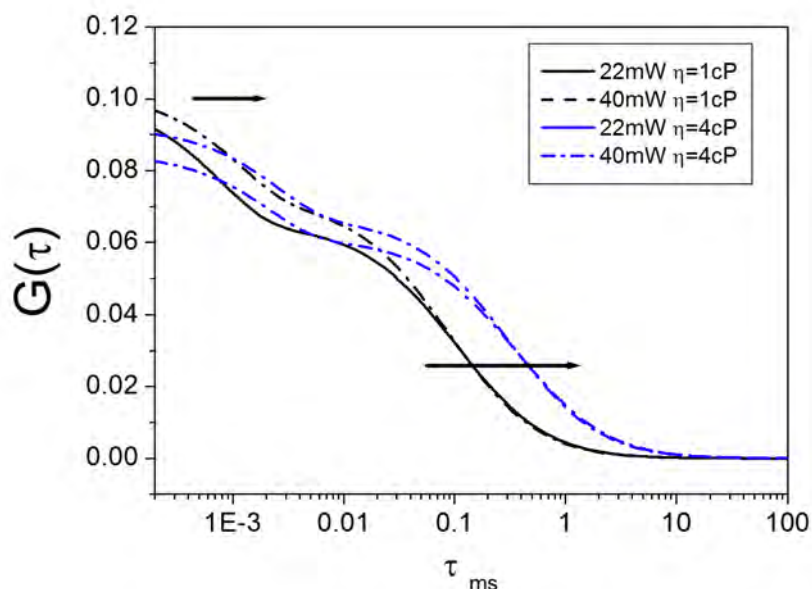


FIGURE 3.13 – Effet de la viscosité sur les courbes d'autocorrélation de (14-39)cTAR-Rh6G.

En outre, il est probable que la distribution des τ_{fast} compatibles avec un ajustement correct des données expérimentales ne puisse se résumer à une estimation ponctuelle. Ceci pourrait être le cas si la dynamique de la sonde ne résulte pas d'un système à deux états. Dans ce cas, la composante additionnelle ne suit plus une loi exponentielle mais plutôt une loi exponentielle étirée. Enfin, il est également possible que la viscosité n'ait pas une influence linéaire sur les mouvements de la sonde (comme c'est le cas par exemple dans le cadre de mesures d'anisotropie de fluorescence résolue en temps où la composante courte, associée au mouvement de la sonde, est en général indépendante de la viscosité du milieu). Une approche d'ajustement des modèles à nos données expérimentales par la méthode du maximum d'entropie permettrait d'accéder à ces distributions de temps de diffusion plausibles et donc, d'apporter certaines réponses à nos questions.

Conclusion :

Ces observations permettent de dire que l'existence d'une composante additionnelle au sein

3.4 Quand la fluorescence n'est plus constante...

des courbes d'autocorrélation d'ODN marqués par la Rh6G ne résulte probablement pas de l'existence d'un état triplet. Au contraire, les fluctuations de la fluorescence de la Rh6G résultent vraisemblablement d'une dynamique complexe de la sonde par rapport à l'ODN auquel elle est liée [387], induisant des changements de la fluorescence de la Rh6G par des mécanismes tels que du quenching. D'autre part, devant l'absence d'effets évidents et facilement accessibles de la NCp7 sur ces mouvements de la sonde, nous n'avons pas cherché à caractériser plus finement ces phénomènes. La parution au cours de la réalisation de ces expériences d'un article décrivant la nature de la dynamique de la Rh6G par rapport à un ODN [388], et reprenant largement des observations similaires à celles que nous avons faites, nous a conforté dans ce choix.

| | NCp7 | Viscosité | Puissance | N | τ_{Diff} | F_{fast} | τ_{fast} | CR1 | CR2 | $N^{-1} \cdot \sum CR$ |
|---------------------|------------|-----------|-----------|------|---------------|------------|---------------|-------|-------|------------------------|
| (14-39)cTAR-Rh6G | - | 1 | 23 | 14.8 | 0.108 | 0.35 | 0.8 | 10.18 | 8.05 | 1.23 |
| | - | 1 | 33 | 14.1 | 0.104 | 0.36 | 0.9 | 18.47 | 15.17 | 2.38 |
| | - | 1 | 40 | 13.9 | 0.092 | 0.29 | 1.1 | 26.63 | 21.78 | 3.48 |
| | - | 1 | 50 | 12.8 | 0.076 | 0.23 | 2.2 | 37.39 | 31.72 | 5.38 |
| | 5 nt :NCp7 | en cP | en mW | | en ms | | en μs | en Hz | en Hz | Hz/molécule |
| (14-39)cTAR-Rh6G | - | 4 | 23 | 18.1 | 0.322 | 0.37 | 1.1 | 11.28 | 9.94 | 1.17 |
| | - | 4 | 33 | 18.0 | 0.300 | 0.34 | 1.0 | 24.75 | 23.39 | 2.68 |
| | - | 4 | 40 | 18.7 | 0.316 | 0.25 | 1.4 | 36.68 | 34.50 | 3.80 |
| | - | 4 | 50 | 17.9 | 0.257 | 0.29 | 2.3 | 54.54 | 51.99 | 5.96 |
| (14-39) cTAR-Rh6G | + | 1 | 23 | 19.6 | 0.135 | 0.23 | 1.1 | 18.25 | 14.8 | 1.69 |
| | + | 1 | 33 | 19.5 | 0.126 | 0.28 | 0.7 | 37.01 | 30.96 | 3.49 |
| | + | 1 | 40 | 19.2 | 0.117 | 0.25 | 1.7 | 53.6 | 45.01 | 5.12 |
| | + | 1 | 50 | 18.0 | 0.101 | 0.22 | 2.6 | 73.67 | 63.95 | 7.64 |
| (14-39)cTAR-Rh6G Rh | + | 4 | 23 | 22.1 | 0.502 | 0.22 | 3.8 | 18.39 | 17.25 | 1.62 |
| | + | 4 | 33 | 22.2 | 0.578 | 0.33 | 5.2 | 40.72 | 39.63 | 3.62 |
| | + | 4 | 40 | 21.4 | 0.379 | 0.25 | 2.9 | 57.2 | 55.6 | 5.27 |
| | + | 4 | 50 | 21.2 | 0.329 | 0.28 | 4.1 | 83.87 | 80.98 | 7.78 |
| Rh6G | - | 1 | 18 | 0.63 | 0.040 | | | 0.6 | 1.3 | 2.9 |
| Rh6G | - | 1 | 23 | 0.61 | 0.034 | | | 2.5 | 2.6 | 7.3 |
| Rh6G | - | 1 | 31 | 0.60 | 0.029 | | | 1.3 | 3.6 | 9.8 |
| Rh6G | - | 1 | 43 | 0.62 | 0.021 | | | 1.3 | 3.8 | 13.0 |
| Rh6G | - | 4 | 18 | 1.43 | 0.127 | | | 1.2 | 2.9 | 2.9 |
| Rh6G | - | 4 | 23 | 1.12 | 0.142 | | | 4.1 | 4.1 | 7.3 |
| Rh6G | - | 4 | 31 | 1.05 | 0.129 | | | 2.5 | 6.7 | 8.7 |
| Rh6G | - | 4 | 43 | 1.05 | 0.146 | | | 3.3 | 9.3 | 12.0 |

TABLE 3.4 – Paramètres de FCS de (14-39)cTAR dans différentes conditions.

N correspond au nombre moyen de molécules dans le volume focal, τ_{Diff} et τ_{fast} au temps de diffusion apparents correspondant à la composante de diffusion et la composante aux temps courts, respectivement. F_{fast} est la fraction

3.5 Cross-Corrélation ou corrélation croisée

Alors que l'autocorrélation mesure la corrélation d'un signal par rapport à lui même, il est possible de corréler deux signaux différents entre eux : c'est ce que l'on appelle la corrélation croisée (ou cross-corrélation). Les signaux peuvent venir soit de fluorophores différents (dual-color FCS), soit de deux volumes focaux différents (dual-focus FCS ou two-focus FCS), soit encore d'un même signal de fluorescence qui serait séparé et détecté sur deux détecteurs.

3.5.1 Corrélation croisée à deux couleurs

Par analogie aux formules d'autocorrélation, l'équation 3.5 peut être généralisée pour 2 fluorophores :

$$\delta F_1(t) = \int_V W_1(r).Cr_1.\delta (C_1(r, t) + C_{12}(r, t)) dV$$

$$\delta F_2(t) = \int_V W_2(r).Cr_2.\delta (C_2(r, t) + C_{12}(r, t)) dV$$

avec

$W_{i(i=1,2)}(r)$ la distribution spatiale d'émission de fluorescence pour les espèces i

$C_{i(i=1,2)}(r, t)$ la concentration en espèce i simplement marquée

$C_{12}(r, t)$ la concentration en espèce doublement marquée

Si l'on considère des conditions idéales dans lesquelles

- $W_1(r) = W_2(r)$
- les spectres d'émission des deux fluorophores sont parfaitement séparables
- il n'existe pas de cross-talk entre les voies de détection

alors les formules donnent :

$$G_i(\tau) = \frac{\langle C_i \rangle M_i(\tau) + \langle C_{12} \rangle M_{12}(\tau)}{V_{eff} (\langle C_i \rangle + \langle C_{12} \rangle)^2} \quad \text{pour l'autocorrélation} \quad (3.20)$$

$$G_{\times}(\tau) = \frac{\langle C_{12} \rangle M_{12}(\tau) + \langle C_{12} \rangle M_{12}(\tau)}{V_{eff} (\langle C_1 \rangle + \langle C_{12} \rangle) (\langle C_2 \rangle + \langle C_{22} \rangle)} \quad \text{pour la cross-corrélation} \quad (3.21)$$

3.5 Cross-Corrélation ou corrélation croisée

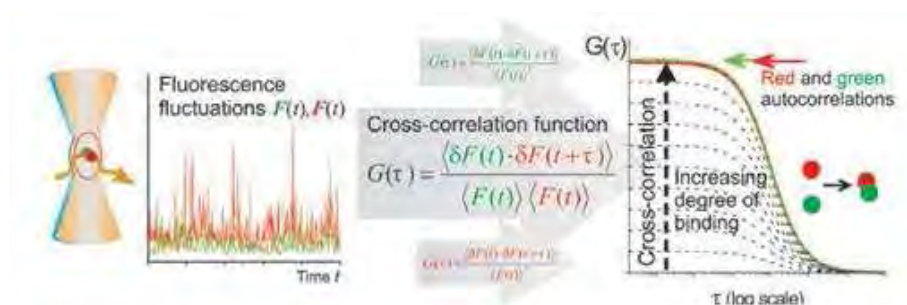


FIGURE 3.14 – Schéma illustrant le principe de la FCCS à deux couleurs

Afin de vérifier expérimentalement que le montage dont nous disposons permet de mesurer fidèlement la cross-corrélation d'espèces doublement marquées, deux séquences ADN complémentaires (cTAR et dTAR) de 55nt, marquées en 5' par de l'OregonGreen 488 pour dTAR et en 5' par du TexasRed pour cTAR, ont été hybridées à la température. Le duplex doublement marqué est séparé des séquences d'ADN simple brin résiduelles par électrophorèse sur gel d'agarose à 2% puis élué à l'aide d'un kit NucleoSpin Extract II. Les spectres d'excitation et d'émission de chacun des fluorophores du duplex doublement marqué sont représentés en figure 3.15, montrant la très bonne séparation des deux spectres d'émission par le miroir dichroïque. Il est à noter que l'excitation biphotonique permet d'exciter ces deux fluorophores avec un seul laser, ce qui permet de s'affranchir des problèmes d'alignement des volumes focaux. Les courbes d'autocorrélation des voies vertes et rouges ainsi que la courbe de cross-corrélation sont présentées figure 3.16.

Le traitement des données avec les équations 3.20 et 3.21 donne les résultats suivants :

| | N | τ | $G(0)$ |
|-------|------|--------|--------|
| Red | 54.6 | 0.348 | 0.018 |
| Green | 40.3 | 0.300 | 0.025 |
| Cross | 40.5 | 0.291 | 0.018 |

Le nombre d'espèces qui cross-corrèlent correspond au nombre de molécule déterminé par autocorrélation de la voie verte, ce qui montre sans ambiguïté que la concentration des espèces qui cross-corrèlent correspond bien à la quantité de duplex doublement marqués en solution.

3.5 Cross-Corrélation ou corrélation croisée

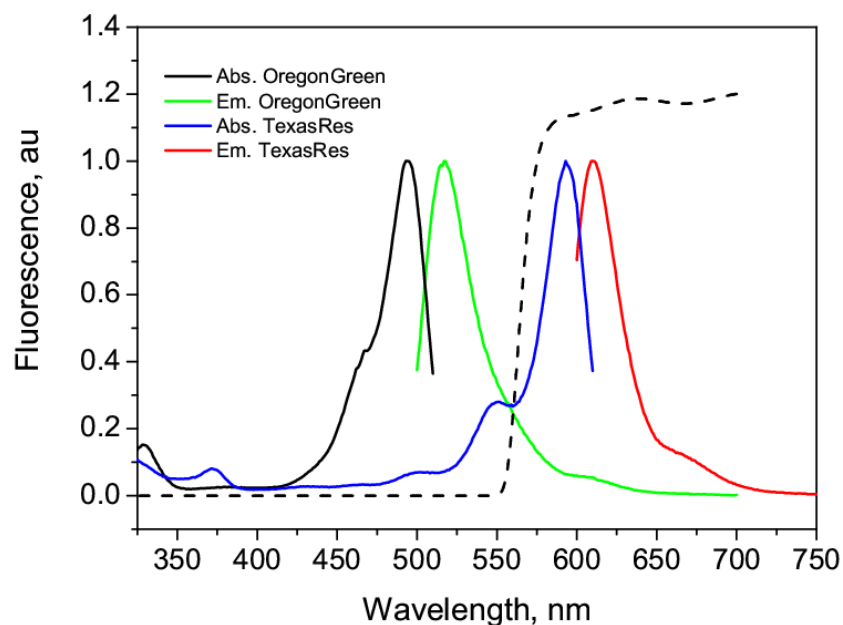


FIGURE 3.15 – Spectre d'excitation (noir) et d'émission de fluorescence (vert) du duplex cTAR/dTAR (bleu et rouge), ainsi que le spectre de transmittance du miroir dichroïque.

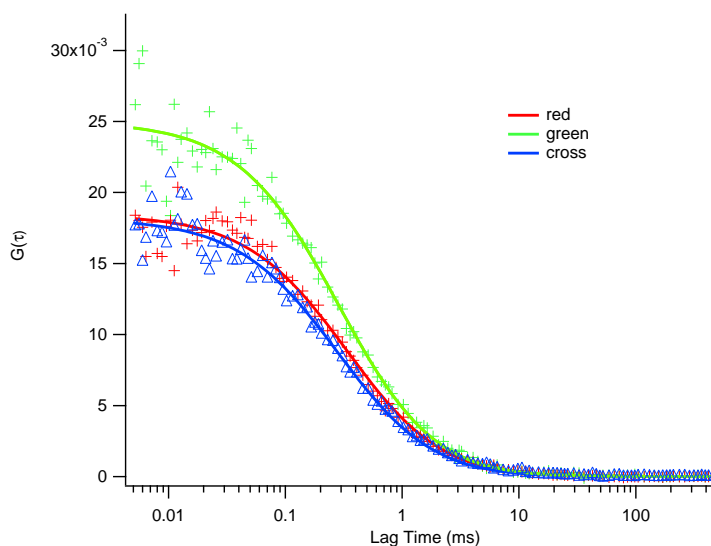


FIGURE 3.16 – Exemple de courbe de cross-corrélation d'un duplexe de 55bp (5'TexasRed cTAR :5'Oregon Green dTAR) (bleu), ainsi que les deux courbes d'autocorrélation des voies vertes (vert) et rouge (rouge)

3.5 Cross-Corrélation ou corrélation croisée

3.5.2 Corrélation croisée à deux volumes focaux

De même que pour la corrélation croisée à deux couleurs, il est possible de corrélérer des signaux provenant de deux volumes focaux distincts, décalés latéralement, partiellement chevauchants et générés par deux lasers. L'introduction d'une distance fixe entre les deux volumes focaux permet de mesurer des coefficients de diffusion absolus avec une haute précision [389]. Cette approche constitue la base de la dual-focus (ou two-focus) FCS. Dans les travaux rapportés dans l'article II, nous avons proposé une approche ne nécessitant qu'un seul laser, avec une distance ajustable entre les volumes focaux. Ceci est rendu possible par l'introduction d'un interféromètre de Michelson sur la voie d'excitation. Cette approche a été validée sur différents systèmes modèles. L'utilisation d'une excitation biphotonique permet d'envisager son utilisation en milieu cellulaire, voir tissulaire, pour caractériser avec précision la diffusion de molécules d'intérêt.

Ces travaux ont été publiés dans *Journal of fluorescence* en 2009.

3.6

Publication II

Two-Photon Two-Focus Fluorescence Correlation Spectroscopy with a Tunable Distance Between the Excitation Volumes

Pascal Didier · Julien Godet · Yves Mély

Received: 30 June 2008 / Accepted: 7 October 2008 / Published online: 16 October 2008
© Springer Science + Business Media, LLC 2008

Abstract In the present work, a Michelson interferometer was combined with a two-photon excitation microscope to perform two-focus Fluorescence Correlation Spectroscopy. This simple and original approach allows us to tune the distance between the two excitation volumes and determine absolute diffusion constants. The technique was validated on different model systems that demonstrate the sensitivity of the approach.

Keywords Two-photon excitation · Fluorescence Correlation Spectroscopy · Michelson interferometer

Introduction

Fluorescence Correlation Spectroscopy (FCS) has emerged as a powerful tool for biological applications [1]. This technique allows characterizing the translational dynamics of fluorescent molecules or complexes in any liquid environment. By analysing the intensity fluctuations of the fluorescent species within a femto-liter volume (defined by the laser excitation), several physical parameters—diffusion time, local concentration, molecular brightness, related to the hydrodynamic and photophysical properties of these species—can be monitored [2]. FCS has been successfully applied in model systems as well as in living cells [3]. However,

one concern of FCS is related to the accurate determination of the diffusion constant. Indeed, in classical FCS this parameter strongly depends on the shape of the observation volume and the experimental conditions (coverslide thickness, refractive index mismatch, optical saturation...) [4, 5]. To circumvent these limitations, several FCS based techniques were developed: standing wave FCS [6], total internal reflectance fluorescence-FCS [7], spatial correlation FCS with two laterally shifted pinholes [8], scanning FCS [9] and two-focus FCS (2fFCS) [10]. With this last method, fluorescence fluctuations from two overlapping excitation volumes are monitored and the resulting cross-correlation is used to determine absolute diffusion constants. The power of the method relies on the introduction of an external ruler, the distance between the overlapping beams generated using a Nomarski prism with two cross polarized laser beams [10].

In this context, we have developed a new 2fFCS approach using a two-photon excitation microscope coupled to a Michelson interferometer to generate two adjustable overlapping beams. This method, that requires only one laser, was validated on different model systems to illustrate its sensitivity.

Materials and methods

Sample preparation

Highly pure NCp7(11-55) peptides were prepared by chemical synthesis [11]. Peptide concentration was determined using an extinction coefficient of $5700 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nm (MW=5138 Da). DNA oligonucleotides (ODN) were synthesized at a $0.2 \mu\text{mol}$

P. Didier (✉) · J. Godet · Y. Mély
Photophysique des interactions Biomoléculaires,
Faculté de Pharmacie, UMR 7175, Université Louis Pasteur,
74 route du Rhin, 67401 Illkirch France
e-mail: pascal.didier@pharma.u-strasbg.fr

scale by IBA (Göttingen, Germany). Labelled oligonucleotides were modified at the 5' terminus with 5/6 TetraMethylRhodamine (TMR) or Rhodamine 6G (Rh6G) via an amino-linker with six-carbon spacer and HPLC purified. Titration curves were obtained by mixing a fixed amount of 5'TMR-AATGCC hexanucleotide with increasing amounts of NCp7(11-55) in 25 mM TRIS-HCl (pH 7.5), 30 mM NaCl, 0.2 mM MgCl₂ at 20°C. The mixtures (400 μL of final volume) were then equilibrated for several minutes in a Tween-coated eight-well Lab-Tek chamber system, before measurements. DNA duplexes were obtained by mixing 5'Rh6G (14-39)cTAR with a 5-fold excess of complementary sequences in 50 mM TRIS (pH 7.8), 75 mM KCl, 7 mM MgCl₂. The samples (10 μL) were heated for two minutes at 85°C and then cooled stepwise and finally diluted in the same buffer before measurement to reach a final concentration of 200 nM.

Fluorescence correlation spectroscopy and two-photon imaging

FCS measurements were performed on a home-built two-photon laser scanning system set-up based on an Olympus IX70 inverted microscope with an Olympus 60x 1.2NA water immersion objective [13]. Two-photon excitation was provided by a titanium-sapphire laser (Tsunami, Spectra Physics) and photons were detected with an Avalanche Photodiode (APD SPCM-AQR-14-FC, Perkin Elmer) connected to an on-line hardware correlator (ALV5000, ALV GmbH, Germany). To generate two overlapping excitation volumes in the sample, a Michelson interferometer was introduced in the optical path of the microscope as explained in "Results and discussion". Typical acquisition time was 10 min (20 × 30 s) with an excitation power around 2.5 mW at the sample level. Imaging was carried out using two fast galvo mirrors in the descanned fluorescence collection mode. Photons were detected with an APD connected to a counter/timer PCI board (PCI6602, National Instrument).

Data analysis

Data analysis was performed by assuming two identical laterally shifted observation volumes. In this case, the autocorrelation function can be expressed in the Green's functions formalism:

$$G(\tau, \delta) = Ac \int dr_1 \int dr_2 U(r_1) \frac{1}{(4\pi Dt)^{\frac{3}{2}}} \times \exp\left[-\frac{(r_2 - r_1 - \delta)^2}{4Dt}\right] U(r_2) \quad (1)$$

where $U(r)$ is the molecule detection function giving the spatial probability to observe a molecule at a position r in one detection volume, D is the diffusion coefficient, c is the concentration of the fluorescent species, δ is the distance between the two beam waists and A a factor that accounts for the detection efficiency of the whole set-up. To obtain an analytical solution for Eq. 1, $U(r)$ is assumed to be a 3D-Gaussian:

$$U(r) = U_0 \exp\left[-\frac{2x^2}{\omega_0^2} - \frac{2y^2}{\omega_0^2} - \frac{2z^2}{\omega_z^2}\right] \quad (2)$$

In this case, Eq. 1 can be written as:

$$G(\tau, \delta) = \frac{1}{N} \left[\frac{1}{1 + \frac{4D\tau}{\omega_0^2}} \right]^{-1} \left[\frac{1}{1 + \frac{4D\tau}{\omega_z^2}} \right]^{-\frac{1}{2}} \times \exp\left[-\frac{\delta^2}{4Dt + 2\omega_0^2 + \omega_z^2}\right] \quad (3)$$

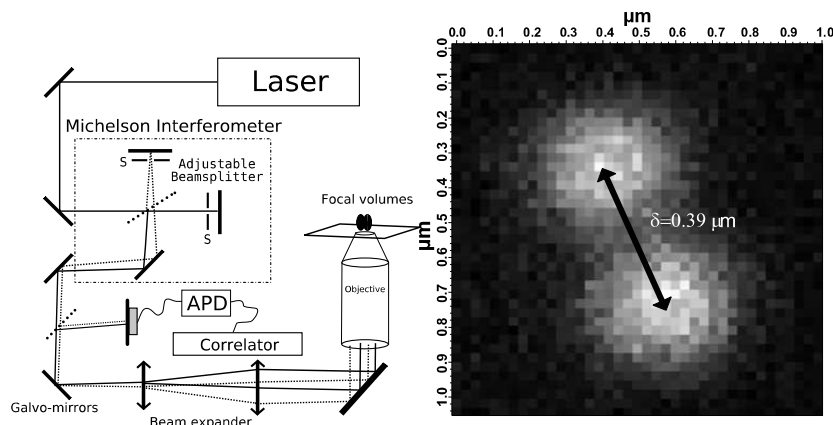
If $\delta = 0$, Eq. 3 reduces to the classical expression for a 3D diffusion model.

Results and discussion

Two overlapping excitation volumes were generated from a single laser source by adding a Michelson interferometer in the optical path of the microscope. When the two arms of the interferometer are spatially and temporally aligned, an interference pattern is generated. On the contrary, if the pulses in the two arms of the interferometer are temporally shifted, they do not interfere. Moreover, it is possible to produce two non-collinear beams by tilting the beamsplitter. The introduction of an angle between the beams leads to a lateral shift in the focal plane of the microscope (Fig. 1a). By changing the tilt angle, the distance between the two foci can thus be easily varied. To determine the absolute distance between the two overlapping beams, shutters (S in Fig. 1a) were placed on the two arms of the interferometer. Two images were obtained by closing alternatively the shutters. The two images were fitted with a 2D gaussian function and the distance between the excitation volumes was deduced from the position of the peaks (Fig. 1b). The point spread function (PSF) of the set-up was determined from a z-scan on one fluorescent bead. The measured lateral (ω_0) and axial (ω_z) resolutions were respectively 0.34 μm and 1.1 μm (Eq. 2).

A 2fFCS measurement was first performed with a 50 nM aqueous solution of TMR. The obtained correlation curve is reported in Fig. 2a. Using an inter-foci distance of 390 nm, a diffusion constant of 421 ±

Fig. 1 **a** Scheme of the experimental set-up. **b** Image of one fluorescent bead obtained in the descanned mode by illuminating the sample with two slightly tilted beams. In this experiment, the measured distance was $0.39 \mu\text{m}$



$7 \mu\text{m}^2\text{s}^{-1}$ was determined, in excellent agreement with the value obtained by scanning FCS for Rh6G [9]. The inter-foci distance can be varied continuously. However, due to the exponential damping factor in Eq. 3 the distance has to be lower than 800 nm to get reliable fit (data not shown). To further highlight the accuracy of the method, 2fFCS measurements were repeated with TMR solutions of increasing viscosity, obtained by adding different amounts of sucrose [12]. As expected from the Stokes-Einstein equation, the diffusion coefficient scales linearly with the inverse of the viscosity (Fig. 2b).

2fFCS was also applied to monitor the binding of the NCp7(11-55) peptide to a 5'-TMR-AATGCC hexanucleotide. In classical FCS, a significant change in the diffusion constant can be evidenced only if the mass ratio between bound and free forms is at least five, due

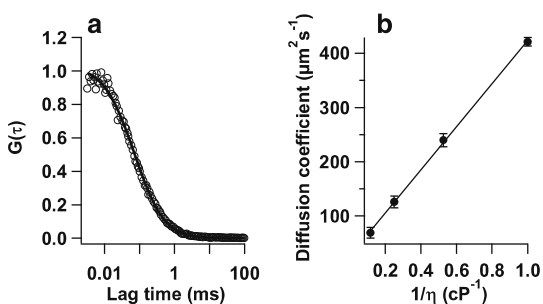


Fig. 2 **a** Correlation curve obtained with a 50 nM aqueous solution of TMR (circles). The fit in solid line provides a diffusion constant of $421 \pm 7 \mu\text{m}^2\text{s}^{-1}$ for $\delta = 0.39 \mu\text{m}$. **b** Variation of the diffusion constant of TMR as a function of the inverse viscosity (calculated from the sucrose concentration [12]). Typical acquisition time per point was 10 min

to the cubic root dependence of the diffusion coefficient with the molecular mass and the uncertainty on the excitation volume. In Fig. 3a, the correlation curves for free 5'-TMR-AATGCC and 5'-TMR-AATGCC bound to NCp7(11-55) were reported. The corresponding diffusion constants were respectively $216 \mu\text{m}^2\text{s}^{-1}$ and

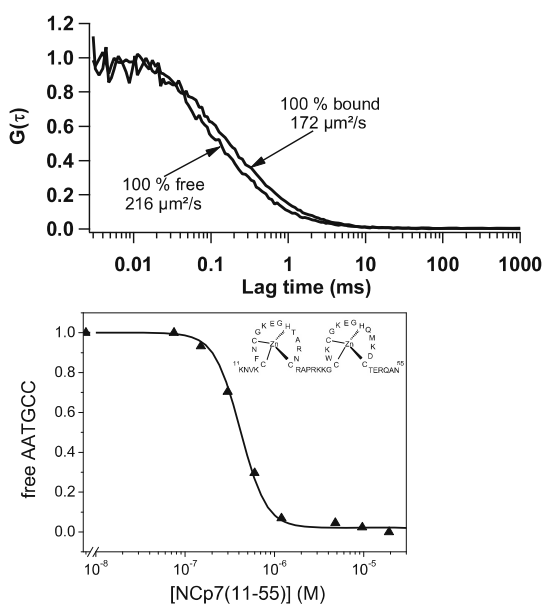


Fig. 3 **a** Correlation curves obtained for 5'-TMR-AATGCC in the absence or in the presence of a saturating concentration ($20 \mu\text{M}$) of NCp7(11-55). **b** Variation of the population of free oligonucleotide as a function of the NCp7(11-55) concentration. From the titration curve, an affinity constant $K = (2.4 \pm 0.8) \times 10^6 \text{ M}^{-1}$ was determined. Typical acquisition time per point was 10 min

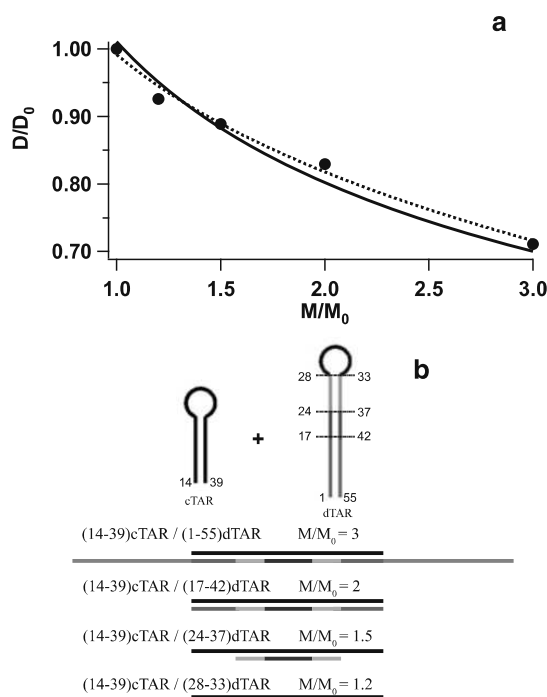


Fig. 4 **a** Variation of the diffusion constant (black circles) as a function of the molecular mass of the duplexes. The diffusion constant D and the molecular mass M of the duplexes were expressed as a ratio to respectively, the diffusion constant D_0 and the molecular mass M_0 of the 5′Rh6G(14-39)cTAR sequence. The solid and dotted lines correspond respectively to the fits obtained with the spherical and the rod-like models, respectively. Typical acquisition time per point was 10 min. **b** Schematic structure of the duplexes and their corresponding mass ratio

$172 \mu\text{m}^2\text{s}^{-1}$, showing that the method can discriminate the two forms, differing by a factor 3 in their molecular mass. The brightness of the free and bound forms were identical indicating that TMR emission was not affected by NCp7(11-55). Next, the correlation curves obtained with different amounts of NCp7(11-55) were analyzed with a two population model and fixing the diffusion constants of the free and bound species. From the variation of the population of free species as a function of the NCp7(11-55) concentration (Fig. 3b), we found an affinity constant $K = (2.4 \pm 0.8) \times 10^6 \text{M}^{-1}$, in excellent agreement with that measured from fluorescence anisotropy titrations (data not shown).

Finally, the technique was further validated by measuring the diffusion constant of 5′Rh6G(14-39)cTAR thermally annealed to complementary oligonucleotides of different molecular weights. In Fig. 4a, we reported the dependence of the diffusion constant of

the different duplexes as a function of their molecular weight, both parameters being expressed relatively to the corresponding parameters of free 5′Rh6G(14-39)cTAR. The solid and dotted lines correspond to fits obtained assuming either spherical ($M_w \propto R^3$) or rod-like ($M_w \propto L$) shape for the complexes, respectively [13, 14]. A reasonable fit was obtained with both models, but the rod-like model was clearly the best one for duplexes with $\frac{M}{M_0} > 1.5$. This rod-like behavior for the two largest duplexes is fully consistent with the predicted dimension of their 25 bp double-stranded central part (length=90 Å and width=24 Å). Thus, the introduction of the exact distance between the two foci allows us to evidence changes as small as 20 % in the molecular mass.

In the present work an original approach, using a two-photon excitation microscope and a Michelson interferometer, was developed to perform 2fFCS. The technique can be easily implemented on any two-photon excitation microscope and allows measuring absolute diffusion constants overcoming thus one of the major limitation of classical FCS.

References

- Haustein E, Schwille P (2007) Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu Rev Biophys Biomol Struct* 36:151–169
- Wachsmuth M, Waldeck W, Langowski J (2000) Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J Mol Biol* 298:677–689
- Garcia-Saez AJ, Schwille P (2007) Single molecule techniques for the study of membrane proteins. *Appl Microbiol Biotech* 76:257–266
- Enderlein J, Gregor I, Patra D, Dertinger T, Kaupp B (2005) Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. *ChemPhysChem* 11:2324–2336
- Berland K, Shen G (2003) Excitation saturation in two-photon fluorescence correlation spectroscopy. *Appl Opt* 42:5566–5576
- Davis SK, Bardeen CJ (2002) Using two-photon standing waves and patterned photobleaching to measure diffusion from nanometers to microns in biological systems. *Rev Sci Instr* 73:2128–2135
- Hansen RL, Harris JM (1998) Total internal reflection fluorescence correlation spectroscopy for counting molecules at solid/liquid interfaces. *Anal Chem* 70:4247–4256
- Jaffiol R, Blancquaert Y, Delon A, Derouard J (2006) Spatial fluorescence cross-correlation spectroscopy. *Appl Opt* 45:1225–1235
- Petrásek P, Schwille P (2008) Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys J* 94:1437–1448
- Dertinger T, Pacheco V, von der Hocht I, Hartmann R, Gregor I, Enderlein J (2007) Two-focus fluorescence correlation

- spectroscopy: a new tool for accurate absolute diffusion measurements. *ChemPhysChem* 8:433–443
11. de Rocquigny H, Ficheux D, Gabus C, Fournié-Zaluski MC, Darlix JL, Roques BP (1991) First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem Biophys Res Commun* 180: 1010–1018
 12. The Chemical Rubber Company (2008) *Handbook of biochemistry*, 2nd edn. J290
 13. Clamme JP, Azoulay J, Mély Y (2003) Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys J* 84:1960–1968
 14. Beltz H, Piémont E, Schaub E, Ficheux D, Roques B, Darlix JL, Mély Y (2004) Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol* 338:711–723

Cinétique et mécanisme de l'hybridation de
cTAR/TAR au cours du premier saut de brin

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR au cours du premier saut de brin

4.1.1 Contexte

Les premières études *in vitro* de transfert du brin (-) étaient basées sur l'hybridation en l'absence de NCp7 de séquences virales relativement peu structurées [390, 94]. De ce fait, la dégradation de la matrice ARN par l'activité RNase H de la Transcriptase Reverse (RT) ainsi que le degré d'homologie entre la matrice ARN et le (-)ssDNA apparaissaient comme les deux paramètres essentiels modulant l'efficacité du transfert de brin. Ultérieurement, l'utilisation de systèmes reconstitués, permettant la formation de complexes nucléoprotéiques *in vitro*, a montré que la NCp7 stimule le transfert de brin en accélérant l'hybridation des régions R [368, 391, 350, 392]. Comme ces régions comportent des structures hautement conservées en tige-boucle, cette réaction de transfert dépend de la capacité de la NCp7 à déstabiliser ces structures secondaires ce qui constitue l'étape limitante de la réaction d'hybridation [278]. En fait, si la taille des régions complémentaires joue un rôle dans l'efficacité du transfert de brin, ce n'est pas la composante essentielle. La relation entre d'une part la structure et la stabilité des acides nucléiques mise en jeu et d'autre part l'activité chaperonne de la NCp7 semble plus importante encore. En effet, alors que la NCp7 exerce son effet déstabilisant de manière plus efficace sur les séquences ADN que ARN [393], l'efficacité de transfert est également sensible aux structures ARN puisqu'un transfert efficace ne se produit que si les structures ARN, en partie déstabilisées, conservent une structure suffisante. Un équilibre thermodynamique entre les partenaires ADN et ARN doit être maintenu pour permettre un transfert efficace [394]. Plus que la stabilité thermodynamique globale des séquences, c'est donc la stabilité de structures locales qui constitue le principal déterminant de l'activité chaperonne de la NCp7 [268, 269]. Les propriétés de déstabilisation de NCp7 sont attribuées aux deux doigts de zinc, à la séquence entre ces deux doigts et à la plateforme hydrophobe constituée par les acides aminés *Val*₁₃, *Phe*₁₆, *Thr*₂₄, *Ala*₂₄ et *Trp*₃₇. Les parties *N*_{ter} et *C*_{ter} contribuent quant à elles seulement à la liaison de NCp7 sur les séquences oligonucléotidiques [226].

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

Dans ce contexte, et afin de mieux comprendre le rôle de NCp7 dans le premier saut de brin, l'objectif de ce travail a été de caractériser les paramètres cinétiques de l'hybridation de cTAR avec TAR ARN et de mettre en évidence les principales relations structure-activité. Nous avons également cherché à déterminer le mécanisme de cette réaction d'hybridation. Pour atteindre ces objectifs, nous avons utilisé essentiellement des techniques de spectroscopie de fluorescence à l'aide d'oligonucléotides synthétiques doublement marqués à leurs extrémités par des sondes fluorescentes. Pour mettre en évidence les déterminants nucléotidiques et protéiques des propriétés chaperonnes de la NCp7, nous avons utilisé différents mutants de cTAR et de la protéine.

4.1.2 Résultats

Rôle de NCp7 dans l'hybridation cTAR/TAR

Lors du premier saut de brin de la transcription inverse, le (-)ssDNA néoformé est transféré de l'extrémité 5' à l'extrémité 3' du brin ARN(+). Cette étape met en jeu l'hybridation de TAR et de cTAR.

Cette hybridation est suivie en utilisant un oligonucléotide cTAR doublement marqué par un couple (Rh6G-DABCYL) ou (TMR-FAM). A 20°C, cTAR forme une tige-boucle. Les deux chromophores flanquants les extrémités de la tige sont alors proches l'un de l'autre induisant un couplage de type excitonique entre les sondes ce qui entraîne une inhibition de la fluorescence [395]. Une augmentation de la distance entre les chromophores se traduit par la restauration de l'émission de fluorescence de la Rh6G ou de la FAM. L'hybridation de cTAR et de son complémentaire peut donc être suivie par la restauration de l'émission de fluorescence en fonction du temps [396].

La réaction d'hybridation est bimoléculaire et NCp7 active cette réaction

En l'absence de protéine, cTAR doublement marqué par le couple (TMR-FAM) présente deux pics de fluorescence correspondant à chacun des deux chromophores. L'ajout de protéine provoque une augmentation de l'intensité de fluorescence (insert de la Figure 2 de [280]) due à la déstabilisation de la structure secondaire qui entraîne une ouverture partielle

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

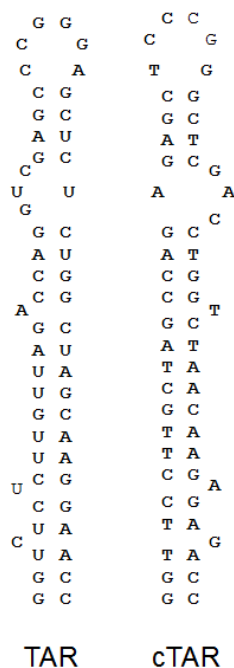


FIGURE 4.1 – Structures secondaires de TAR ARN et cTAR ADN de la souche Mal. La structure secondaire de cTAR est prédite à partir du programme [mfold](#)

de la tige.

La réaction d'hybridation de cTAR avec la séquence complémentaire ARN ou son équivalent ADN (dTAR) a d'abord été réalisée en solution en absence ou en présence de NCp7. L'hybridation de cTAR avec dTAR se traduit par une augmentation de l'intensité de fluorescence en fonction du temps puis par l'atteinte d'un plateau correspondant à la fluorescence du duplex formé (Figure 2a de [280]). La fluorescence du duplex est 7 fois plus intense que celle de la tige-boucle cTAR. Le traitement des données expérimentales des courbes de cinétiques nécessite un modèle biexponentiel pour décrire correctement les données expérimentales (Figure 2b de [280]).

En utilisant des concentrations en séquence complémentaire au moins dix fois supérieure à l'espèce marquée, il nous est possible de travailler dans des conditions de pseudo premier ordre. En analysant les vitesses à des concentrations croissantes d'espèce complémentaire, il apparaît que la vitesse d'hybridation augmente avec la concentration de dTAR (Figure 3 b) ou de TAR ARN (données non présentées). Il s'agit donc bien d'une réaction bimoléculaire.

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

laire et on peut caractériser les vitesses de formation k_e^{on} et de dissociation k_e^{off} du duplex cTAR/TAR selon l'équation :

$$k_e^{obs} = k_e^{on} \times [TAR] + k_e^{off} \quad (4.1)$$

Nos conclusions sur la bimolécularité de l'hybridation sont en désaccord avec des résultats antérieurs [278] qui ont montré que l'hybridation est monomoléculaire du fait de l'étape limitante de déstabilisation des différentes structures secondaires par NCp7. Par contre, nos résultats sont en accord avec des études d'hybridation en présence de NCp7 d'autres acides nucléiques comme PBS [312] et sont cohérents avec les modèles proposés pour de nombreuses autres protéines chaperonnes comme hnRNP A1 [397]. Les désaccords avec les données de You et McHenry [278] peuvent être attribués en partie au fait que ces derniers ont analysé l'hybridation par électrophorèse sur gel, une méthode qui ne permet pas de suivre la réaction en temps réel et où les phénomènes d'agrégation (qui peuvent influencer l'ordre des réactions d'hybridation) sont difficiles à mettre en évidence.

En l'absence de NCp7, la réaction d'hybridation de cTAR avec dTAR est très lente ($\sim 40 M^{-1}.s^{-1}$) et la constante de vitesse réverse (k_e^{off}) est très proche de 0. L'ajout de protéine NCp7(12-53) induit un gain de 4 ordres de grandeur de la vitesse de réaction. Cette augmentation dépend de la concentration en protéine et plus précisément du ratio nucléotide/protéine. Une vitesse optimale est obtenue dans des conditions saturantes de protéine où la protéine recouvre l'ensemble de la séquence des acides nucléiques ($r=5$). L'augmentation de la vitesse est encore plus importante lorsque l'on substitue NCp7(12-53) par NCp7(1-55) où la partie N_{term} de la protéine est présente. La NCp7(1-55) augmente la vitesse de réaction d'au moins 2 ordres de grandeur par rapport à NCp7(12-53) soit au moins 6 ordres de grandeur par rapport à l'hybridation en absence de protéine. Contrairement à la composante de déstabilisation de l'activité chaperonne, la composante d'hybridation apparaît donc fortement dépendante de la partie N_{term} . En règle générale, l'étape limitante d'une hybridation est constituée par la nucléation qui correspond à l'appariement des n premières paires de bases entre les séquences complémentaires. La nucléation est un phénomène limité par la diffusion des séquences complémentaires. Les forces de répulsions

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

électrostatiques entre les brins complémentaires ainsi que la faible probabilité d'interaction entre des brins correctement positionnés limitent la vitesse de nucléation. Par ses charges positives, la NCp7(1-55) neutralise les charges négatives des acides nucléiques, et permet leur rapprochement, facilitant ainsi leur hybridation.

L'existence de deux composantes exponentielles pour l'hybridation de cTAR avec TAR peut s'expliquer par deux populations A1 et A2 de cTAR qui seraient capables de réagir avec l'espèce complémentaire avec des vitesses différentes. Aucune interconversion n'est envisagée entre ces deux populations de cTAR. Cette hypothèse est cependant à l'encontre des mesures de fluorescence résolue en temps qui montrent qu'en solution, cTAR existe sous différentes conformations (caractérisées par des temps de vie de fluorescence différents) en équilibre [393]. Ces différentes conformations correspondent à des formes de cTAR où la tige est partiellement ouverte. Ces conformations sont en équilibre avec des espèces noires, non fluorescentes, correspondant à une structure fermée pour laquelle les deux sondes, très proches l'une de l'autre forment un hétérodimère [396]. L'addition de NCp7 à cTAR induit un déplacement de l'équilibre conformationnel de cTAR vers les formes ouvertes [393]. Pour tenir compte de cet équilibre entre les différentes conformations de cTAR et en se fondant sur l'hypothèse raisonnable que seules les conformations partiellement ouvertes de cTAR sont réactives, nous avons proposé un modèle bicompartimental ouvert pour traiter les données de l'hybridation. Dans ce modèle, la conformation fermée non réactive de cTAR est en équilibre avec les conformations où la tige est partiellement fondue (A2). Les espèces A2 réagissent avec la séquence complémentaire (B) pour donner un duplex étendu (Figure 6 de [280]). L'intégration du système d'équations différentielles correspondantes donne, en fonction des conditions initiales, l'équation 14 de [280] qui a la forme d'une équation biexponentielle.

Rôle de la stabilité de la tige de cTAR dans l'hybridation

L'hybridation de cTAR avec TAR est significativement plus lente (10 fois) que l'hybridation des deux partenaires ADN (Table 1). Cette diminution de la vitesse d'hybridation pourrait s'expliquer par la plus grande stabilité de la tige-boucle ARN de TAR par rapport à la tige-boucle ADN de dTAR. La stabilité des structures en tige-boucle serait donc une

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

composante limitante de la réaction.

Pour évaluer l'incidence de la stabilité des oligonucléotides dans la réaction d'hybridation, nous avons utilisé trois mutants : cTAR1, cTAR2 et cTAR1,2¹ (Figure 1) dans lesquels les *bulges* de la partie basse de cTAR ont été remplacés par une paire de base (bp). Ces mutants présentent une stabilité plus grande que cTAR et la déstabilisation de ces composés par la NCp7 est limitée [268].

La stabilisation de la tige de cTAR diminue la vitesse de réaction, voire inhibe la réaction dans le cas du cTAR1,2. Concernant cTAR1 et cTAR2, la relation entre la stabilité et les vitesses d'hybridation induites par NCp7 n'est pas linéaire. La stabilité de cTAR et de ses mutants a été obtenue à partir des courbes de fusion réalisées par spectrophotométrie UV en suivant les absorbances des composés à 260nm ou en suivant le pic d'absorbance du TMR [268]. Bien que cTAR1 et cTAR2 présentent la même stabilité, la vitesse d'hybridation est deux fois plus élevée pour cTAR1 que pour cTAR2. Par ailleurs, la NCp7 est incapable d'hybrider cTAR1,2 dont la structure secondaire est beaucoup plus stable. Ceci confirme que la stimulation de l'hybridation par NCp7 est limitée par son aptitude à fondre que des séquences double brin de taille limitée [269]. En effet, cTAR1 conserve une séquence double brin terminale de 3 bp alors que dans cTAR2, elle passe à 6 bp. Ces expériences mettent en évidence le rôle primordial du bulge G52 dans la déstabilisation de cTAR par NCp7. Ce dernier bulge étant fortement conservé, ceci renforce l'hypothèse d'une co-évolution entre TAR et l'activité de NCp7 [268].

Rôle de la boucle apicale

Il a été montré que l'efficacité de l'hybridation repose essentiellement sur la complémentarité de séquences entre les boucles apicales de cTAR et TAR et que la boucle apicale de TAR serait le site d'initiation de l'hybridation [398]. Afin de confirmer ce point, nous avons déterminé les cinétiques d'hybridation avec des mutants de la boucle apicale : ΔL dTAR et (14-39) ΔL cTAR.

Les hybridations de (14-39) ΔL cTAR (un mutant de la partie supérieure de cTAR dans lequel la boucle apicale a été remplacée par une chaîne hexaéthylène glycol) avec dTAR

1. Les données de cTAR2 et de cTAR1,2 ne figurent pas dans la publication [280]

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

ou le (17-42)dTAR n'ont pas lieu, montrant que la boucle est primordiale pour ces réactions. *A contrario*, l'hybridation de ΔL cTAR avec dTAR se déroule avec une vitesse similaire à la vitesse correspondant aux séquences natives. Il apparaît donc que pour les séquences longues, la partie basse de la tige initie la réaction d'hybridation. Ceci a été confirmé par l'analyse de mutants ciblant plus particulièrement cette zone.

Ainsi, l'hybridation de cTAR/TAR nécessite que la partie basse de l'un des deux partenaires soit ouverte par NCp7, ce qui permet l'invasion de la tige de la séquence complémentaire et la nucléation du duplex étendu. Par la suite, la réaction se propagerait à l'ensemble de la séquence selon un processus de zipping qui permettrait de dépasser la barrière énergétique essentiellement instituée par les régions doubles brins stables de la partie supérieure de la tige.

Rôle des doigts de zinc de NCp7²

Pour évaluer l'importance des doigts de zinc de NCp7 dans l'hybridation, nous avons comparé la protéine NCp7 ((1-55)NCp7) avec un mutant (SSHS)₂NCp7. Dans ce mutant, les différentes cystéines permettant de complexer le zinc ont été mutées par des sérines, ce qui empêche la formation des doigts de zinc. Ce mutant n'est pas capable de déstabiliser cTAR [226]. La réaction d'hybridation de cTAR/dTAR en présence de NCp7 à une concentration saturante est très rapide ($> 10^6 M^{-1}.s^{-1}$). Par conséquent, les peptides sont ajoutés à un ratio nucléotide/protéine de 10 avec un ratio cTAR/dTAR de 1 : 2 pour permettre de suivre la réaction et comparer l'effet des peptides. Dans ces conditions, la réaction d'hybridation entre cTAR et TAR en présence de NCp7 reste rapide ($\sim 7 \times 10^5 M^{-1}.s^{-1}$). (SSHS)₂NCp7 est également capable de promouvoir l'hybridation de cTAR et de TAR avec une vitesse comparable ($9 \times 10^5 M^{-1}.s^{-1}$). D'autre part, une étude de la dépendance des vitesses d'hybridation en fonction de la température a été effectuée (tableau 4.1). Celle-ci a permis de montrer que NCp7 diminue légèrement l'énergie d'activation, en accord avec sa capacité à déstabiliser la tige de cTAR. L'entropie d'activation associée au mutant (SSHS)₂NCp7 est presque deux fois supérieure à celle de NCp7. Ainsi, la capacité d'agrégation de (SSHS)₂NCp7, supérieure à celle de NCp7, permet de favoriser la nucléation des séquences cTAR et dTAR. Le

2. Non présenté dans le manuscrit [280]

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

fait que les vitesses d'hybridation soient relativement similaires indique que la stimulation de la nucléation induite par (SSHS)₂NCp7 compense largement la perte de l'activité de déstabilisation. Cette hypothèse a déjà été évoquée dans la littérature [312].

| | k $M^{-1}.s^{-1}$ | ΔG^* kJ/mol | Ea | ΔH^* kJ/mol | ΔS^* $K.cal/mol$ |
|--------------------------|---------------------------|--------------------------|------------|--------------------------|-----------------------------|
| (1-55)NCp7 | $6.6 \pm 0.6 \times 10^5$ | 44 | 58 ± 4 | 56 ± 3 | 9 ± 1 |
| (SSHS) ₂ NCp7 | $9.0 \pm 0.7 \times 10^5$ | 44 | 66 ± 4 | 64 ± 3 | 16 ± 1 |

TABLE 4.1 – Influence de la température sur l'hybridation de cTAR / dTAR en présence de NCp7 ou de (SSHS)₂NCp7 à un ratio protéine/nt de 1 : 10

Rôle du magnésium dans l'hybridation de cTAR/TAR ³

La concentration en magnésium module fortement les vitesses et le mécanisme de l'hybridation de cTAR avec TAR en stabilisant les tige-boucles [393]. Ceci peut être mis en évidence facilement par spectroscopie d'absorption d'un cTAR doublement marqué en exploitant les variations d'absorption liées au couplage excitonique entre les sondes [395, 396]. Le Mg^{2+} induit également une diminution de l'activité de déstabilisation de cTAR par NCp7 (figure 4.2). Ceci résulte de l'action combinée du Mg^{2+} qui (i) stabilise la tige de cTAR, favorisant donc les formes fermées [267], et qui (ii) déplace la NCp7 de la séquence cTAR en diminuant son affinité. La liaison de la NCp7 est en effet largement dépendante de la force ionique [246].

En accord, avec le mécanisme d'hybridation que nous avons décrit et qui repose sur la déstabilisation de la séquence cTAR, la vitesse d'hybridation est donc largement diminuée en présence de Mg^{2+} (figure 4.3a).

D'autre part, nous avons pu mettre en évidence que l'agrégation des acides nucléiques par NCp7 est elle aussi largement dépendante de la concentration en Mg^{2+} . En effet, pour des concentrations de Mg^{2+} de plus de 5 mM, celle-ci n'est plus observée. Enfin, l'agrégation est totalement réversible lorsque l'on augmente la concentration de Mg^{2+} (figure 4.3

3. Présenté dans [332]

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

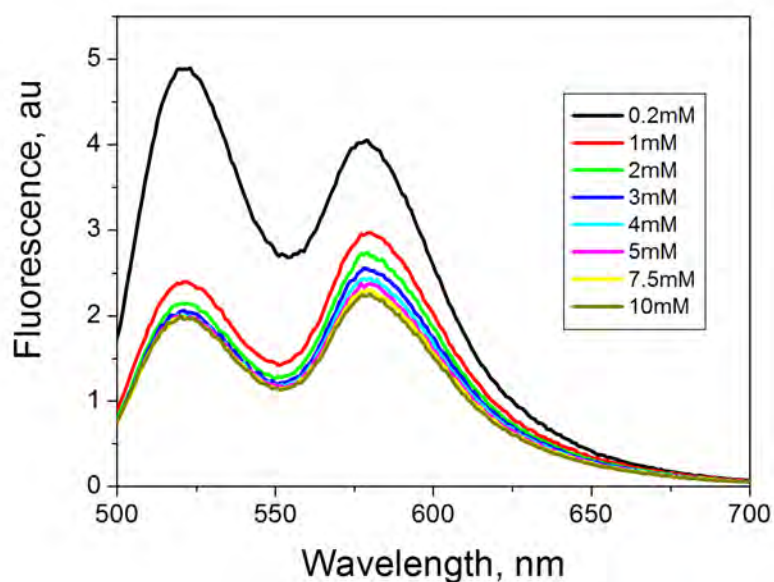


FIGURE 4.2 – Effet de la concentration en magnésium sur la destabilisation de cTAR marqué par le couple (TMR,F1) induite par (11-55)NCp7.

La destabilisation du 5'TMR-cTAR-3'F1 induite par NCp7 se traduit principalement par une diminution des espèces "noires" [395, 393], résultant en une augmentation simultanée des deux pics de fluorescence. Le magnésium inhibe très largement la destabilisation induite par NCp7 dès 2 mM.

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

et 4.3c).

Ces travaux ont été publiés dans *Journal of Molecular Biology* en 2006 et dans *Biochemistry* en 2008.

4.1 Cinétique et mécanisme d'hybridation de cTAR/TAR

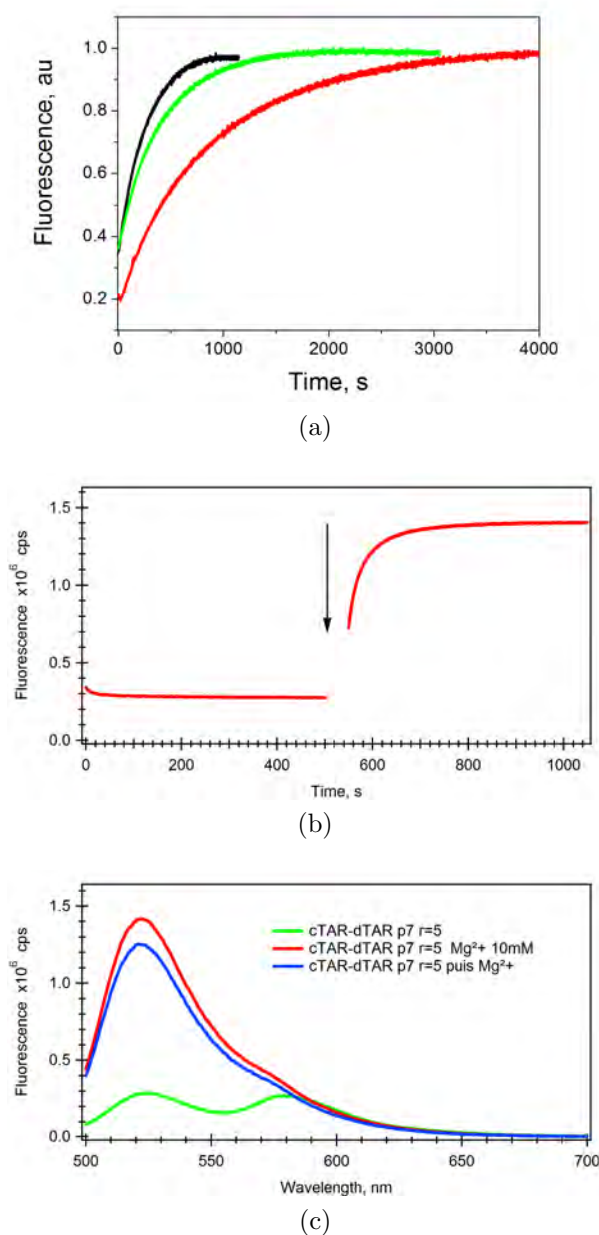


FIGURE 4.3 –

(a) Évolution de la fluorescence de la FAM au cours du temps lors de l'hybridation de 5'-TMR-cTAR-3'-FAM avec dTAR en présence de (11-55)NCp7 et 0.2 mM (noir), 5 mM (vert) et 10 mM (rouge) de $MgCl_2$

(b) Evolution de la fluorescence de la FAM au cours du temps lors de l'hybridation de 5'-TMR-cTAR-3'-FAM avec dTAR en présence de (1-55)NCp7 à 0.2 mM $MgCl_2$, l'ajout de QSP 10 mM de $MgCl_2$ (flèche), induit la restauration de la fluorescence de la FAM

(c) Spectre du duplex cTAR/dTAR résultant de l'hybridation en présence de 0.2 mM (vert) ou 10 mM (rouge) de $MgCl_2$. L'ajout de 10 mM de $MgCl_2$ en fin de réaction permet de recouvrer une fluorescence de la FAM identique à celle obtenue pour l'hybridation à 10 mM de $MgCl_2$

4.2

Publication III

During the Early Phase of HIV-1 DNA Synthesis, Nucleocapsid Protein Directs Hybridization of the TAR Complementary Sequences *via* the Ends of their Double-stranded Stem

Julien Godet¹, Hugues de Rocquigny¹, Chinnappan Raja¹,
Nicole Glasser¹, Damien Ficheux², Jean-Luc Darlix³ and Yves Mély^{1*}

¹*Photophysique des interactions moléculaires, UMR 7175 CNRS Institut Gilbert Laustriat Faculté de Pharmacie, Université Louis Pasteur Strasbourg 1, 74, Route du Rhin 67401 Illkirch Cedex, France*

²*IBCP, 7, passage du Vercors 69367 Lyon Cedex 07, France*

³*LaboRétro, Unité de Virologie Humaine INSERM #412 Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon, France*

Reverse transcription of HIV-1 genomic RNA requires two obligatory strand transfers. During the first strand transfer reaction, the minus strand strong-stop DNA (ss-cDNA) is transferred by hybridization of complementary sequences located at the 3' ends of the ss-cDNA and genomic template, respectively. In HIV-1, the major components of ss-cDNA transfer are the terminally redundant structured TAR elements and the nucleocapsid protein NCp7, which actively chaperones the hybridization of cTAR DNA to TAR. In the present study, we investigated the annealing kinetics of TAR with fluorescently labelled cTAR derivatives both in the absence and in the presence of NC(12-55), a peptide that contains the finger and C-terminal domains of NCp7. The annealing of TAR with cTAR involves two second-order kinetic components that are activated by at least two orders of magnitude by NC(12-55). The NC-promoted activation of cTAR-TAR annealing was correlated with its ability to destabilize the lower half of TAR stem, in order to generate the single-stranded complementary regions for nucleating the duplex structures. The two kinetics components have been assigned to two different pathways. The rapid one does not lead to extended duplex formation but is associated with a limited annealing of the terminal bases of cTAR to TAR. On the other hand, extended duplex formation follows a slower pathway that is limited kinetically by the nucleation of residues located mainly within the central double-stranded segment of both cTAR and TAR stems. An alternative mechanism involving an interaction through TAR and cTAR loops has been observed but is a minor pathway in the present conditions.

© 2006 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: fluorescence; HIV NCp7; TAR; hybridization; kinetics

Introduction

In retroviruses, DNA synthesis starts by reverse transcriptase (RT)-directed extension of a cellular

Abbreviations used: HIV-1, human immunodeficiency virus type 1; NC, nucleocapsid protein; RT, reverse transcriptase; TAR, transactivation response element; cTAR, complementary DNA sequence to TAR; dTAR, DNA analogue to TAR; SL, stem-loop; FRET, fluorescence resonance energy transfer; ssDNA, strong-stop DNA; Fl, 5 (and-6)carboxyfluorescein; TMR, 5 (and-6)carboxytetramethylrhodamine; PBS, primer binding site.

E-mail address of the corresponding author: mely@pharma.u-strasbg.fr

primer tRNA annealed to a small sequence, the primer-binding site (PBS), located close to the 5' end of the genomic RNA template.¹ Synthesis of the complementary DNA (cDNA) stops rapidly because RT reaches the template 5' end, generating the so-called minus-strand strong-stop DNA (ss-cDNA). cDNA synthesis can resume only if the ss-cDNA is transferred to the template 3' end. This cDNA transfer process is essential to generate the complete double-stranded proviral DNA flanked by the long terminal repeat (LTR) and relies on the genetic structure of the retroviral genome that contains terminally redundant R sequences. In human immunodeficiency virus type 1 (HIV-1), the major components of ss-cDNA transfer onto the

genomic RNA were found to be the terminally redundant structured TAR elements and the nucleocapsid protein (NC) NCp7.²⁻¹⁵

HIV-1 NC in its mature form is generated by sequential cleavage of the structural Gag polyprotein by the viral protease.¹⁶⁻¹⁸ NC is a small basic protein (Figure 1(a)) containing two highly conserved CCHC zinc fingers that coordinate zinc ions with high affinity.^{19,20} The two zinc fingers are connected by a short flexible linker and flanked by poorly folded N and C-terminal basic domains.^{4,21-24} NC is a nucleic acid-binding and chaperoning protein that plays critical roles in HIV-1 virion assembly and structure, and replication. By chaperoning the rearrangement of nucleic acid molecules into their most stable conformation and the rapid hybridization of complementary

sequences in physiological conditions, NC is viewed as the major viral factor required for primer tRNA annealing to the PBS and the two obligatory strand transfers during viral DNA synthesis.^{25,26} In the minus strand DNA transfer reaction, the ss-cDNA is translocated to the 3' end of the genomic RNA by means of the hybridization of the repeat R sequences present at the 3' ends of the ss-cDNA and genomic RNA, respectively. *In vitro*, this transfer reaction is slow and inefficient in the absence of NC, mainly due to stable stem-loop (SL) structures present in R, notably in TAR (Figure 1(b)) at the very 3' end of the genomic RNA that does not hybridize readily with the complementary cTAR of the ss-cDNA.^{6,8,11,14,27,28} NC chaperones ss-cDNA transfer by increasing the rate and extent of the

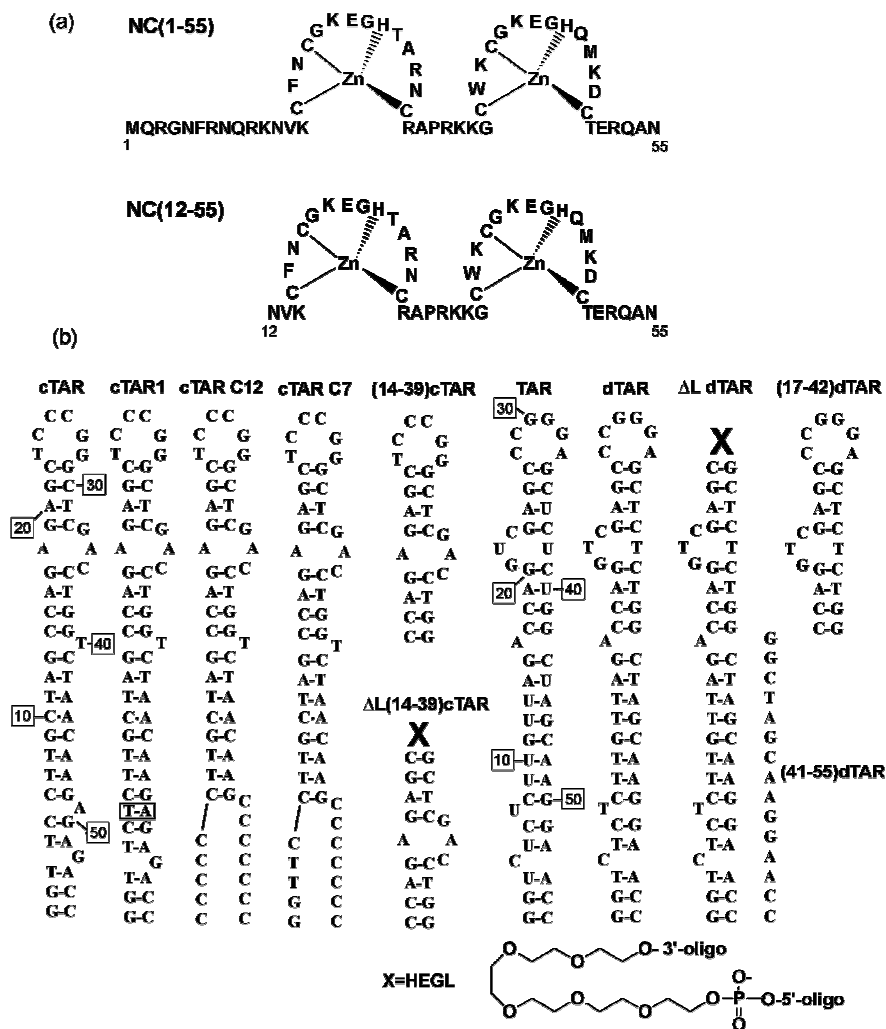


Figure 1. Structure of the (a) peptides and (b) oligonucleotides used in this study. The TAR RNA and cTAR DNA sequences are from the HIV-1 MAL strain. The secondary structures of the oligonucleotides were predicted from the structure of TAR²⁷ and the *mfold* program (<http://www.bioinfo.rpi.edu/applications/mfold/old/dna/form1.cgi>).

annealing reaction,^{3,4,7,11,14,29} and by preventing non-specific self-primed cDNA synthesis by RT.^{6-8,10,28,30} Also, NC interacts with RT in a specific manner,^{31,32} which enhances both RT polymerization and associated RNase H activities.^{12,33-36}

The first step of NC chaperoning TAR-cTAR annealing corresponds to the destabilization of TAR and cTAR structures upon binding of NC molecules.³⁷⁻⁴⁰ Indeed, NC activates the transient opening (fraying) of cTAR terminal base-pairs leading to a partial melting of the cTAR ends^{37,39} that propagates up to the middle of the stem.^{41,42} NC also melts TAR but less efficiently than cTAR. This is clearly related to the higher stability of the TAR RNA structure, since the DNA analogue of TAR (dTAR) (Figure 1(b)) is melted by NC as efficiently as cTAR.³⁹ The NC fingers in their proper context and the basic 29-RAPRKKG-35 linker support the destabilizing activity.⁴¹ Indeed, these NC structural elements form a hydrophobic platform that is critical for oligonucleotide binding.^{41,43-45} The second step in NC chaperoning of minus-strand transfer most probably corresponds to the hybridization of cTAR with TAR while being coated by NC molecules.^{3,11,14} This step appears to be promoted by the numerous basic residues in NC (Figure 1(a)) that may neutralize the negative phosphate charges of the DNA, thus decreasing the electrostatic repulsion between the nucleic acid molecules. However, the mechanistic aspects of NC-promoted annealing of cTAR to TAR are still poorly understood.

To characterize in depth how HIV-1 NC chaperones the minus-strand DNA transfer, we investigated the annealing kinetics of TAR with cTAR labelled at its 3' and 5' ends by two chromophores exhibiting fluorescence resonance energy transfer (FRET). Moreover, the structural determinants of the hybridization reaction were investigated using cTAR and dTAR derivatives. Taken together, our results indicate that NC chaperones TAR-cTAR annealing according to two distinct bimolecular reactions, a rapid one involving a limited annealing of TAR and cTAR terminal stem and a slower reaction leading to extended duplex formation.

Results and Discussion

Kinetics of cTAR DNA annealing to TAR RNA in the absence of NC

To investigate the kinetics of cTAR annealing to TAR, non-labelled TAR was mixed with cTAR labelled at its 5' and 3' ends by 5-(and 6)-carboxyfluorescein (Fl) and 5-(and 6)-carboxytetramethylrhodamine (TMR), respectively. In the absence of NC and TAR, cTAR exhibits a conformational equilibrium between closed and partially melted species.^{38,39} In the closed species, the Fl and TMR dyes are in close proximity, forming a non-fluorescent heterodimer. In the partly melted

species, Fl and TMR are away from each other and thus exhibit a FRET that partly quenches the Fl fluorescence in a distance-dependent manner. Since the closed species are by far the major species, and because most melted species experience only a limited stem opening, the fluorescence of TMR-5'-cTAR-3'-Fl is very low (Figure 2(a) inset, continuous line). In contrast, formation of the 55 bp TAR/cTAR extended duplex (ED) strongly increases the interchromophore distance, inducing the disappearance of the TMR emission peak at 580 nm and a full recovery of the Fl emission peak at 520 nm (Figure 2(a) inset, dotted line).⁴⁶ Consequently, real-time annealing kinetics can be monitored in solution with high sensitivity.⁴⁶

As a representative example, the annealing kinetics of 10 nM TMR-5'-cTAR-3'-Fl with 1 μ M TAR in 25 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.2 mM MgCl₂ at 20 °C is shown in Figure 2(a). In agreement with previous reports on the hybridization of the replication primer tRNA^{Lys,3} to the genomic PBS,⁴⁷ and the annealing of the 3' R RNA to the complementary 5' R DNA,¹⁴ TAR was found to anneal to its complementary sequence with a slow kinetics, since completion of the reaction

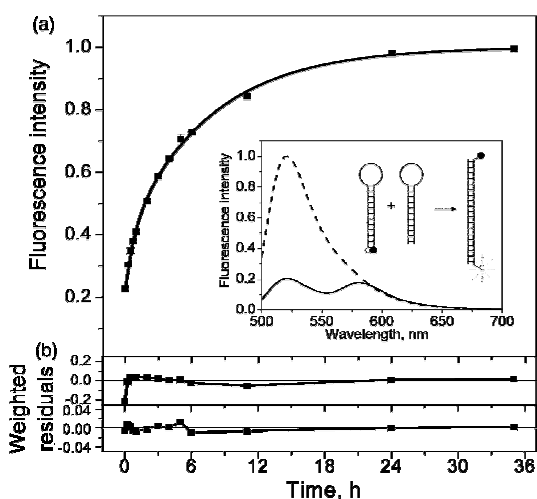


Figure 2. Kinetics of TAR-cTAR annealing in the absence of protein. (a) Kinetic trace of 10 nM doubly labelled cTAR with 1 μ M TAR in 25 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.2 mM MgCl₂ at 20 °C. Excitation and emission wavelengths were 480 nm and 520 nm, respectively, to monitor the Fl emission. The continuous line corresponds to the best fit of the data according to equation (1) with the k_{obs1} , k_{obs2} and x values given in the text. Inset: emission spectra of 10 nM doubly labelled cTAR in the absence of TAR (continuous) and after completion of the annealing reaction with TAR (broken). The low fluorescence of cTAR in the absence of TAR is related to the close proximity of Fl (open circle) and TMR (filled circle). Formation of the extended duplex strongly increases the interchromophore distance and thus, restores Fl emission. (b) Weighted residuals for a single-exponential (upper) and a two-exponential (lower) fit of the experimental data.

needed about 36 h, although the concentration of TAR was 100 times that of cTAR. The annealing reaction can be considered complete, since the fivefold fluorescence increase reached at the plateau was independent of the concentration of TAR (from 1 μM to 10 μM) and close to that obtained when annealing was promoted by a heating-cooling cycle.⁴⁶ As a first approach, the kinetics trace was fitted with a single-exponential function:

$$I(t) = I_f - (I_f - I_0)e^{-k_{\text{obs}}(t-t_0)}$$

where t_0 is the dead-time, k_{obs} is the observed kinetics rate constant, and I_0 and I_f are the fluorescence intensities of the stem-loop and the ED, respectively. A reasonable fit with a k_{obs} value of $5.4 \times 10^{-5} \text{ s}^{-1}$ was obtained by using this equation. However, the non-random deviations of the residuals from the zero value, especially at short times (Figure 2(b), upper curve), suggested that an additional fast kinetics component was present. To take this into account, a two-exponential function was used to fit the kinetics trace:

$$I(t) = I_f - (I_f - I_0)(x e^{-k_{\text{obs1}}(t-t_0)} + (1-x)e^{-k_{\text{obs2}}(t-t_0)}) \quad (1)$$

An improved fit was obtained using this equation, as could be seen from the nearly random distribution of the residuals around the zero value (Figure 2(b), lower curve) and the significantly improved reduced χ^2 value (0.999 *versus* 0.933 for the single-exponential fit). The rate constant k_{obs1} ($1.8 \times 10^{-3} \text{ s}^{-1}$) of the fast component was about two orders of magnitude higher than the rate constant k_{obs2} ($5.1 \times 10^{-5} \text{ s}^{-1}$) of the slow component. The amplitude x of the fast component was about 0.25. In a next step, the kinetics was repeated with increasing concentrations of TAR in pseudo-first-order conditions. While the amplitude x was independent of the concentration of RNA (data not shown), both the k_{obs2} and k_{obs1} values varied linearly with the concentration of TAR (Figure 3 and inset), indicating that both components followed second-order kinetics. Such a kinetics implies that the rate-limiting step of both components is probably the formation of an initial short, double-stranded nucleus between the two complementary TAR sequences.⁴⁸

The second-order kinetics of TAR-cTAR hybridization differs from the first-order kinetics between the 5' R DNA of 81 nt in length (encompassing the cTAR DNA sequence) and the complementary 3' R RNA (encompassing the TAR sequence).¹⁴ This first-order kinetics has been attributed to an unusual slow unfolding of the secondary structure of the R sequence followed by a more rapid nucleation. The discrepancy between the R data and ours might be due to the additional nucleotides of the R sequences and/or the different experimental conditions that may induce a different mechanism or prevent the formation of the nucleus

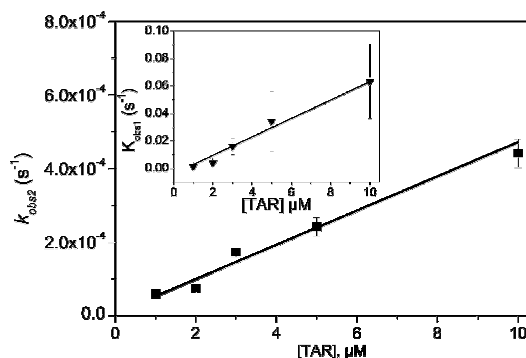


Figure 3. Kinetic parameters of cTAR-TAR annealing in the absence of protein. The observed annealing rate constants for the fast (k_{obs1}) (\blacktriangledown) and slow (k_{obs2}) (\blacksquare) components were determined in pseudo-first order conditions as described for Figure 2(a) and plotted as a function of the TAR concentration. The continuous line corresponds to the fit of the data with equation (2) and the k_{on} values given in Table 1. The k_{off} values are close to 0.

between the TAR and cTAR sequences. Alternatively, since the R and TAR sequences were from different HIV strains, a dependence of the annealing kinetics on the base composition of the TAR and cTAR sequences can not be excluded too.

According to the linear dependence of the $k_{\text{obs1,2}}$ values as a function of the RNA concentration (Figure 3), the association ($k_{\text{on1,2}}$) and dissociation ($k_{\text{off1,2}}$) rate constants of both slow and fast components were determined according to:

$$k_{\text{obs1,2}} = k_{\text{on1,2}}[\text{TAR}] + k_{\text{off1,2}} \quad (2)$$

The bimolecular rate constants k_{on1} and k_{on2} were found to be $7200(\pm 700) \text{ M}^{-1} \text{ s}^{-1}$ and $43(\pm 3) \text{ M}^{-1} \text{ s}^{-1}$ (Table 1), respectively. Both k_{on} values were several orders of magnitude smaller than the 10^5 – $10^7 \text{ M}^{-1} \text{ s}^{-1}$ rate constants reported for the annealing of unstructured sequences.⁴⁸ Interestingly, an even slower rate constant ($4 \text{ M}^{-1} \text{ s}^{-1}$) was reported for primer tRNA annealing to the PBS.⁴⁷ This was attributed to the requirement of the sequences involved in the formation of the PBS-tRNA nucleus to be in a single-stranded reactive conformation and a proper orientation. Since the sequences involved in the nucleus are double-stranded in PBS and tRNA, the slow reaction kinetics of PBS-tRNA annealing is related to the low probability of these double-stranded regions to be melted at room temperature. It follows that cTAR-TAR annealing is also probably nucleated through nucleotides of the double-stranded regions of both cTAR and TAR sequences. Moreover, since the k_{on1} value associated with the fast component is two orders of magnitude higher than the k_{on2} value of the slow component, the nucleotides involved in the nucleation of the fast component are probably in a less stable context

Table 1. Kinetic parameters of cTAR-TAR and cTAR-dTAR annealing

| Labeled sequence | Complementary sequence | NC(12-55) <i>r</i> =5 | k_{on1} ($M^{-1} s^{-1}$) | k_{on2} ($M^{-1} s^{-1}$) | χ | a_{lim} | K_{lim} (M^{-1}) |
|------------------|------------------------|--------------------------|-------------------------------|-------------------------------|--------|-----------|------------------------|
| cTAR | TAR | – | 7200 ± 700 | 43 ± 3 | 0.26 | – | – |
| | dTAR | – | 9000 ± 800 | 85 ± 20 | 0.22 | – | – |
| | TAR | + | $> 10^6$ | $1.8(\pm 0.2) \times 10^4$ | 0.4 | 0.64 | 7×10^6 |
| | dTAR | + | $> 10^6$ | $1.3(\pm 0.1) \times 10^5$ | 0.21 | 0.54 | 2.6×10^6 |
| cTAR1 | dTAR | + | $> 10^6$ | $2.1(\pm 0.5) \times 10^4$ | 0.25 | 0.55 | 1.3×10^6 |
| cTARC12 | dTAR | + | – | $6.0(\pm 0.3) \times 10^4$ | 0 | – | – |
| cTARC7 | dTAR | + | – | $1.1(\pm 0.1) \times 10^5$ | 0 | – | – |
| cTAR | Δ LdTAR | + | $> 10^6$ | $1.1(\pm 0.2) \times 10^5$ | 0.23 | 0.54 | 3.0×10^6 |
| cTAR | (41-55)dTAR | + | $> 10^6$ | $1.2(\pm 0.2) \times 10^5$ | 0.36 | 0.62 | 5.3×10^6 |
| (14-39)cTAR | dTAR | + | $4.9(\pm 0.6) \times 10^4$ | $1.7(\pm 0.2) \times 10^3$ | 0.19 | – | – |
| (14-39)cTAR | (17-42)dTAR | + | $5.4(\pm 0.4) \times 10^4$ | $4.4(\pm 0.5) \times 10^3$ | 0.26 | – | – |

The kinetic rate constants for the fast (k_{on1}) and the slow (k_{on2}) components were calculated from the dependence of the k_{obs} values on the concentration of the unlabeled oligonucleotide, as described in Figure 3. The parameter χ corresponds to the amplitude of the fast component with a standard deviation of 0.02. The parameters a_{lim} and K_{lim} designate the lower limits for the fractional fluorescence and the equilibrium association constant associated with the unproductive complex. These parameters were not calculated for the annealing reactions with (14-39)cTAR, since the mechanism associated with these species is different (see the text).

than those involved in the nucleation of the slow process. The large difference in the k_{on} values also implies that the two kinetic components are clearly separated on the time scale and, consequently, the fast component is completed at early time-points of the kinetics (less than 2 h in Figure 2). Noticeably, since only a limited fluorescence increase is observed at early times, this indicates that the fast component involves only a limited population of the cTAR molecules and/or leads to a complex where the chromophores are close in order to allow an efficient FRET.

Moreover, the intercepts, and thus the k_{off} values, of both kinetics components are close to zero (Figure 3), indicating that the reverse reaction, the dissociation of the ED, is negligible, in line with the large differences in stability between the ED and the initial stem-loop structures.

Kinetics of cTAR-TAR annealing in the presence of NC protein

Next, we performed experiments with the NC(12-55) peptide, which contains the zinc finger motifs and the C-terminal domain. This NC peptide exhibits wild-type activity on cTAR destabilization,⁴¹ but does not aggregate oligonucleotides. This last property is critical, since aggregates preclude direct analyses in solution.⁴⁹ To perform the annealing reaction, NC(12-55) was first added to the doubly labelled cTAR and the unlabelled TAR, separately. The peptide was added at a ratio, r , of one NC molecule per 5 nt in order to ensure a complete coating of TAR and cTAR, and an optimal destabilization of both sequences.^{39,50} As reported,^{38,39} NC(12-55) markedly increased the fluorescence of the doubly labelled cTAR species and notably of the Fl peak at 520 nm (Figure 4 inset, dash-dot line) by shifting the equilibrium from the closed dark species to the open fluorescent species.

A further increase in fluorescence was observed when the two NC(12-55)-coated complementary

TAR and cTAR sequences were mixed together. The final fluorescence reached the same plateau value as in the absence of NC, suggesting that the reaction went to completion and that the same ED was obtained. Moreover, NC(12-55) induced a dramatic increase in the reaction kinetics, since the reaction was complete within few minutes

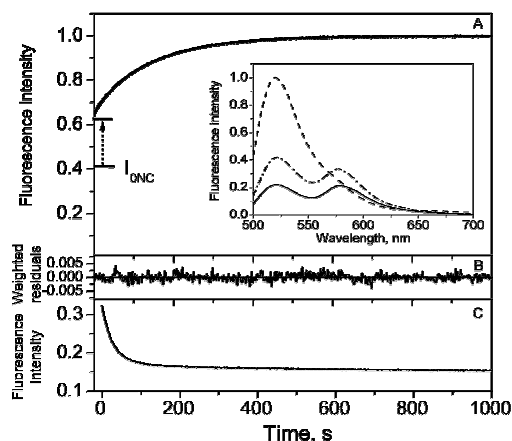


Figure 4. Kinetics of TAR-cTAR annealing in the presence of NC(12-55) or NC(1-55). (a) Kinetic trace of 10 nM doubly labelled cTAR with 400 nM TAR in the presence of NC(12-55) added at a ratio r of five nucleotides per molecule. Experimental conditions are as in Figure 2. The continuous line corresponds to the best fit of the data according to equation (3). The arrow shows the fluorescence jump due to the fast non-resolvable kinetic component. I_{0NC} designates the fluorescence intensity of the NC-coated cTAR molecules in the absence of TAR. Inset: emission spectra of doubly labelled cTAR in the absence of TAR and NC(12-55) (continuous), in the presence of NC(12-55) added at $r=5$ (dash-dot) and after completion of the annealing reaction with TAR and NC(12-55) (broken). (b) Weighted residuals for the fit of the experimental data to equation (3). (c) Kinetic trace of 10 nM doubly labelled cTAR with 100 nM TAR in the presence of NC(1-55) protein at $r=5$.

(compare Figures 2(a) and 4(a)). Due to this increase, it became impossible to kinetically resolve the fast component. Indeed, this fast component occurred mainly during the mixing time, as evidenced by the large fluorescence jump (Figure 4(a), arrow) between I_{0NC} , the fluorescence of NC(12-55)-coated cTAR in the absence of TAR and the initial time-point of the kinetic trace. Consequently, the kinetics traces were fit with a simplified version of equation (1):

$$I(t) = I_f - (I_f - I_{0NC})(1 - x)e^{-k_{obs2}(t-t_0)} \quad (3)$$

A good fit was obtained with this equation, as seen from the nearly random distribution of the residuals (Figure 4(b)). While the x value was independent of the concentration of RNA (data not shown), the k_{obs2} values varied linearly as a function of the concentration of RNA, suggesting that the slow component followed a second-order reaction. The k_{on2} value ($1.8 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) was about three orders of magnitude larger than that without NC (Table 1). Moreover, the k_{off2} value was close to zero, indicating that the reverse reaction was negligible and thus, that NC was unable to melt the ED. Assuming that the fast component is second-order, it can be further inferred from the dead-time of the measurements (10–20 s) and the initial fluorescence jump that $k_{on1} \geq 10^6 \text{ M}^{-1} \text{ s}^{-1}$, a value that is at least two orders of magnitude larger than the corresponding one in the absence of peptide. This large increase in the kinetic rate constants is in full agreement with the fact that NC strongly activates the annealing of cTAR to TAR during cDNA synthesis.^{3,14} Moreover, since the NC-promoted strand annealing is largely dependent on the stability of the initial oligonucleotides,^{8,30,39,50} cTAR/TAR annealing by NC(12-55) is correlated directly with the ability of NC to melt cTAR and TAR secondary structure and thus increase the availability of the single-stranded segments that nucleate the ED.

In a next step, the full-length NC(1–55) was substituted for NC(12-55) in order to evaluate the influence of the N-terminal basic domain on the annealing reaction. Even at the lowest concentrations of cTAR and TAR, we observed a kinetically non-resolvable fluorescence increase followed by a rapid decrease (Figure 4(c)). This decrease suggested that NC caused aggregation of nucleic acid in the form of nucleoprotein complexes in which, due to the high local concentration of condensed oligonucleotides, the FI dyes will be quenched by the TMR dyes on the same cTAR species and by TMR dyes of neighbour cTAR species in the nucleoprotein complex. This strong condensing activity of the full-length NC is in agreement with previous reports,^{49,51,52} and indicates that the NC N-terminal basic domain is a major determinant of this activity. This in turn favours cTAR/TAR annealing, as indicated by the fact that full-length NC is more efficient than

NC(12-55) in chaperoning cTAR-TAR annealing and minus strand DNA transfer.¹⁰

Effect of TAR sequence and stability on NC-promoted annealing kinetics

To get further insights into the mechanism of the NC(12-55)-promoted annealing reaction, we investigated in a next step the dependence of the annealing kinetics on the oligonucleotide sequence and stability. In a first step, TAR RNA has been substituted by its less stable DNA analogue, dTAR. Mixing of dTAR with cTAR in the presence of NC(12-55) led to biphasic kinetic traces (Figure 5). As with TAR, the amplitude x of the non-resolvable fast component was independent of the concentration of dTAR and the slow component followed a second-order reaction, suggesting that the NC(12-55)-promoted cTAR-TAR and cTAR-dTAR annealing reactions followed similar pathways. The k_{on2} rate constant ($1.3 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$) for dTAR is about tenfold faster than that for TAR (Table 1), in line with the lower stability of dTAR as compared with TAR that enables a more efficient melting by NC(12-55).³⁹ In addition, since cTAR and dTAR have similar stabilities and are similarly destabilized by NC(12-55),³⁹ and NC(12-55) mainly destabilizes the lower half of the cTAR stem,⁵⁰ we speculate that the lower half of cTAR and dTAR stems plays a critical role in the annealing reaction.

To confirm this hypothesis as well as the close connection between NC(12-55)-directed destabilization of the oligonucleotides and their subsequent

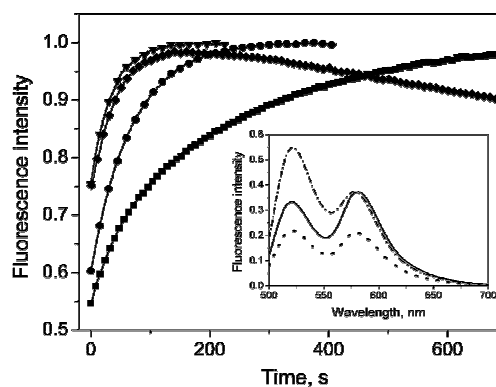


Figure 5. Kinetics of NC(12-55)-promoted cTAR-dTAR annealing. Kinetic traces of 10 nM doubly labelled cTAR (triangle), cTAR-C12 (circle) and cTAR1 (square) with 300 nM dTAR. Annealing of 10 nM doubly labelled cTAR with 300 nM (41-55)dTAR is also reported (diamond). NC(12-55) was added at $r=5$ in all the cases and experimental conditions were as in Figure 2. Inset: emission spectra of the doubly labelled cTAR12 either in the absence (continuous) or in the presence (broken) of NC(12-55). cTAR emission spectrum at the same concentration in the absence of protein is represented for comparison (dotted).

annealing, cTAR was substituted by the cTAR1 sequence in which the bulge at position 49 has been replaced by a base-pair (Figure 1(b)). This sequence has been shown to be more stable and less susceptible than cTAR to destabilization by NC(12-55).³⁸ The kinetics of the doubly labelled cTAR1 with dTAR in the presence of NC(12-55) shows again two components, one being non-kinetically resolved and one being second-order. The k_{on2} value was one order of magnitude lower than that of cTAR/dTAR (Table 1), indicating that the slow component was affected by the increased stability of the lower half of cTAR stem in line with an important role of this region in the annealing process.

To further investigate the role of the lower half of cTAR and dTAR stems in NC(12-55)-directed annealing, we investigated the interaction of the doubly labelled cTAR with (41-55)dTAR, the 3'-terminal sequence of dTAR (Figure 1(b)). This sequence is complementary to the 15 nt at the 5' terminus of cTAR and poorly folded (as predicted by mfold). The kinetics of cTAR(41-55)-dTAR was strikingly similar to that of cTAR-dTAR, since we observed similar fluorescence increases, k_{on2} values and amplitudes (Figure 5 and Table 1). This striking similarity suggests that the two kinetic components of the NC(12-55)-promoted cTAR-dTAR annealing rely on the interaction between the lower half of cTAR and dTAR stems. Moreover, the fluorescence of the cTAR-(41-55)dTAR complex at its maximum value is similar to that of cTAR-dTAR (Figure 5, compare triangles and diamonds), which indicates that the initial secondary structure of cTAR is fully melted in the cTAR-(41-55)dTAR complex. NC(12-55)-promoted invasion of the cTAR stem by (41-55)dTAR is probably driven by the perfect complementarities of (41-55)dTAR with the 15 nt of the 5' terminus of cTAR, leading to a perfect duplex that is more stable than the lower half of cTAR stem, which contains three bulges and one mismatch. This NC-promoted invasion of cTAR is fully consistent with the NC-promoted strand exchange that has been shown when an imperfect duplex was mixed with a perfect complement.⁵³

In line with its much smaller number of base-pairs, the cTAR-(41-55)dTAR complex appeared less stable than the cTAR-dTAR ED in the presence of NC. This lower stability is evidenced by the measurable k_{off2} value ($3(\pm 2) \times 10^{-3} \text{ s}^{-1}$) of the cTAR-(41-55)dTAR complex, which indicates that it can be dissociated by NC. Moreover, in contrast with the cTAR-dTAR complex, the fluorescence of the cTAR-(41-55)dTAR complex was found to decrease with time (see Figure 5 at $t > 150 \text{ s}$). This decrease of fluorescence may be due either to a slow NC-promoted conformational change that rearranges the (41-55)dTAR/cTAR duplex or to an additional NC-promoted slow dissociation pathway. Interestingly, a similar reversible strand invasion was observed when a short 13 nt RNA sequence complementary to the 3' end of cTAR was mixed with the full-length cTAR,⁴² indicating that

the strand invasion mechanism operates with both 5' and 3'-terminal sequences of cTAR. These data further confirm the essential role of the bulges,^{8,38} which limit the complementarity between the two strands of the hairpin and thus facilitate the invasion of the double-stranded sequences.

To get further insight into the role of the lower half of the cTAR stem in cTAR-dTAR annealing, we investigated the contribution of the two terminal double-stranded segments of cTAR by substituting either all the 12 terminal residues or only the 3'-terminal seven residues by C residues in the cTARC12 and cTARC7 mutants, respectively (Figure 1(b)). In both mutants, the two terminal double-stranded segments are prevented but, in contrast to cTARC12, the 5'-terminal residues of cTARC7 can still anneal to the 3'-terminal complementary residues of dTAR. As expected, due to the absence of the terminal double-stranded segments, the initial fluorescence intensity of both mutants was significantly higher than that of cTAR (Figure 5 inset). Interestingly, NC(12-55) increased the fluorescence of both mutants significantly, confirming that NC(12-55) is able to melt the central double-stranded segment (with seven base-pairs and one mismatch) in line with the cTAR data.^{38,39} With both mutants, the most striking feature of their NC(12-55)-directed hybridization with dTAR was the disappearance of the fast kinetic component ($x=0$, Table 1). This suggests that the fast component of the NC-promoted cTAR-TAR or cTAR-dTAR annealing reaction is associated with the formation of a complex where annealing is limited to the residues of the two terminal double-stranded segments. The limited annealing in this complex is in line with the limited increase of fluorescence associated with the fast component of the NC-promoted cTAR-TAR or cTAR-dTAR annealing reactions. This complex is clearly not an obligatory intermediate for the formation of ED, since its absence affects the formation of the cTARC12-dTAR ED only marginally. Indeed, both the k_{on2} value (Table 1) and the fluorescence value at the plateau (Figure 5) of the cTARC12/dTAR system are close to those of the cTAR/dTAR system. As a consequence, the complex associated with the fast component appears as an unproductive complex (UC) that does not lead to the ED. Formation of this UC would still be possible with cTARC7 that contains a native 5' terminus. However, since it would not further increase the distance between the chromophores, it could not be detected.

In addition, since the rate-limiting step of a second-order annealing reaction is usually the formation of a short double-strand nucleus between the two complementary oligonucleotides, the similar k_{on2} values for cTARC7-dTAR, cTARC12-dTAR and cTAR-dTAR (Table 1) suggest that (i) the residues nucleating the ED are located mainly in the central double-stranded segment and (ii) the terminal residues of cTAR stem contribute marginally to the nucleation of the ED.

Role of the TAR loops in NC(12-55)-directed annealing

The above results and notably those with cTAR-(41-55)dTAR further suggest that the loops do not play a major role in the NC(12-55)-promoted annealing of cTAR with TAR or dTAR. To further investigate the specific role of the loops in the annealing reaction, additional mutants have been used. In a first step, the doubly labelled cTAR was reacted with Δ LdTAR, a dTAR mutant in which the loop has been substituted by a flexible hexaethylenglycol tether (Figure 1(b)). The NC(12-55)-promoted annealing of cTAR with Δ LdTAR was very similar to that of cTAR with dTAR (Table 1), confirming that the loops play only a marginal role in the NC(12-55)-directed annealing of cTAR with dTAR. This loop-independent annealing pathway is fully consistent with the previously reported annealing between the full-length cTAR and an RNA oligonucleotide representing the lower half of the TAR sequence.⁴²

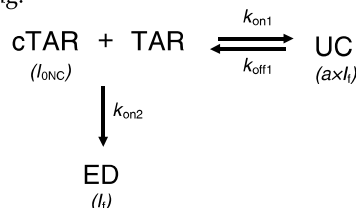
In a next step, we checked the existence of a NC(12-55)-directed pathway through the upper half of the stem by investigating the interaction of the doubly labelled (14-39)cTAR derivative corresponding to cTAR upper half (Figure 1(b)) with either (17-42)dTAR, the corresponding complementary sequence or the full-length dTAR. In both cases, we observed similar fluorescence increases and kinetics (Table 1). Again, a two-exponential equation was required to fit the data. Both slow and fast components were found to be second-order. The k_{on2} values were about two orders of magnitude smaller than that for cTAR-dTAR. Since a difference of at least two orders of magnitude could also be deduced for the k_{on1} value, this suggests that NC(12-55)-promoted hybridization through the upper half of cTAR stem is possible but not kinetically favoured with respect to the invasion pathway through the lower half of the stem. Moreover, since the upper half of cTAR stem is much less susceptible than the lower half to the destabilizing effect of NC(12-55),⁵⁰ our data confirmed the critical relationship between the annealing kinetics and the NC(12-55)-promoted destabilization of the oligonucleotides. The annealing pathway through the upper half of cTAR stem is in line with the previously reported annealing between the full-length cTAR sequence and the upper half of the TAR sequence.⁴² Moreover, this upper pathway was found to be less efficient than the lower pathway, though the differences were smaller than in our study.

Finally, to check if the loops participate in the pathway involving the upper half of cTAR and dTAR stems, we monitored the annealing of Δ LdTAR with the doubly labelled (14-39)cTAR derivative as well as the annealing of dTAR with the doubly labelled Δ L(14-39)cTAR, a mutant of the upper half of cTAR in which the loop has been substituted by a hexaethylenglycol tether (Figure 1(b)). In both systems, no fluorescence

increase was observed (data not shown), indicating that no ED formed under these conditions. As a consequence, it can be inferred that the annealing pathway through the upper half of the cTAR stem requires a loop-loop interaction.

A proposed reaction mechanism

On the basis of our kinetic data and the conclusions inferred with the cTAR mutants, the simplest reaction scheme describing the NC(12-55)-promoted cTAR-TAR annealing reaction is the following:



where ED and UC are the extended duplex and the unproductive complex, respectively. I_{0NC} , $a \times I_f$ and I_f are the fluorescence intensities associated with the different species.

Since the formation of UC, associated with the fast component has been found to be several orders of magnitude faster than the formation of ED associated with the slow component, the two steps may be considered as sequential and can be treated separately. Since we are working in pseudo first-order conditions ($[\text{TAR}] \gg [\text{cTAR}]$), the following differential equation can be written for the formation of UC:

$$\frac{d(\text{cTAR}(t))}{dt} = -\frac{d(\text{UC}(t))}{dt} = -k_{on1}[\text{TAR}]_0 \text{cTAR}(t) + k_{off1} \text{UC}(t) \quad (4)$$

with:

$$\text{UC}(t) = [\text{cTAR}]_0 - \text{cTAR}(t)$$

Integration of this differential equation provides:

$$\begin{aligned} \text{cTAR}(t) = & \frac{1}{K[\text{TAR}]_0 + 1} [\text{cTAR}]_0 \\ & + \left(1 - \frac{1}{K[\text{TAR}]_0 + 1}\right) [\text{cTAR}]_0 e^{-k_{obs1} \times t} \end{aligned} \quad (5)$$

with:

$$k_{obs1} = k_{on1}[\text{TAR}]_0 + k_{off1}$$

and:

$$K = k_{on1}/k_{off1}$$

As a consequence, the fluorescence intensity, $I(t)$, as a function of time could be described by:

$$I(t) = \frac{\text{cTAR}(t)}{[\text{cTAR}]_0} I_{0NC} + \frac{\text{UC}(t)}{[\text{cTAR}]_0} a I_f \quad (6)$$

By substituting equation (5) into equation (6), we obtain:

$$I(t) = \frac{K[\text{TAR}]_0 a I_f + I_{0\text{NC}}}{K[\text{TAR}]_0 + 1} - \frac{(I_f K[\text{TAR}]_0 a - K[\text{TAR}]_0 I_0)}{K[\text{TAR}]_0 + 1} e^{-k_{\text{obs1}} t} \quad (7)$$

which can be expressed as:

$$I(t) = I_i - (I_f - I_0)x e^{-k_{\text{obs1}} t} \quad (8)$$

with:

$$x = \frac{I_i - I_0}{I_f - I_0} \quad \text{and} \quad I_i = \frac{K[\text{TAR}]_0 a I_f + I_0}{K[\text{TAR}]_0 + 1} \quad (9)$$

where I_i describes the fluorescence intensity obtained when the equilibrium in the fast step is reached.

For the second step corresponding to the ED formation, the concentrations of cTAR and UC may be considered as equilibrated and the following differential expression can be written:

$$\frac{d(\text{cTAR}(t))}{dt} = -k_{\text{on2}}[\text{TAR}]_0 \text{cTAR}(t) \quad (10)$$

Integration of this equation provides:

$$\text{cTAR}(t) = \frac{1}{K[\text{TAR}]_0 + 1} [\text{cTAR}]_0 e^{-k_{\text{obs2}} t} \quad (11)$$

with:

$$k_{\text{obs2}} = k_{\text{on2}}[\text{TAR}]_0$$

Consequently, the concentrations UC(t) and ED(t) are given by:

$$\text{UC}(t) = \frac{K[\text{TAR}]_0}{K[\text{TAR}]_0 + 1} [\text{cTAR}]_0 e^{-k_{\text{obs2}} t} \quad \text{and} \quad (12)$$

$$\text{ED}(t) = [\text{cTAR}]_0 (1 - e^{-k_{\text{obs2}} t})$$

The fluorescence intensity can then be expressed by:

$$I(t) = \frac{\text{cTAR}(t)}{[\text{cTAR}]_0} I_{0\text{NC}} + \frac{\text{UC}(t)}{[\text{cTAR}]_0} a I_f + \frac{\text{ED}(t)}{[\text{cTAR}]_0} I_f \quad (13)$$

Substituting equation (12) in equation (13), we obtain:

$$I(t) = I_f - \left(I_f - \frac{1}{K[\text{TAR}]_0 + 1} I_{0\text{NC}} - \frac{K[\text{TAR}]_0 a}{K[\text{TAR}]_0 + 1} I_f \right) e^{-k_{\text{obs2}} t} \quad (14)$$

This equation is formally identical with equation (3), with x taking the value given in equation (9). From these equations, it appears that the k_{on1} and k_{on2} values reported in Table 1 directly describe the association rate constants for the formation of UC and ED, respectively. Moreover, from equation (9),

it can be further concluded that the x values (Table 1) depend both on the equilibrium constant K for the UC formation and the UC fluorescence intensity $a \times I_f$. From equation (9), the equilibrium constant K can be expressed as:

$$K = \frac{x(I_f - I_{0\text{NC}})}{[\text{TAR}]_0 [I_f(a - x) - I_{0\text{NC}}(1 - x)]} \quad (15)$$

By substituting $[\text{TAR}]_0$, I_f , $I_{0\text{NC}}$ and x by their values, it can be inferred that $a > 0.64$. Moreover, since the UC fluorescence cannot exceed the ED fluorescence, $I_f 0.64 < a \leq 1$. As expected, $a \times I_f > I_{0\text{NC}}$ ($\approx 0.4 I_f$), indicating that UC formation further decreases the population of closed dark cTAR species. This confirms that the fluorescence of NC-coated doubly labelled cTAR species increased when TAR is present.⁴⁰ The a value is not modified strongly when TAR is substituted by dTAR, ΔL dTAR or (41-55)dTAR (Table 1), suggesting that the UC complexes are similar. In addition, since x is constant at the concentrations of oligonucleotide used in the present experiments, the lower limit, K_{lim} , of the equilibrium constant for UC formation can be calculated from equation (15). The K_{lim} values are larger than 10^6 M^{-1} for all the sequences tested and are thus several orders of magnitude higher than those calculated by mfold (data not shown) for an annealing limited at the terminal residues of cTAR and TAR species. This suggests that NC(12-55) stabilizes UC through its binding. This stabilization probably results from the neutralization of the TAR and cTAR phosphate charges by the positively charged NC(12-55) peptide that decreases the repulsive interactions between the single-stranded segments involved in the formation of UC. Moreover, as expected from the higher stability of RNA/DNA duplexes with respect to DNA/DNA duplexes,⁴⁸ the largest K_{lim} value is obtained for the UC between TAR and cTAR.

In a next step, based on this kinetic scheme and our data, we propose a molecular model in Figure 6. The upper pathway (steps a and b), which is associated with the fast kinetic component, involves a limited annealing of cTAR and TAR termini in UC and does not lead to the formation of ED. Since this pathway is not observed with the cTAR12 mutant, the annealing involves only residues belonging to either or both of the two terminal double strands of cTAR and TAR. This pathway is in line with the reported NC(12-55)-promoted fraying of the cTAR terminal double-stranded segments (step a), which increases the concentration of reactive single-stranded nucleotides needed for nucleating the UC.^{37,38}

The second pathway (steps a', c and d), which is associated with the slow kinetic component, leads to ED. The rate-limiting step of this invasion pathway is thought to be the nucleation of the duplex through residues of the central double-stranded segment of both cTAR and TAR stems (step c). Then, this nucleation is followed by a fast

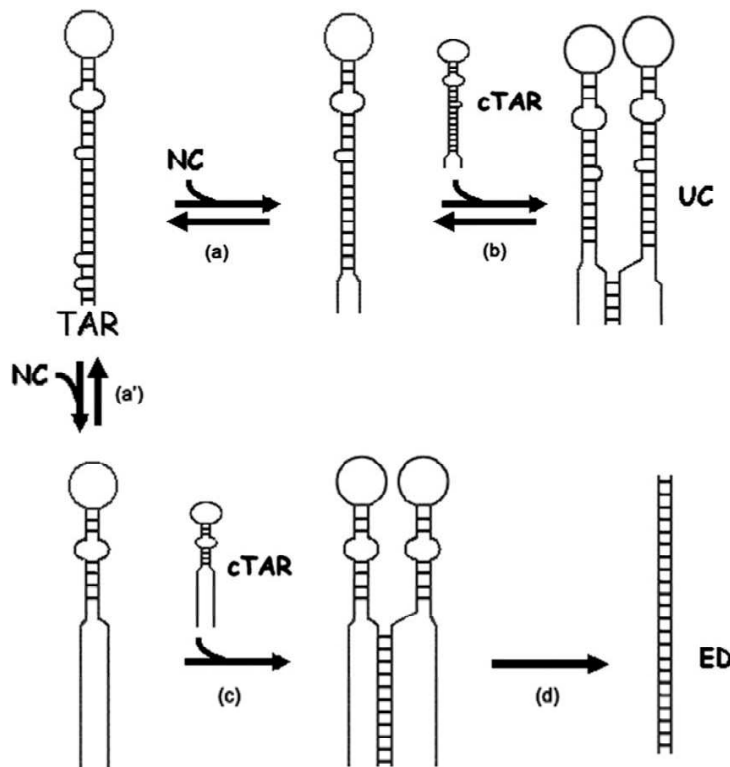


Figure 6. Proposed mechanism for NC(12-55)-promoted cTAR-TAR annealing. The annealing reaction is thought to involve two pathways. The upper pathway (steps a and b) is associated with the fast kinetic component and corresponds to the formation of the unproductive complex (UC). In this pathway, NC(12-55) melts the terminal double-stranded segments of the TAR stem (step a) and promotes a limited annealing with the corresponding segments of cTAR stem (step b). The lower pathway (steps a', c and d) is associated with the slow kinetic component and corresponds to the formation of the extended duplex (ED). In this pathway, NC(12-55) propagates the melting up to the middle of the stem (step a') and nucleates ED through residues located mainly within the central double-stranded segment of both cTAR and TAR stems (step c). Formation of the rate-limiting nucleus is followed by a fast zipping reaction (step d).

zipping reaction (step d) leading to ED. Formation of the nucleus requires that the central double-stranded segment should be melted (step a'). This requirement is fully consistent with the ability of NC(12-55) to promote the melting of cTAR up to the middle of the stem.^{38,39,54} Consequently, the TAR and cTAR species where the lower half of the stem is melted are thought to be the reactive species for ED formation. The slower kinetics of the ED pathway as compared with the UC one is fully consistent with the destabilization kinetics, since the NC-promoted fraying is in the microsecond range,^{37,50} while the NC-promoted melting of the cTAR lower half is in the millisecond range.⁵⁴ The ED kinetic pathway is very similar to the "zipping mechanism" proposed recently for the NC-promoted annealing of TAR and cTAR sequences from the NL4-3 strain.⁴² Since the TAR sequences from the NL4-3 and MAL sequences differ in their base composition as well as in the number and position of their bulges, this suggests that the invasion mechanism is of general relevance for NC-promoted cTAR-TAR annealing in HIV-1. In contrast, the UC pathway has not been observed with the TAR sequence from the NL4-3 strain.⁴² This may be related to the increased length and stability of its terminal double-stranded segment (5 bp *versus* only 3 bp in the MAL strain) that probably limit its fraying by NC.

No critical role for the TAR or cTAR loops was apparent in the strand invasion pathway. However,

a still slower pathway (not represented here) that involves the upper half of cTAR and TAR species and relies on loop-loop interaction has been observed with truncated cTAR mutants (data not shown).⁴² Due to its slow rate constant, the contribution of this loop-loop pathway is probably negligible in the NC(12-55)-promoted cTAR-TAR annealing under our conditions. However, this does not exclude the possibility that this pathway could be favoured in a different context (depending on the oligonucleotide length and sequence, the nature of the NC peptide, the buffer composition, etc.), since the TAR and cTAR loops have been shown to increase the minus-strand transfer reaction in previous reports,^{55,56}. This is under investigation.

Conclusion

In the present study, the real-time kinetics of cTAR-TAR annealing was investigated both in the absence and in the presence of NC protein. The annealing reaction involves two distinct kinetic components characterized by second-order rate constants. NC(12-55) does not modify the number of kinetic components and the molecular nature of the annealing reaction but instead, activates the kinetic rate constants by several orders of magnitude. Our data indicate further that NC activation of cTAR-TAR annealing is correlated with its ability to destabilize the lower half of the more stable TAR

sequence (Figure 6). Therefore, the nature of the TAR structure is a major determinant for TAR-cTAR annealing kinetics. This is in line with the inability of NC to chaperone the formation of an RNA/DNA duplex when the stability of the acceptor RNA is increased, regardless of the stability of the complementary DNA.^{2,3,8} Moreover, since NC has been shown to destabilize the least stable parts of a stem-loop structure,^{38,50} the kinetics of the NC-promoted annealing of two complementary sequences are likely governed by the destabilization of the least stable secondary structure of the most stable of the two complementary oligonucleotides.

The destabilization step is required to provide the single-stranded complementary regions for nucleating the duplex structures. The two kinetic components have been assigned to two sequential pathways. The first one is associated with a limited annealing of the terminal bases of cTAR to TAR but does not lead to ED formation. The biological relevance of this step is unclear and may be related to our experimental conditions. Further experiments will be required to further check this point. The ED is formed in a second phase that is limited kinetically by the nucleation of residues located mainly within the central double-stranded segment of both cTAR and TAR stems. This invasion mechanism through the lower half of TAR and cTAR stems is fully consistent with previous reports showing that NC promotes the invasion of the donor-cDNA by the acceptor RNA at the base of the TAR sequence.^{34,57} Finally, an alternative mechanism involving the interaction through the loops has been observed but is a minor pathway in the present conditions.

Materials and Methods

The NC(12-55) and NC(1-55) peptides (Figure 1(a)) were synthesized as described⁵⁸ and stored lyophilized in their zinc-bound form. Their purity was greater than 98%. An extinction coefficient of $5700 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm was used to determine their concentration.

The unlabelled and the doubly labelled oligonucleotides (Figure 1(b)) were synthesized at a 1 μmol and a 0.2 μmol scale, respectively, by IBA GmbH Nucleic Acids Product Supply (Göttingen, Germany). The 5' terminus of the oligonucleotides was labelled by 5/6 tetramethylrhodamine (TMR) *via* an amino-linker with six-carbon spacer. The 3' terminus was labelled with 5/6 fluorescein (Fl). Oligonucleotides were purified by the manufacturer by reverse-phase HPLC and polyacrylamide gel electrophoresis.

Experiments were performed in 25 mM Tris-HCl (pH 7.5), 30 mM NaCl, 0.2 mM MgCl_2 at 20 °C. Emission spectra and kinetic traces were registered with a Fluorolog and FluoroMax spectrofluorimeters (Jobin Yvon Instruments, S.A. Inc.) equipped with a temperature-controlled cell compartment. All fluorescence intensities were corrected for buffer emission and lamp fluctuations.

Kinetic measurements were performed in pseudo first-order conditions by using TAR or dTAR concentrations at least tenfold higher than the concentration of the labelled

complementary sequence. Excitation and emission wavelengths were 480 nm and 520 nm, respectively, to monitor the Fl fluorescence. All reported concentrations of reactants correspond to those after mixing. To avoid high local concentrations during mixing, both reactants were of the same volume. When the reaction was performed in the presence of NC(12-55), the peptide was added to each reactant separately at a nucleotide to peptide ratio $r=5$. The reaction was then triggered by mixing the NC-coated oligonucleotides together. Since the kinetics in the presence of NC is fast, the fluorescence intensities could be monitored continuously without photo-bleaching. In contrast, due to the slow kinetics without NC, measurements were performed at selected times. The apparent rate constants k_{obs} and the amplitudes were determined from the kinetics data by including a dead-time correction t_0 to take into account the delay between the mixing of reactants and the start of the measurements. All fitting procedures were carried out with Microcal™ Origin™ 6.1 software based on the non-linear, least-squares method and applying the Levenberg-Marquardt algorithm.

Acknowledgements

This work was supported by grants from the Agence Nationale de Recherches sur le SIDA (ANRS) and the European Community (TRIOH integrated project). C.R. is a fellow from Sidaction.

References

- Arts, E. J. & Wainberg, M. A. (1996). Human immunodeficiency virus type 1 reverse transcriptase and early events in reverse transcription. *Advan. Virus Res.* **46**, 97–163.
- Allain, B., Lapadat-Tapolsky, M., Berlioz, C. & Darlix, J. L. (1994). Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *EMBO J.* **13**, 973–981.
- Darlix, J. L., Vincent, A., Gabus, C., de Rocquigny, H. & Roques, B. (1993). Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. *C.R. Acad. Sci. III*, **316**, 763–771.
- Davis, W. R., Gabbara, S., Hupe, D. & Peliska, J. A. (1998). Actinomycin D inhibition of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase and nucleocapsid protein. *Biochemistry*, **37**, 14213–14221.
- DeStefano, J. J. (1995). Human immunodeficiency virus nucleocapsid protein stimulates strand transfer from internal regions of heteropolymeric RNA templates. *Arch. Virol.* **140**, 1775–1789.
- Guo, J., Henderson, L. E., Bess, J., Kane, B. & Levin, J. G. (1997). Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J. Virol.* **71**, 5178–5188.
- Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J. *et al.* (2000). Zinc finger structures

- in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.* **74**, 8980–8988.
8. Heilman-Miller, S. L., Wu, T. & Levin, J. G. (2004). Alteration of nucleic acid structure and stability modulates the efficiency of minus-strand transfer mediated by the HIV-1 nucleocapsid protein. *J. Biol. Chem.* **279**, 44154–44165.
 9. Kim, J. K., Palaniappan, C., Wu, W., Fay, P. J. & Bambara, R. A. (1997). Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1. *J. Biol. Chem.* **272**, 16769–16777.
 10. Lapadat-Tapolsky, M., Gabus, C., Rau, M. & Darlix, J. L. (1997). Possible roles of HIV-1 nucleocapsid protein in the specificity of proviral DNA synthesis and in its variability. *J. Mol. Biol.* **268**, 250–260.
 11. Lapadat-Tapolsky, M., Pernelle, C., Borie, C. & Darlix, J. L. (1995). Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1. *Nucl. Acids Res.* **23**, 2434–2441.
 12. Peliska, J. A., Balasubramanian, S., Giedroc, D. P. & Benkovic, S. J. (1994). Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. *Biochemistry*, **33**, 13817–13823.
 13. Rodriguez-Rodriguez, L., Tsuchihashi, Z., Fuentes, G. M., Bambara, R. A. & Fay, P. J. (1995). Influence of human immunodeficiency virus nucleocapsid protein on synthesis and strand transfer by the reverse transcriptase *in vitro*. *J. Biol. Chem.* **270**, 15005–15011.
 14. You, J. C. & McHenry, C. S. (1994). Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J. Biol. Chem.* **269**, 31491–31495.
 15. Berkhout, B. (2000). Multiple biological roles associated with the repeat (R) region of the HIV-1 RNA genome. *Advan. Pharmacol.* **48**, 29–73.
 16. Henderson, L. E., Bowers, M. A., Sowder, R. C., 2nd, Serabyn, S. A., Johnson, D. G., Bess, J. W., Jr *et al.* (1992). Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. *J. Virol.* **66**, 1856–1865.
 17. Mervis, R. J., Ahmad, N., Lillehoj, E. P., Raum, M. G., Salazar, F. H., Chan, H. W. & Venkatesan, S. (1988). The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors. *J. Virol.* **62**, 3993–4002.
 18. Swanstrom, R. & Wills, J. W. (1997). Retroviral gene expression: Synthesis, Assembly and Processing of Viral Proteins. In *Retroviruses* (Varmus, H. E., Coffin, J. M. & Hughes, S. H., eds), pp. 263–334, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
 19. Mely, Y., De Rocquigny, H., Morellet, N., Roques, B. P. & Gerard, D. (1996). Zinc binding to the HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence spectroscopy. *Biochemistry*, **35**, 5175–5182.
 20. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., Jr, South, T. L. *et al.* (1992). Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. *Protein Sci.* **1**, 563–574.
 21. Morellet, N., de Rocquigny, H., Mely, Y., Jullian, N., Demene, H., Ottmann, M. *et al.* (1994). Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by ¹H NMR. *J. Mol. Biol.* **235**, 287–301.
 22. Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J. L. & Roques, B. P. (1992). Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by ¹H NMR. *EMBO J.* **11**, 3059–3065.
 23. Omichinski, J. G., Clore, G. M., Sakaguchi, K., Appella, E. & Gronenborn, A. M. (1991). Structural characterization of a 39-residue synthetic peptide containing the two zinc binding domains from the HIV-1 p7 nucleocapsid protein by CD and NMR spectroscopy. *FEBS Letters*, **292**, 25–30.
 24. South, T. L., Blake, P. R., Sowder, R. C., 3rd, Arthur, L. O., Henderson, L. E. & Summers, M. F. (1990). The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers. *Biochemistry*, **29**, 7786–7789.
 25. Cristofari, G. & Darlix, J. L. (2002). The ubiquitous nature of RNA chaperone proteins. *Prog. Nucl. Acid Res. Mol. Biol.* **72**, 223–268.
 26. Rein, A., Henderson, L. E. & Levin, J. G. (1998). Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem. Sci.* **23**, 297–301.
 27. Baudin, F., Marquet, R., Isel, C., Darlix, J. L., Ehresmann, B. & Ehresmann, C. (1993). Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J. Mol. Biol.* **229**, 382–397.
 28. Driscoll, M. D. & Hughes, S. H. (2000). Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences. *J. Virol.* **74**, 8785–8792.
 29. Guo, J., Wu, T., Bess, J., Henderson, L. E. & Levin, J. G. (1998). Actinomycin D inhibits human immunodeficiency virus type 1 minus-strand transfer in *in vitro* and endogenous reverse transcriptase assays. *J. Virol.* **72**, 6716–6724.
 30. Golinelli, M. P. & Hughes, S. H. (2003). Secondary structure in the nucleic acid affects the rate of HIV-1 nucleocapsid-mediated strand annealing. *Biochemistry*, **42**, 8153–8162.
 31. Lener, D., Tanchou, V., Roques, B. P., Le Grice, S. F. & Darlix, J. L. (1998). Involvement of HIV-1 nucleocapsid protein in the recruitment of reverse transcriptase into nucleoprotein complexes formed *in vitro*. *J. Biol. Chem.* **273**, 33781–33786.
 32. Druillennec, S., Caneparo, A., de Rocquigny, H. & Roques, B. P. (1999). Evidence of interactions between the nucleocapsid protein NCp7 and the reverse transcriptase of HIV-1. *J. Biol. Chem.* **274**, 11283–11288.
 33. Cameron, C. E., Ghosh, M., Le Grice, S. F. & Benkovic, S. J. (1997). Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. *Proc. Natl Acad. Sci. USA*, **94**, 6700–6705.
 34. Chen, Y., Balakrishnan, M., Roques, B. P. & Bambara, R. A. (2003). Steps of the acceptor invasion mechanism for HIV-1 minus strand strong stop transfer. *J. Biol. Chem.* **278**, 38368–38375.
 35. Chen, Y., Balakrishnan, M., Roques, B. P., Fay, P. J. & Bambara, R. A. (2003). Mechanism of minus strand strong stop transfer in HIV-1 reverse transcription. *J. Biol. Chem.* **278**, 8006–8017.

36. Roda, R. H., Balakrishnan, M., Hanson, M. N., Wohrl, B. M., Le Grice, S. F., Roques, B. P. *et al.* (2003). Role of the reverse transcriptase, nucleocapsid protein, and template structure in the two-step transfer mechanism in retroviral recombination. *J. Biol. Chem.* **278**, 31536–31546.
37. Azoulay, J., Clamme, J. P., Darlix, J. L., Roques, B. P. & Mely, Y. (2003). Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J. Mol. Biol.* **326**, 691–700.
38. Beltz, H., Azoulay, J., Bernacchi, S., Clamme, J. P., Ficheux, D., Roques, B. *et al.* (2003). Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. *J. Mol. Biol.* **328**, 95–108.
39. Bernacchi, S., Stoylov, S., Piemont, E., Ficheux, D., Roques, B. P., Darlix, J. L. & Mely, Y. (2002). HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol. Biol.* **317**, 385–399.
40. Hong, M. K., Harbron, E. J., O'Connor, D. B., Guo, J., Barbara, P. F., Levin, J. G. & Musier-Forsyth, K. (2003). Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer. *J. Mol. Biol.* **325**, 1–10.
41. Beltz, H., Clauss, C., Piemont, E., Ficheux, D., Gorelick, R. J., Roques, B. *et al.* (2005). Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J. Mol. Biol.* **348**, 1113–1126.
42. Liu, H. W., Cosa, G., Landes, C. F., Zeng, Y., Kovaleski, B. J., Mullen, D. G. *et al.* (2005). Single-molecule FRET studies of important intermediates in the nucleocapsid protein chaperoned minus-strand transfer step in HIV-1 reverse transcription. *Biophys. J.* **89**, 3470–3479.
43. Amarasinghe, G. K., Zhou, J., Miskimon, M., Chancellor, K. J., McDonald, J. A., Matthews, A. G. *et al.* (2001). Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition. *J. Mol. Biol.* **314**, 961–970.
44. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N. & Summers, M. F. (1998). Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science*, **279**, 384–388.
45. Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M. C. & Roques, B. P. (1998). Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J. Mol. Biol.* **283**, 419–434.
46. Bernacchi, S., Piemont, E., Potier, N., Dorsselaer, A. & Mely, Y. (2003). Excitonic heterodimer formation in an HIV-1 oligonucleotide labelled with a donor-acceptor pair used for fluorescence resonance energy transfer. *Biophys. J.* **84**, 643–654.
47. Hargittai, M. R., Gorelick, R. J., Rouzina, I. & Musier-Forsyth, K. (2004). Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J. Mol. Biol.* **337**, 951–968.
48. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. J. (2000). *Nucleic Acids Structures, Properties, and Functions*, University Science Books, Sausalito, CA.
49. Stoylov, S. P., Vuilleumier, C., Stoylova, E., De Rocquigny, H., Roques, B. P., Gerard, D. & Mely, Y. (1997). Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. *Biopolymers*, **41**, 301–312.
50. Beltz, H., Piemont, E., Schaub, E., Ficheux, D., Roques, B., Darlix, J. L. & Mely, Y. (2004). Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J. Mol. Biol.* **338**, 711–723.
51. Krishnamoorthy, G., Roques, B., Darlix, J. L. & Mely, Y. (2003). DNA condensation by the nucleocapsid protein of HIV-1: a mechanism ensuring DNA protection. *Nucl. Acids Res.* **31**, 5425–5432.
52. Le Cam, E., Coulaud, D., Delain, E., Petitjean, P., Roques, B. P., Gerard, D. *et al.* (1998). Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron microscopy investigation. *Biopolymers*, **45**, 217–229.
53. Urbaneja, M. A., Wu, M., Casas-Finet, J. R. & Karpel, R. L. (2002). HIV-1 nucleocapsid protein as a nucleic acid chaperone: spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J. Mol. Biol.* **318**, 749–764.
54. Cosa, G., Harbron, E. J., Zeng, Y., Liu, H. W., O'Connor, D. B., Eta-Hosokawa, C. *et al.* (2004). Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys. J.* **87**, 2759–2767.
55. Berkhout, B., Vastenhout, N. L., Klasens, B. I. & Huthoff, H. (2001). Structural features in the HIV-1 repeat region facilitate strand transfer during reverse transcription. *Rna*, **7**, 1097–1114.
56. Kanevsky, I., Chaminade, F., Ficheux, D., Moumen, A., Gorelick, R., Negroni, M. *et al.* (2005). Specific interactions between HIV-1 nucleocapsid protein and the TAR element. *J. Mol. Biol.* **348**, 1059–1077.
57. Chen, Y., Balakrishnan, M., Roques, B. P. & Bambara, R. A. (2005). Acceptor RNA cleavage profile supports an invasion mechanism for HIV-1 minus strand transfer. *J. Biol. Chem.* **280**, 14443–14452.
58. de Rocquigny, H., Ficheux, D., Gabus, C., Fournie-Zaluski, M. C., Darlix, J. L. & Roques, B. P. (1991). First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem. Biophys. Res. Commun.* **180**, 1010–1018.

Edited by M. F. Summers

(Received 21 October 2005; received in revised form 7 December 2005; accepted 12 December 2005)
Available online 27 December 2005

4.3

Publication IV

HIV-1 Nucleocapsid Traps Reverse Transcriptase on Nucleic Acid Substrates[†]Dina Grohmann,^{‡,§} Julien Godet,^{||} Yves Mély,^{||} Jean-Luc Darlix,[⊥] and Tobias Restle^{*,‡}

Institut Gilbert Laustriat, Photophysique des interactions moléculaires, UMR 7175 CNRS, Faculté de Pharmacie, Université Louis Pasteur, Strasbourg 1, 74, Route du Rhin, 67401 Illkirch, France, LaboRetro, INSERM 758, IFR 128, ENS de Lyon, 46 Allée d'Italie, 69364 Lyon, France, and Institut für Molekulare Medizin, Universitätsklinikum Schleswig-Holstein, Universität zu Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

Received July 24, 2008; Revised Manuscript Received September 21, 2008

ABSTRACT: Conversion of the genomic RNA of human immunodeficiency virus (HIV) into full-length viral DNA is a complex multistep reaction catalyzed by the reverse transcriptase (RT). Numerous studies have shown that the viral nucleocapsid (NC) protein has a vital impact on various steps during reverse transcription, which is crucial for virus infection. However, the exact molecular details are poorly defined. Here, we analyzed the effect of NC on RT-catalyzed single-turnover, single-nucleotide incorporation using different nucleic acid substrates. In the presence of NC, we observed an increase in the amplitude of primer extension of up to 3-fold, whereas the transient rate of nucleotide incorporation (k_{pol}) dropped by up to 50-fold. To unravel the underlying molecular mechanism, we carefully analyzed the effect of NC on RT–nucleic acid substrate dissociation. The studies revealed that NC considerably enhances the stability of RT–substrate complexes by reducing the observed dissociation rate constants, which more than compensates for the observed drop in k_{pol} . In conclusion, our data strongly support the concept that NC not only indirectly assists the reverse transcription process by its nucleic acid chaperoning activity but also positively affects the RT-catalyzed nucleotide incorporation reaction by increasing polymerase processivity presumably via a physical interaction of the two viral proteins.

The process of reverse transcription during the HIV-1 replication cycle is a complex interplay of different proteins and nucleic acid intermediates. Among these, the reverse transcriptase (RT) is first and foremost responsible for the conversion of the viral RNA into full-length viral DNA. Interestingly, the nucleocapsid protein (NC) chaperones this complex multistep process at the levels of primer tRNA annealing to the PBS (primer binding site) and of the two obligatory DNA strand transfers to generate the long terminal repeats (LTRs) (for reviews, see refs 1–4). NC is synthesized as part of the Gag precursor and subsequently cleaved by the viral protease into polypeptides of 55 and 72 amino acids. The exact biological function of these two forms is poorly understood, though the predominant form of NC found in mature HIV-1 particles mainly consists of 55 amino acids (5, 6 and references therein). NC is composed of two zinc fingers flanked by basic residues and is able to bind nucleic acids in a specific (e.g., stem loops SL2 and SL3 of the genomic Psi packaging signal) and nonspecific way. NMR studies of NC/SL–RNA complexes revealed that hydropho-

bic residues within the zinc fingers interact with the single-stranded loop region of the RNA, while the N-terminal basic domain of the protein undergoes electrostatic interactions with the double-stranded stem region of the RNA (7–9). NC assists the folding of nucleic acid molecules into their most stable conformation, a property termed nucleic acid chaperoning activity (10–12). The chaperone activity of NC is comprised of two properties: destabilization of nucleic acid interactions and the ability to promote annealing. The first one is specific, being dependent mainly on the folded fingers, while the second one is less specific and depends for the most part on positively charged residues (13–14). Due to its ability to destabilize and re-form base pairs, NC participates in a wide range of steps during replication, including the annealing of the tRNA^{Lys3} primer to the viral RNA (15–18), strand transfer reactions (19–23), and packaging of the viral genome (24–25). NC allows efficient polymerization on RT pausing sites (19, 26–28), regions within the viral RNA genome that expose secondary structure elements. Moreover, NC has also been shown to enhance the RNase H activity of RT and to protect the full-length viral DNA LTR ends (1, 29–31). However, for the most part, the molecular mechanisms of these binding and chaperoning activities are still poorly understood and to a certain extent quite puzzling. Besides its general nucleic acid binding and chaperoning activity, there are some hints that a physical interaction between RT and NC might play a role in facilitating reverse transcription (32–33).

Numerous detailed transient kinetic as well as structural studies have considerably improved our understanding of the complex interplay of RT with the different nucleic acid

[†] This work was supported in part by ANRS grants to J.-L.D. and Y.M. T.R. acknowledges funding by EC Grant LSHG-CT-2003-503480.

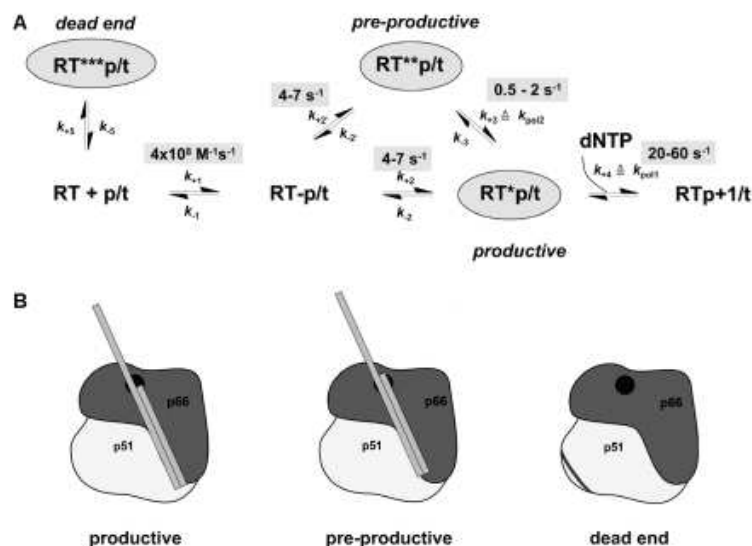
* To whom correspondence should be addressed. Telephone: +49-451-500-2745. Fax: +49-451-500-2729. E-mail: restle@imm.uni-luebeck.de.

[‡] Universität zu Lübeck.

[§] Present address: Institute of Structural and Molecular Biology (ISMB), Division of Biosciences, Faculty of Life Sciences, University College London, Darwin Building, Gower Street, London WC1E 6BT, United Kingdom.

^{||} Université Louis Pasteur.

[⊥] INSERM 758, IFR 128.

Scheme 1: (A) Minimal Kinetic Model of RT-p/t Interaction and Nucleotide Incorporation and (B) Illustrating Cartoon Based on Pre-Steady-State Kinetic and Single-Molecule Fluorescence Data (36, 37)^a

^a (A) Three distinct RT-primer/template complexes are formed: a productive complex that is capable of nucleotide incorporation, a preproductive complex which has to undergo a conformational change to incorporate nucleotides, and, additionally, a dead-end complex incapable of nucleotide incorporation. The latter can incorporate nucleotide only upon dissociation followed by re-association. The rate constants k_{-1} , k_1 , k_{-2} , k_2 , and k_{-3} for RT-p/t association were determined by stopped flow analysis (36). Additionally, monitoring of p/t binding via nucleotide incorporation, k_{-4} , k_4 , and k_{-5} , could be confirmed independently by performing quench flow experiments (36). k_{-3} and k_{-4} are the rate-limiting steps of nucleotide incorporation (termed k_{pol} and k_{pol}) of the preproductive and productive complex, respectively. At present, no firmly established data concerning an assignment of individual rate constants are available for the back reactions. One, two, and three asterisks distinguish structural states of the system that otherwise have the same composition. (B) The productive complex interacts with the p/t (gray bars) in a state closely resembling known X-ray structures (60, 61). In the preproductive complex, the primer terminus is shifted such that it now occupies the dNTP-binding pocket (dark circle in p66). For the dead-end complex, the dark line in p51 shows the area of contact of the 5'-terminus of the primer with the protein.

substrates encountered during reverse transcription and the subsequent polymerization reaction (for a comprehensive review, see ref 34). Previously, we showed that pre-steady-state single-turnover, single-nucleotide incorporation by HIV-1 RT occurs in a biphasic burst followed by a linear phase. This was eventually interpreted on the basis of three different RT-primer/template (p/t) substrate complexes, i.e., a productive enzyme-substrate complex which is capable of nucleotide incorporation and a preproductive complex which has to undergo an isomerization before dNTP incorporation can occur (35-37). The linear part represents the steady-state phase and is due to a catalytically incompetent dead end complex. In this case, dissociation of RT from the faulty bound nucleic acid substrate followed by proper rebinding is necessary to enable nucleotide incorporation (see Scheme 1 for an illustration). On the basis of these results, we investigated the effect of NC on transient RT-catalyzed single-turnover, single-nucleotide incorporation.

Here we present a further piece of a puzzle along the way to understanding how NC facilitates the process of HIV-1 reverse transcription. The data presented clearly show that NC extends the lifetime of RT-nucleic acid complexes, which in turn affects the processivity of the viral polymerase.

MATERIALS AND METHODS

Proteins. Recombinant, heterodimeric wild-type HIV-1_{BH10} RT was expressed in *Escherichia coli* cells and purified as described previously (38). The expression system and purification protocol allowed the preparation of large quanti-

ties of the heterodimeric enzyme in a homogeneous form. Furthermore, an RT mutant containing a single accessible cysteine at position 287 of the p66 subunit was used (39). This cysteine, situated within the thumb subdomain of p66, was labeled with the green donor fluorophore Alexa⁴⁸⁸-C5 maleimide (37). Highly pure NC protein [amino acids 1-55 (NC¹⁻⁵⁵)] was prepared by chemical synthesis as described by de Rocquigny and colleagues (40). Lyophilized protein powder (1-2 mg) was first dissolved in 200-400 μL of freshly degassed water at a pH value of ~ 3 to maintain the thiol groups in a reduced state followed by the addition of zinc sulfate (3 molar equiv) and adjustment of the pH to 7.5 with 50 mM Hepes buffer. Proteins were stored at -80°C . Protein concentrations were determined using extinction coefficients at 280 nm of $260450 \text{ M}^{-1} \text{ cm}^{-1}$ (HIV-1 RT) and $5500 \text{ M}^{-1} \text{ cm}^{-1}$ (NC¹⁻⁵⁵).

Buffers. Unless noted otherwise, all experiments were routinely carried out at 25°C in a buffer containing 50 mM Tris-HCl (pH 8.0), 50 mM KCl, and 10 mM MgCl_2 (standard RT buffer). Nucleic acid annealing buffer consisted of 20 mM Tris-HCl (pH 7.5) and 50 mM NaCl.

Synthetic Oligonucleotides. PAGE-purified oligodeoxynucleotides were purchased from IBA (Göttingen, Germany). Concentrations were determined by UV absorbance at 260 nm. The 18- and 36-mer DNA/DNA primer/template (p/t) (linear substrate) sequences were 5'-TCC CTG TTC GGG CGC CAC and 5'-TGT GGA AAA TCT CTA GAC GTG GCG CCC GAA CAG GGA, respectively. The complementary region is equivalent to the primer binding site (PBS) of

HIV-1. The 20- and 48-mer DNA/DNA p/t (hairpin substrate) sequences were 5'-TCT GCT CTG AAG AAA ATT CC and 5'-ATC TGG CCT TCC TAC AAA GGA AGG CCA GGG AAT TTT CTT CAG AGC AGA, respectively. The sequence is derived from the HIV-1 RNA genome and forms a secondary structure with a stem of 10 bp and a loop of six nucleotides (41). The abasic site substrate had the following sequence: template, 5'-AAA TCA XCC TAT CCT CCT TCA GGA CCA ACG TAC; and primer, 5'-CGT TGG TCC TGA AGG AGG ATA GGA. Before annealing, DNA primers were ^{32}P -labeled with T4 polynucleotide kinase at the 5' end as described previously (42).

FRET measurements were carried out with either a Cy5-labeled 19-mer (5'-TTG TCC CTG TTC GGG CGC C) or a 20-mer primer complementary to the 36- and 48-mer, respectively, described above. In both cases, the dye was covalently attached via a C6 amino link to the 5' end of the corresponding primer strand. Primer and template oligodeoxynucleotides were annealed by heating equimolar amounts in annealing buffer at 90 °C for 2 min, followed by cooling to room temperature over several hours in a heating block. To ensure complete annealing, the hybrids were analyzed by nondenaturing gel electrophoresis.

Gel Retardation Assay. Radioactively labeled substrate (100 nM) was incubated with NC at different molar ratios (see the corresponding Supporting Information figure legend) for 10 min at 25 °C. Next gel loading buffer (20% ficoll, 0.1% bromophenol blue, and 0.1% xylene cyanol) was added, and the samples were analyzed by nondenaturing gel electrophoresis (8% PAA) at 4 °C and subsequent phosphorimaging (Typhoon 8600, GE Healthcare).

Analysis of NC Chaperone Activity. The nucleic acid destabilizing activity of NC was characterized using a doubly labeled cTAR oligonucleotide (5'-GGT TCC TTG CTA GCC AGA GAG CTC CCG GGC TCG ACC TGG TCT AAC AAG AGA GAC C) with 6-carboxyrhodamine (Rh6G) attached to the 5' end and the quencher DABCYL to the 3' end (43). Experiments were performed in the standard buffer [50 mM Tris-HCl (pH 8.0), 50 mM KCl, and 10 mM MgCl₂] and in NC binding buffer [25 mM Tris-HCl (pH 7.5), 30 mM NaCl, and 200 μM MgCl₂] using a FluoroMax-3 spectrofluorometer (Horiba Jobin Yvon) with a thermostated cuvette holder at 25 °C in a 600 μL quartz cuvette. Excitation was performed at 480 nm with a 5 nm slit, and emission spectra were recorded from 500 to 700 nm with a 5 nm slit. cTAR (100 nM) was mixed with 1.1 μM NC and incubated for 5 min prior to the emission spectrum being recorded. The NC preparations were judged as active if the quotient of the fluorescence at the emission maximum (550 nm) in the presence and absence of NC was between 6.5 and 7.5 (Figure 1).

NC-mediated annealing of cTAR/dTAR was monitored as described previously (22). Briefly, the kinetic measurements were performed under pseudo-first-order conditions by using unlabeled dTAR at a concentration which was at least 10-fold higher than the concentration of cTAR labeled with carboxytetramethylrhodamine (TMR) at the 5' end and with 5 (and 6)-carboxyfluorescein (Fl) at the 3' end (44). Excitation and emission wavelengths were 480 and 520 nm, respectively, for monitoring the Fl fluorescence. NC was added to each oligonucleotide separately at a protein:nucleotide ratio of 2. Mixing the NC-coated oligonucleotides

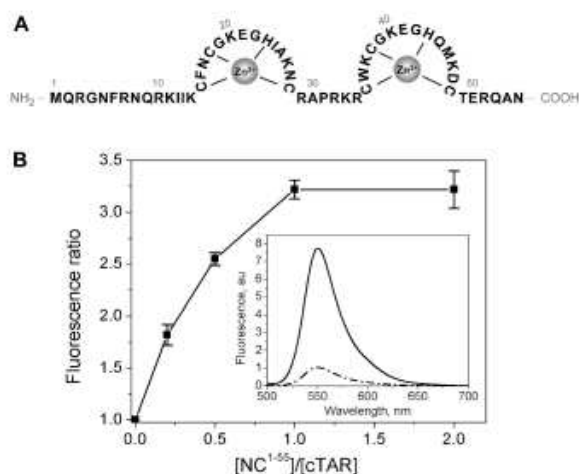


FIGURE 1: Structure and nucleic acid destabilization activity of HIV-1 nucleocapsid protein. (A) Amino acid sequence of HIV-1 NC¹⁻⁵⁵ and coordination of zinc ions within the two zinc fingers. (B) NC-induced destabilization of cTAR. The destabilization of 0.1 μM doubly labeled Rh6G-5'-cTAR-3'-DABCYL in the standard buffer [50 mM Tris-HCl (pH 8.0), 50 mM KCl, and 10 mM MgCl₂] was monitored at increasing NC:nucleotide ratios. Error bars are for three independent experiments. The inset shows emission spectra of 0.1 μM Rh6G-5'-cTAR-3'-DABCYL in the absence (---) and presence (—) of HIV-1 NC (1.1 μM). The spectra were recorded in 25 mM Tris-HCl (pH 7.5), 30 mM NaCl, and 200 μM MgCl₂. Excitation was at 480 nm.

started the annealing reaction. Experiments were performed in 50 mM Tris-HCl (pH 8.0) and 50 mM KCl with either 0.2 or 10 mM MgCl₂ at 20 °C. Emission spectra and kinetic traces were recorded with a Fluorolog spectrofluorometer (Jobin Yvon) equipped with a thermostated cell compartment. All fluorescence intensities were corrected for buffer emission and lamp fluctuations.

RT-Nucleic Acid Hairpin Substrate Binding. Binding affinities of RT-hairpin substrate complexes were determined by competitive fluorescence titrations (i.e., displacing a fluorescently labeled 5'-FAM-18/36-mer DNA/DNA p/t bound to RT) using a FluoroMax-3 spectrofluorometer. To monitor the fluorescence change upon displacement of the labeled p/t from RT, the samples were excited at 492 nm, and the emission intensity was measured at 516 nm. The competitive titration was evaluated using Scientist (Micro-Math Scientific Software), which allows the user to define the system under investigation as a series of parallel equations defining (in this case) each discrete equilibrium, the relationship between the total and free concentrations of the components, and the way in which the observable signal is generated. The K_d of the 18/36-mer DNA/DNA p/t was independently determined (45) and kept constant during the fit procedure.

Amplitude of Primer Extension after Single-Turnover, Single-Nucleotide Incorporation. HIV-1 RT was preincubated with radiolabeled p/t substrates for 5 min at 25 °C to allow complex formation. In the case of experiments where NC was present, it was added next and an additional incubation step of 5 min was performed to enable the formation of RT-substrate/NC complexes. The polymerization reaction was started by the addition of the next correct nucleotide to be incorporated (final concentration of 250 μM), and samples were incubated at 25 °C. To ensure single-

turnover conditions and prevent RT of multiple rounds of nucleotide incorporation, a large excess (25 μ M) of an anti-RT pseudoknot RNA aptamer (46, 47; see Figures 4 and 5 of the Supporting Information for further details) was added together with the nucleotide. Finally, reactions were stopped at different times in the second to minute range by addition of trifluoroacetic acid (TFA, 0.6% final concentration). The samples were mixed with formamide loading buffer, and the elongation of the primer was analyzed by denaturing PAGE. Data were quantified by phosphorimaging.

Rapid Kinetics of Nucleotide Incorporation. Rapid quench experiments were carried out in a chemical quench flow apparatus (RQF-3, KinTek Corp.). Reactions were started by rapidly mixing the reactants [RT–substrate complexes in the presence or absence of NC with nucleotide and RNA aptamer trap (46, 47), 15 μ L each] and then quenched with 0.6% TFA at defined time intervals in the millisecond to second range. All concentrations reported are final concentrations after mixing in the rapid quench apparatus. Products were analyzed by denaturing gel electrophoresis (10% polyacrylamide and 7 M urea in TBE buffer) and quantified by scanning the dried gel using a phosphor imager. The experimental data were fitted to the double-exponential equation [product] = $A[1 - \exp(-k_{\text{pol1}}t)] + B[1 - \exp(-k_{\text{pol2}}t)]$ using Grafit (Erithacus Software). A and B are the amplitudes of the biphasic burst, which reflects the concentration of catalytically competent p/t-bound enzyme at time zero. The effective pre-steady-state constants (k_{pol1} and k_{pol2}) are derived from the exponential rates.

Analysis of RT–Nucleic Acid Substrate Dissociation in the Presence of NC by Fluorescence Resonance Energy Transfer (FRET) Measurements. RT–nucleic acid substrate dissociation was followed using a FRET system described previously (37, 39; see Figure 6 of the Supporting Information for further details). Briefly, the p66 subunit of the HIV-1 RT was site specifically labeled with the donor fluorophore Alexa⁴⁸⁸ at position 287. The acceptor fluorophore Cy5 was introduced at the 5' end of a 19-mer (linear substrate) or 20-mer (hairpin substrate) primer annealed to the corresponding template oligonucleotide (see Synthetic Oligonucleotides). Formation of RT–substrate complexes results in fluorescence energy transfer from Alexa⁴⁸⁸ to Cy5.

All measurements were taken at 25 °C using a FluoroMax-3 spectrofluorometer and a 70 μ L fluorescence cuvette (Hellma). Excitation of the labeled RT was performed at 493 nm, and the FRET signal was recorded at 670 nm (in both cases with a 1 nm slit). As described above for the nucleotide incorporation experiments, preincubation of RT and substrate was followed by a second incubation step with NC. Next, the sample was rapidly mixed with a competitor solution containing the anti-HIV-1 RT RNA aptamer at 25 μ M (47), which prevents a rebinding of the primer/template motifs to RT. Time-resolved measurements of the change in the FRET signal after addition of the competitor solution allowed the calculation of dissociation rate constants using a double- or triple-exponential equation.

RESULTS

Nucleic Acid Binding and Chaperoning Properties of HIV-1 NC Protein. In this study, chemically synthesized nucleocapsid protein consisting of 55 (NC^{1–55}) amino acids

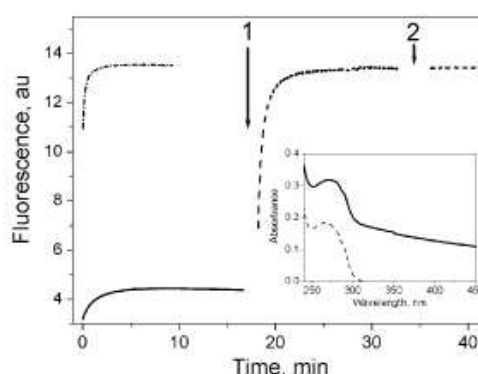


FIGURE 2: Effect of the Mg^{2+} concentration on NC-induced cTAR–dTAR annealing and aggregation (see Figure 2 of the Supporting Information for further details). The annealing kinetics of 0.1 μ M TMR-5'-cTAR-3'-F1 with 2 μ M dTAR in the presence of NC at a protein:nucleotide ratio of 2 were monitored in 50 mM Tris-HCl (pH 8.0), 50 mM KCl, and either 0.2 mM (—) or 10 mM MgCl_2 (---). At low Mg^{2+} concentrations, the fluorescence increase was limited. At the time indicated by arrow 1, the concentration of Mg^{2+} was increased to 10 mM, leading to a full restoration of the fluorescence emission. A further increase in the Mg^{2+} concentration to 20 mM at the time indicated by arrow 2 did not further affect the fluorescence emission. Excitation and emission wavelengths were 480 and 520 nm, respectively. The inset shows absorption spectra. The absorption spectrum recorded at 0.2 mM Mg^{2+} (—) shows a strong light scattering, indicative of aggregation. Addition of 10 mM Mg^{2+} fully suppresses the light scattering (---), suggesting that aggregates dissociated.

was used (Figure 1A). First, the NC preparations were tested in terms of chaperoning activities to control the functionality of the protein preparations. Experiments were initially performed in 25 mM Tris-HCl (pH 7.5), 30 mM NaCl, and 200 μ M MgCl_2 , where NC chaperone properties are optimal (15). A fluorescently labeled cTAR oligonucleotide which folds into a stem loop structure was used as described by Bernacchi and colleagues (43). Here the fluorescence of Rh6G attached to the 5' end of cTAR is quenched by 3' end-linked DABCYL. Upon addition of active NC, the stem loop structure is partly melted, resulting in a separation of the labeled termini and subsequently in an increase in fluorescence. An increase of a factor of 6.5–7.5-fold was seen (Figure 1B, inset), indicating chaperoning competent, fully active, and folded NC (13). Only active NC preparations were used in subsequent studies.

In a next step, the nucleic acid binding capability of NC in the standard buffer [50 mM Tris-HCl (pH 8.0), 50 mM KCl, and 10 mM MgCl_2] used for the RT assays described below was investigated using gel retardation electrophoresis. Radioactively labeled RT–nucleic acid substrates (termed linear or hairpin substrate; see Figure 3), which later were used for nucleotide incorporation experiments (see below), were mixed with increasing amounts of NC. In both cases, a concentration-dependent gel shift of the primer/template substrates, linear as well as hairpin substrate, was observed with complete retardation only at a NC:nucleotide molar ratio of 5:1 (Figure 1 of the Supporting Information). This high molar ratio suggests that the affinity of NC for nucleic acids is rather low in the standard buffer. This is likely due to the high Mg^{2+} concentration (10 mM), since an increase in the Mg^{2+} concentration from 1 to 10 mM was previously shown to decrease the affinity of NC for its target oligonucleotides by 2 orders of magnitude (48).

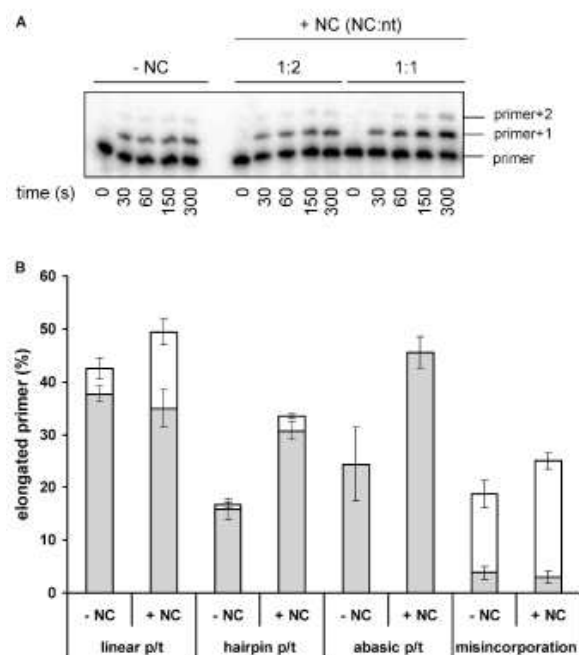


FIGURE 4: Effect of NC on the amplitude of RT-catalyzed single-turnover, single-nucleotide incorporation. For single-nucleotide incorporation experiments, 100 nM radiolabeled substrates [linear, hairpin, and abasic site substrate (see Figure 3)] were preincubated with HIV-1 RT (100 nM in the case of the linear substrate and 2 μ M in the case of the hairpin substrate) followed by a second preincubation step with HIV-1 NC. The polymerization reaction was started by addition of nucleotide (250 μ M dGTP, dCTP, and dATP in the cases of linear, hairpin, and abasic site substrate, respectively, and 2 mM dTTP in the case of misincorporation into the linear substrate) and 25 μ M anti-HIV-1 RT RNA aptamer as a competitor to prevent multiple-turnover events (see Figures 4 and 5 of the Supporting Information for further details). The elongation of the primer was analyzed by denaturing PAGE and quantified by phosphorimaging. In panel A, the polymerization products at different incubation times of RT–hairpin substrate complexes with different NC concentrations (1:2 molar ratio, 3.4 μ M NC; 1:1 molar ratio, 6.8 μ M NC) are shown. (B) Amplitudes of different primer extensions after incubation for 5 min. The relative distribution of primer+1 (gray bars) and primer+2 (white bars) products in the absence or presence of HIV-1 NC (NC:nucleotide ratio of 2:1) is shown. Details are given in the text. Results were obtained from three independent experiments each. The standard deviation is given by the error bars.

largely unaffected. RT-catalyzed incorporation of dCTP into the hairpin substrate in the presence of NC affected both amplitudes (primer+1/template nt G and primer+2/template nt A) which increased by 2- and 3-fold, respectively. In line with the dependence of NC chaperone activity on the NC:nucleotide ratio (Figure 1B), a ratio of 2:1 yielded the best results. Choosing higher or lower NC:nucleotide ratios led to a reduction of the observed effect (data not shown). Furthermore, we were also interested in whether NC is able to promote misincorporation at position primer+1 and to facilitate incorporation opposite abasic template sites. Misincorporation of dTTP opposite template nt C using the same linear substrate that was used previously was only slightly but significantly enhanced by a factor of 1.5 for primer+2 extension, which essentially reflects incorporation at position primer+1 as position+2 corresponds to correct incorporation and should therefore be much faster than the preceding step.

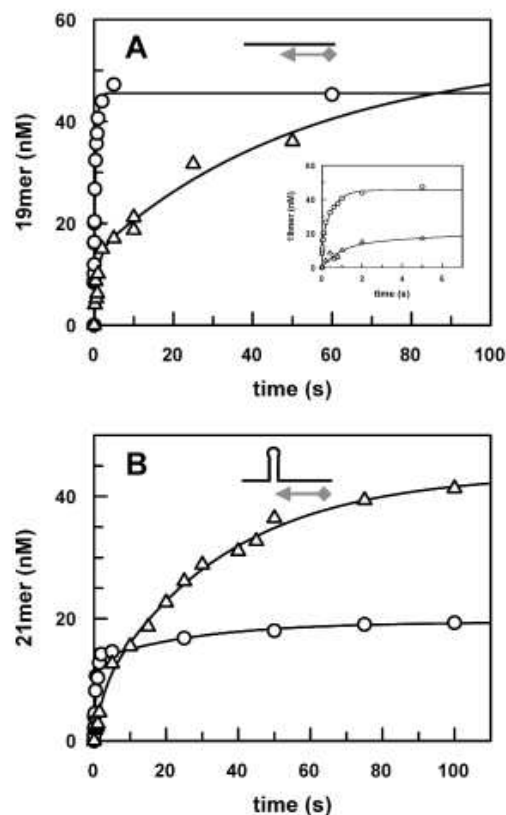


FIGURE 5: Rapid kinetics of single-turnover, single-nucleotide incorporation into linear or hairpin DNA/DNA p/t (see Figure 3) by HIV-1 RT. A preformed complex of 100 nM (A, linear substrate) or 2 μ M (B, hairpin substrate) RT and 100 nM p/t in the absence (O) or presence (Δ) of HIV-1 NC (NC:nucleotide ratio of 2:1) was rapidly mixed with either 250 μ M dGTP (A) or dCTP (B) and 25 μ M anti HIV-1 RT RNA aptamer (see Figures 4 and 5 of the Supporting Information for further details). The curves show the best fit of the data to a double-exponential equation yielding k_{pol1} and k_{pol2} values of 56 ± 12 and 1.9 ± 0.2 s $^{-1}$, and 2.3 ± 0.3 and 0.03 ± 0.01 s $^{-1}$ for the linear and hairpin substrate, respectively, in the absence of NC. In the presence of NC, we observed rates of 1.1 ± 0.3 and 0.02 ± 0.03 s $^{-1}$, and 0.4 ± 0.2 and 0.03 ± 0.003 s $^{-1}$, for the linear and hairpin substrate, respectively. The inset shows the reaction on a shorter time scale. Data are summarized in Table 1.

Incorporation of dATP in the presence of NC into a substrate with an abasic site at template position+1 resulted in an increase in the level of primer+1 extension of 1.9-fold, whereas no incorporation at position primer+2 could be observed.

Influence of NC on the Transient Polymerization Rate of HIV-1 RT. Next we analyzed the effect of NC on the pre-steady-state, single-nucleotide incorporation rate. Again, preformed RT–substrate complexes in the presence or absence of NC were rapidly mixed with the incoming nucleotide together with the aptamer trap, and the polymerase reaction was monitored in the millisecond to second range using a quench flow apparatus (Figure 5). As briefly outlined in the introductory section and previously described in detail (36, 37), single-turnover, single-nucleotide incorporation by HIV-1 RT occurs under such experimental conditions with a biphasic exponential burst where the sum of the amplitudes of the burst is equivalent to the amount of

Table 1: Summary of the Effects of NC on the HIV-1 RT-Catalyzed Rate of Single-Nucleotide Incorporation and RT-p/t Dissociation^a

| substrate | NC | k_{pol1} (s ⁻¹) | k_{pol2} (s ⁻¹) | k_{diss1} (s ⁻¹) | k_{diss2} (s ⁻¹) | k_{diss3} (s ⁻¹) |
|-----------|----|--------------------------------------|--------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| linear | - | 56 ± 12 (15) | 1.9 ± 0.2 (30) | 3 ± 0.14 (67) | 0.24 ± 0.006 (24) | - |
| linear | + | 1.1 ± 0.3 (14) | 0.02 ± 0.03 (40) | 0.04 ± 0.0006 (46) | 0.006 ± 0.0001 (30) | - |
| hairpin | - | 2.3 ± 0.3 (13) | 0.03 ± 0.01 (6) | 2.4 ± 0.08 (62) | 0.5 ± 0.02 (20) | 0.04 ± 0.0005 (12) |
| hairpin | + | 0.4 ± 0.2 (7) | 0.03 ± 0.003 (37) | 0.17 ± 0.002 (87) | 0.02 ± 0.002 (6) | - |

^aThe numbers in parentheses show the distribution of the different amplitudes (given in nanomolar for k_{pol} and in arbitrary fluorescence units for k_{diss}). The corresponding experiments are shown in Figures 5 and 6. Refer to Scheme 1 for an overview of how the numbers given above are related to the minimal kinetic model of RT-p/t interaction and nucleotide incorporation. As outlined in the text, due to the complexity of the system, it is currently not possible to precisely assign k_{diss1} , k_{diss2} , and k_{diss3} to individual steps for the backward reaction given in Scheme 1 (i.e., k_{-1} , k_{-2} , k_{-3} , and k_{-5}).

catalytically competent p/t-bound enzyme at the start of the reaction and the rates of nucleotide incorporation are given by the exponential terms (Figure 5 of the Supporting Information). With the linear substrate, we observed incorporation rates (k_{pol1} and k_{pol2}) of 56 ± 12 and 1.9 ± 0.2 s⁻¹, respectively, highly consistent with previous findings (35, 36, 50). On the other hand, the hairpin substrate exhibited substantially slower incorporation rates of 2.3 ± 0.3 and 0.03 ± 0.01 s⁻¹, respectively, consistent with a study by Suo and Johnson (41). Upon addition of NC, the pre-steady-state nucleotide incorporation rates in the case of the linear substrate dropped dramatically to 1.1 ± 0.3 and 0.02 ± 0.03 s⁻¹, respectively, while for incorporation into the hairpin substrate, only the faster rate was reduced by a factor of ~6 with the slower rate being unaffected. Moreover, for the linear substrate, the amplitude of nucleotide incorporation (sum of the two burst phases) remained largely unaffected by NC, whereas we observed a more than 2-fold increase in the case of the hairpin sequence, consistent with the findings described above for the primer extension studies. Data are summarized in Table 1.

Effect of NC on the Stability of the RT-Substrate Complex.

In the preceding sections, we could show that NC causes an increase in the amplitude of primer extension and at the same time a reduction of the rate of nucleotide incorporation. A rational explanation for this finding could be a NC-mediated increased stability of RT-nucleic acid complexes. For this reason, we investigated RT-nucleic acid complex dissociation. Analyzing a given protein-nucleic acid interaction in the context of a second nucleic acid binding protein is a challenging task since a differentiation of the two proteins in terms of substrate binding is technically difficult. To resolve this problem, we applied a FRET-based system enabling us to exclusively observe RT-nucleic acid interactions (Figure 6 of the Supporting Information). Here, the FRET signal of Cy5, a fluorophore attached to the 5' end of the primer molecule, is a direct measure of the number of RT-substrate complexes as the transfer of energy from the donor fluorophore Alexa⁴⁸⁸ attached to the thumb domain of RT is only possible within a RT-nucleic acid complex. Preformed RT-nucleic acid substrate complexes (linear or hairpin substrate) in the presence or absence of NC were rapidly mixed with an excess of competitor (Figure 6; see Table 1 for a summary). Analysis of the dissociation process of the RT-linear substrate complex revealed two phases with dissociation rate constants (k_{diss1} and k_{diss2}) of 3 ± 0.14 and 0.24 ± 0.006 s⁻¹, respectively. The presence of NC considerably decelerated dissociation of the complex. Here dissociation rate constants of 0.04 ± 0.0006 and 0.006 ± 0.0001 s⁻¹ were observed. Similar data were determined with

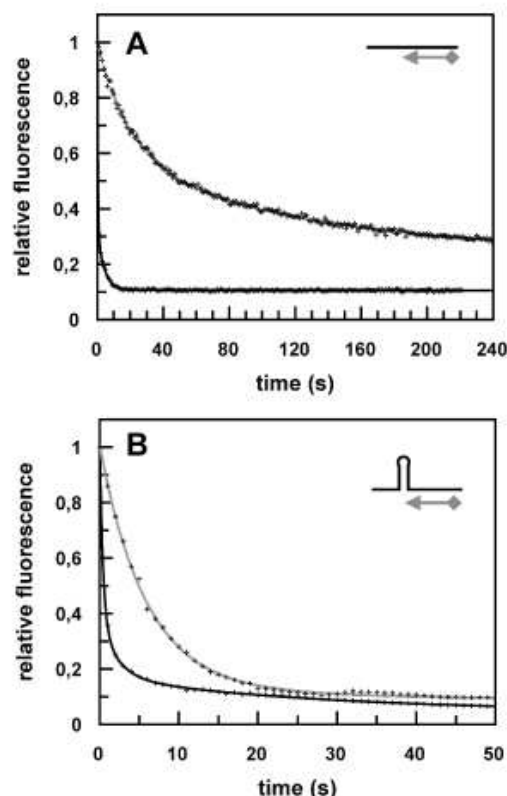


FIGURE 6: FRET measurements of RT-p/t dissociation kinetics. (A) Linear 5'-Cy5-p/t (100 nM) was preincubated with 100 nM HIV-1 RTA⁴⁸⁸ (see Figure 6 of the Supporting Information for a graphical illustration) and rapidly mixed with 25 μM aptamer competitor in the absence (black curve) or presence (gray curve) of NC (10.8 μM). (B) Dissociation of RT (2 μM) from 100 nM hairpin p/t after being mixed with 25 μM RNA aptamer in the absence (black curve) or presence (gray curve) of NC (13.6 μM). Data were fitted to a double- or triple-exponential equation. The corresponding rates and amplitudes are given in Table 1.

the hairpin substrate. However, in contrast to the linear substrate data, these data best fit an equation with three exponentials giving rates of 2.4 ± 0.08 , 0.5 ± 0.02 , and 0.04 ± 0.0005 s⁻¹ in the absence of NC and 0.17 ± 0.002 and 0.02 ± 0.002 s⁻¹ in the presence of NC. In the latter case, data were fitted to a double-exponential equation. Appropriate control experiments showed that the Alexa⁴⁸⁸ label at position 287 did not significantly alter the kinetic parameters of RT (refs 51 and 52 and unpublished data of T. Restle).

DISCUSSION

Over the years, many studies provided evidence that NC assists tRNA annealing, reverse transcription initiation,

minus-strand transfer, processivity of reverse transcription, plus-strand transfer, strand displacement synthesis, and moreover 3' processing of viral DNA by integrase, and integrase-mediated strand transfer (for recent reviews, see refs 3 and 4). It seems rather peculiar that such a small protein can affect so many seemingly different processes. Interestingly, all these essential viral reactions involve protein–nucleic acid interactions, and NC may act as a “macromolecular crowding” factor to enhance the stability and activity of such RNA and DNA replicating complexes (1, 3, 4, 53). However, in most cases, the molecular details of this facilitator activity remain poorly understood. The aim of this study was to explore possible effects of HIV-1 NC on RT-catalyzed single-turnover, single-nucleotide incorporation applying pre-steady-state techniques and to unravel some of the underlying mechanistic characteristics.

In a first set of kinetic experiments, we analyzed the effect of NC on the amplitude of single-nucleotide primer extension under strict single-turnover conditions. Depending on the nucleic acid substrate (linear, hairpin, or abasic site DNA/DNA primer/template, i.e., p/t), an increase in the level of extended primers by a factor of 1.5–3 was observed with the distribution of primers elongated by one or two nucleotides varying considerably. It should be noted that extension of the primers by more than one nucleotide reflects misincorporation as the experimental setup as well as the sequences was designed to allow the incorporation of one single Watson–Crick base-paired nucleotide. In all but one case, NC considerably affected the apparent nucleotide incorporation fidelity of RT. Interestingly, with a substrate containing an abasic site at template position +1, we observed only primer+1 extension, albeit at a level increased by a factor of ~1.9 in the presence of NC, which would indicate that the orientation of the 3' end of the primer in the active site of the polymerase is highly unfavorable to allow for the incorporation of a second mismatched nucleotide. As a first approximation, one would have expected an effect of NC in the case of the hairpin substrate due to a destabilization of its stable stem structure caused by the NC chaperoning activity. However, our results with the linear p/t substrate clearly show the stimulating effect of NC is not restricted to those kinds of substrates.

Maximal stimulation of primer extension was detected at a NC:nucleotide molar ratio of 2:1. As in the virion, the viral RNA is coated with NC proteins at an approximate ratio of 1:7 (48, 54, 55), and the concentration of NC necessary to stimulate the polymerization reaction and to achieve full retardation of the p/t substrates in the gel shift assay (Figure 1 of the Supporting Information) was rather high. This finding most likely was due to the MgCl₂ concentration of 10 mM that was used. High concentrations (i.e., >1 mM) of this divalent cation have been shown to considerably reduce the level of NC binding to nucleic acids (48, 56). Along these lines, Anthony et al. described similar observations, where the production of long-range DNAs was stimulated only at high NC concentrations (57). These results are in excellent agreement with NC chaperoning assays performed at 10 mM MgCl₂ shown in Figures 1B and 2. For NC-mediated nucleic acid destabilization as well as annealing under high-MgCl₂ conditions, we also had to use a protein:nucleotide ratio of ~2:1 to reach maximal activity. In contrast, at low MgCl₂ concentrations (i.e., 0.2 mM), full activity has already been

observed at a NC:nucleotide ratio of 0.2:1. Now one could argue that the nucleotide incorporation experiments should be performed under more physiologically relevant concentrations (i.e., <1 mM). To ensure that the single-turnover rate of nucleotide incorporation observed is limited by internal rate-limiting kinetic parameters, rather than by binding parameters for p/t, dNTP, and MgCl₂, which occurs when concentrations below the saturation level are used, we were forced to use such high concentrations of MgCl₂. Moreover, in all previously conducted experiments regarding single-turnover nucleotide incorporation, we used a concentration of this divalent cation of 10 mM. Accordingly, we would not have been able to relate the data of this study with the comprehensive data sets derived before concerning a kinetic model of the complex RT-catalyzed nucleotide incorporation reaction (35–37). For this matter, it was essential for this study to show that NC is active even at 10 mM MgCl₂, albeit at a much higher protein:nucleotide ratio, but does not induce any aggregation of the nucleic acid substrate (Figure 2).

Next we analyzed the transient rate (k_{pol}) of single-nucleotide incorporation in the presence of NC. As briefly outlined in the introductory section, incorporation of a nucleotide by HIV-1 RT is a complicated process, which involves at least three conformationally distinct enzyme–nucleic acid complexes (Scheme 1). As a consequence, one observes a biphasic burst of nucleotide incorporation with an amplitude of considerably less than 100% due to the formation of dead-end complexes, although under the chosen experimental conditions one theoretically would expect a single burst of product formation with an amplitude corresponding to the amount of active enzyme present at the start of the reaction (36). Our studies revealed that the presence of NC decreased the rate of fast pre-steady-state nucleotide incorporation (k_{pol1}) of RT by up to 50-fold (linear substrate) and was most likely due to NC molecules, which completely cover the substrate and in this way hinder translocation of the enzyme along the nucleic acid. The decrease of the second slow phase (k_{pol2}) by a factor of ~90 is even more pronounced. This nicely fits the model of two catalytically competent RT–p/t complexes, as the second complex has to translocate from a product to an educt state before incorporation can occur (i.e., the educt state is a complex in which the nucleic acid is positioned to allow nucleotide incorporation, whereas the product state is formed immediately after nucleotide incorporation but before RT translates to the next nucleotide). In the case of the hairpin substrate, the situation is different. Here two effects of NC might outweigh each other, so in the end k_{pol1} is decreased by only ~6-fold. It is likely that the hampering effect of NC is compensated by a destabilization of the hairpin structure, which in turn would accelerate nucleotide incorporation. The second slow phase is already so slow that it remained unaffected by NC. Another interesting aspect is the observed shift in the relative distribution of the two burst phases. While there is only a slight reduction in the amplitude of the fast phase from 33 to 26% (amplitude of the fast phase relative to the sum of two amplitudes) in the case of the linear substrate, the hairpin p/t shows a dramatic reduction in the percentage of this phase from 68 to 15% (Table 1). This indicates that NC might have an effect on the relative distribution of the two different catalytically competent RT–p/t complexes [i.e., preproductive and productive enzyme–substrate complexes (Scheme 1)].

Two opposing effects are caused by NC: an increase in the amplitude of primer extension on one hand and a drastically reduced transient rate of nucleotide incorporation on the other. How does this fit together? A rather simple explanation would be an enhanced stability of the enzyme–substrate complexes, which would compensate for the reduced rate of a catalytic or precatalytic step and consequently increased processivity. To address this question, we performed experiments to analyze the RT–nucleic acid substrate dissociation process (Table 1). For this purpose, we applied an established FRET-based system (37, 39), which enabled us to exclusively observe RT–p/t complexes. For the dissociation of RT from a DNA/DNA p/t substrate, we detected two or three phases with distinct rates of $k_{\text{diss}1}$, $k_{\text{diss}2}$, and $k_{\text{diss}3}$ consistent with previous work (ref 36 and unpublished data of T. Restle). Addition of NC to a preformed RT–linear p/t complex reduced the dissociation rate constants ($k_{\text{diss}1}$ and $k_{\text{diss}2}$) of the two observed steps by ~ 75 - and ~ 40 -fold, respectively. For the hairpin substrate, we could resolve three phases. Again, we observed considerable reductions of the two fast rates in the presence of NC of approximately 14- and 25-fold, respectively. Deriving precise numbers concerning the degree of deceleration of complex dissociation is not trivial as an exact assignment of the individual phases observed to individual steps (i.e., RT–p/t complexes; also see Scheme 1) is very difficult. However, what can be stated with certainty is that NC substantially increased the stability of RT–nucleic acid complexes, which more than compensates for the reduction of k_{pol} and thus overall leads to an increase in processivity. Recently, Bampi et al. described an analogous observation (58). Here in the presence of NC^{1–71} an approximately 16-fold reduction in the apparent dissociation rate constant (from 0.059 to 0.0036 s⁻¹) was determined indirectly by performing steady-state single-nucleotide incorporation measurements. Moreover, they report a more efficient extension of mutated cDNAs with a higher rate of corrected mutations in the presence of NC. Essentially, these observations fit very nicely with our results.

With respect to the underlying molecular mechanism of the NC-mediated stabilization of RT–p/t complexes, we currently can only speculate. One likely scenario would be a direct physical interaction of the two viral proteins (32). According to such a model, NC might hold RT onto the nucleic acid substrate. Attempts to map such potential sites of interaction on the polymerase with the help of 23 anti-RT monoclonal antibodies (59) (i.e., testing their potential to block the effect of NC on RT–substrate dissociation) have so far never been successful.

In conclusion, our data further contribute to the concept that NC assists RT during viral replication. The observed NC-mediated increase in the lifetime of RT–nucleic acid complexes enhances processivity overall but possibly even more importantly facilitates the overcoming of exceptional situations (e.g., nucleic acid lesions or particular secondary structures) which ultimately ensures synthesis of the full-length viral DNA flanked by the LTR. Furthermore, the proposed physical interplay of the two proteins might offer a promising target for the development of novel antiviral compounds.

ACKNOWLEDGMENT

We thank Hugues de Rocquigny for synthesis of the NC proteins and Francesca Di Pasquale for the abasic site substrate.

SUPPORTING INFORMATION AVAILABLE

Nucleic acid binding properties of HIV-1 NC, emission spectra of TMR-5'-cTAR-3'-Fl, displacement titration of fluorescent p/t bound to RT with the hairpin substrate, structure of the HIV-1 RT–RNA aptamer complex, trap experiments with the RNA aptamer, and the FRET-based system used for RT–nucleic acid complex dissociation experiments. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Bampi, C., Jacquenet, S., Lener, D., Decimo, D., and Darlix, J. L. (2004) The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. *Int. J. Biochem. Cell Biol.* 36, 1668–1686.
- Levin, J. G., Guo, J., Rouzina, I., and Musier-Forsyth, K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* 80, 217–286.
- Darlix, J. L., Garrido, J. L., Morellet, N., Mely, Y., and de Rocquigny, H. (2007) Properties, functions, and drug targeting of the multifunctional nucleocapsid protein of the human immunodeficiency virus. *Adv. Pharmacol.* 55, 299–346.
- Thomas, J. A., and Gorelick, R. J. (2008) Nucleocapsid protein function in early infection processes. *Virus Res.* 134, 39–63.
- Chertova, E., Chertov, O., Coren, L. V., Roser, J. D., Trubey, C. M., Bess, J. W., Jr., Sowder, R. C., Barsov, E., Hood, B. L., Fisher, R. J., Nagashima, K., Conrads, T. P., Veenstra, T. D., Lifson, J. D., and Ott, D. E. (2006) Proteomic and biochemical analysis of purified human immunodeficiency virus type 1 produced from infected monocyte-derived macrophages. *J. Virol.* 80, 9039–9052.
- Coren, L. V., Thomas, J. A., Chertova, E., Sowder, R. C., Gagliardi, T. D., Gorelick, R. J., and Ott, D. E. (2007) Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. *J. Virol.* 81, 10047–10054.
- Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M. C., and Roques, B. P. (1998) Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J. Mol. Biol.* 283, 419–434.
- De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* 279, 384–388.
- Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K. J., Wu, Z. R., and Summers, M. F. (2000) NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* 301, 491–511.
- Tsuchihashi, Z., and Brown, P. O. (1994) DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* 68, 5863–5870.
- Darlix, J. L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol.* 254, 523–537.
- Herschlag, D. (1995) RNA chaperones and the RNA folding problem. *J. Biol. Chem.* 270, 20871–20874.
- Beltz, H., Clauss, C., Piemont, E., Ficheux, D., Gorelick, R. J., Roques, B., Gabus, C., Darlix, J. L., de Rocquigny, H., and Mely, Y. (2005) Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J. Mol. Biol.* 348, 1113–1126.
- Williams, M. C., Rouzina, I., Wenner, J. R., Gorelick, R. J., Musier-Forsyth, K., and Bloomfield, V. A. (2001) Mechanism for nucleic

- acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6121–6126.
15. Lapadat-Tapolksy, M., Pernelle, C., Borie, C., and Darlix, J. L. (1995) Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1. *Nucleic Acids Res.* 23, 2434–2441.
 16. Li, X., Quan, Y., Arts, E. J., Li, Z., Preston, B. D., de Rocquigny, H., Roques, B. P., Darlix, J. L., Kleiman, L., Parniak, M. A., and Wainberg, M. A. (1996) Human immunodeficiency virus Type 1 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis primed by human tRNA(Lys3) in vitro: Studies of viral RNA molecules mutated in regions that flank the primer binding site. *J. Virol.* 70, 4996–5004.
 17. Feng, Y. X., Campbell, S., Harvin, D., Ehresmann, B., Ehresmann, C., and Rein, A. (1999) The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: Possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J. Virol.* 73, 4251–4256.
 18. Hargittai, M. R., Gorelick, R. J., Rouzina, L., and Musier-Forsyth, K. (2004) Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J. Mol. Biol.* 337, 951–968.
 19. Wu, T., Guo, J., Bess, J., Henderson, L. E., and Levin, J. G. (1999) Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer: Analysis in reconstituted and endogenous reverse transcription systems. *J. Virol.* 73, 4794–4805.
 20. Johnson, P. E., Turner, R. B., Wu, Z. R., Hairston, L., Guo, J., Levin, J. G., and Summers, M. F. (2000) A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry* 39, 9084–9091.
 21. Muthuswami, R., Chen, J., Burnett, B. P., Thimmig, R. L., Janjic, N., and McHenry, C. S. (2002) The HIV plus-strand transfer reaction: Determination of replication-competent intermediates and identification of a novel lentiviral element, the primer over-extension sequence. *J. Mol. Biol.* 315, 311–323.
 22. Godet, J., de Rocquigny, H., Raja, C., Glasser, N., Ficheux, D., Darlix, J. L., and Mely, Y. (2006) During the early phase of HIV-1 DNA synthesis, nucleocapsid protein directs hybridization of the TAR complementary sequences via the ends of their double-stranded stem. *J. Mol. Biol.* 356, 1180–1192.
 23. Ramalanjaona, N., de Rocquigny, H., Millet, A., Ficheux, D., Darlix, J. L., and Mely, Y. (2007) Investigating the mechanism of the nucleocapsid protein chaperoning of the second strand transfer during HIV-1 DNA synthesis. *J. Mol. Biol.* 374, 1041–1053.
 24. Berkowitz, R., Fisher, J., and Goff, S. P. (1996) RNA packaging. *Curr. Top. Microbiol. Immunol.* 214, 177–218.
 25. D'Souza, V., and Summers, M. F. (2005) How retroviruses select their genomes. *Nat. Rev. Microbiol.* 3, 643–655.
 26. Ji, X., Klarmann, G. J., and Preston, B. D. (1996) Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro. *Biochemistry* 35, 132–143.
 27. Drummond, J. E., Mounts, P., Gorelick, R. J., Casas-Finet, J. R., Bosche, W. J., Henderson, L. E., Waters, D. J., and Arthur, L. O. (1997) Wild-type and mutant HIV type 1 nucleocapsid proteins increase the proportion of long cDNA transcripts by viral reverse transcriptase. *AIDS Res. Hum. Retroviruses* 13, 533–543.
 28. Klasens, B. L., Huthoff, H. T., Das, A. T., Jeeninga, R. E., and Berkhout, B. (1999) The effect of template RNA structure on elongation by HIV-1 reverse transcriptase. *Biochim. Biophys. Acta* 1444, 355–370.
 29. Peliska, J. A., Balasubramanian, S., Giedroc, D. P., and Benkovic, S. J. (1994) Recombinant HIV-1 nucleocapsid protein accelerates HIV-1 reverse transcriptase catalyzed DNA strand transfer reactions and modulates RNase H activity. *Biochemistry* 33, 13817–13823.
 30. Chen, Y., Balakrishnan, M., Roques, B. P., and Bambara, R. A. (2003) Steps of the acceptor invasion mechanism for HIV-1 minus strand strong stop transfer. *J. Biol. Chem.* 278, 38368–38375.
 31. Roda, R. H., Balakrishnan, M., Hanson, M. N., Wöhrl, B. M., Le Grice, S. F., Roques, B. P., Gorelick, R. J., and Bambara, R. A. (2003) Role of the Reverse Transcriptase, Nucleocapsid Protein, and Template Structure in the Two-step Transfer Mechanism in Retroviral Recombination. *J. Biol. Chem.* 278, 31536–31546.
 32. Lener, D., Tanchou, V., Roques, B. P., Le Grice, S. F., and Darlix, J. L. (1998) Involvement of HIV-1 nucleocapsid protein in the recruitment of reverse transcriptase into nucleoprotein complexes formed in vitro. *J. Biol. Chem.* 273, 33781–33786.
 33. Druillennec, S., Caneparo, A., de Rocquigny, H., and Roques, B. P. (1999) Evidence of interactions between the nucleocapsid protein NCp7 and the reverse transcriptase of HIV-1. *J. Biol. Chem.* 274, 11283–11288.
 34. Furman, P. A., Painter, G. R., and Anderson, K. S. (2000) An analysis of the catalytic cycle of HIV-1 reverse transcriptase: Opportunities for chemotherapeutic intervention based on enzyme inhibition. *Curr. Pharm. Des.* 6, 547–567.
 35. Thrall, S. H., Krebs, R., Wöhrl, B. M., Cellai, L., Goody, R. S., and Restle, T. (1998) Pre-steady-state kinetic characterization of RNA-primed initiation of transcription by HIV-1 reverse transcriptase and analysis of the transition to a processive DNA-primed polymerization mode. *Biochemistry* 37, 13349–13358.
 36. Wöhrl, B. M., Krebs, R., Goody, R. S., and Restle, T. (1999) Refined model for primer/template binding by HIV-1 reverse transcriptase: Pre-steady-state kinetic analyses of primer/template binding and nucleotide incorporation events distinguish between different binding modes depending on the nature of the nucleic acid substrate. *J. Mol. Biol.* 292, 333–344.
 37. Rothwell, P. J., Berger, S., Kensch, O., Felekyan, S., Antonik, M., Wöhrl, B. M., Restle, T., Goody, R. S., and Seidel, C. A. (2003) Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase: primer/template complexes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 16655–16660.
 38. Müller, B., Restle, T., Weiss, S., Gautel, M., Sczakiel, G., and Goody, R. S. (1989) Co-expression of the subunits of the heterodimer of HIV-1 reverse transcriptase in *Escherichia coli*. *J. Biol. Chem.* 264, 13975–13978.
 39. Kensch, O., Restle, T., Wöhrl, B. M., Goody, R. S., and Steinhoff, H. J. (2000) Temperature-dependent equilibrium between the open and closed conformation of the p66 subunit of HIV-1 reverse transcriptase revealed by site-directed spin labelling. *J. Mol. Biol.* 301, 1029–1039.
 40. de Rocquigny, H., Ficheux, D., Gabus, C., Fournie-Zaluski, M. C., Darlix, J. L., and Roques, B. P. (1991) First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem. Biophys. Res. Commun.* 180, 1010–1018.
 41. Suo, Z., and Johnson, K. A. (1998) DNA secondary structure effects on DNA synthesis catalyzed by HIV-1 reverse transcriptase. *J. Biol. Chem.* 273, 27259–27267.
 42. Krebs, R., Immendörfer, U., Thrall, S. H., Wöhrl, B. M., and Goody, R. S. (1997) Single-step kinetics of HIV-1 reverse transcriptase mutants responsible for virus resistance to nucleoside inhibitors zidovudine and 3-TC. *Biochemistry* 36, 10292–10300.
 43. Bernacchi, S., Stoylov, S., Piemont, E., Ficheux, D., Roques, B. P., Darlix, J. L., and Mely, Y. (2002) HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J. Mol. Biol.* 317, 385–399.
 44. Beltz, H., Azoulay, J., Bernacchi, S., Clamme, J. P., Ficheux, D., Roques, B., Darlix, J. L., and Mely, Y. (2003) Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. *J. Mol. Biol.* 328, 95–108.
 45. Divita, G., Müller, B., Immendörfer, U., Gautel, M., Rittinger, K., Restle, T., and Goody, R. S. (1993) Kinetics of interaction of HIV reverse transcriptase with primer/template. *Biochemistry* 32, 7966–7971.
 46. Jaeger, J., Restle, T., and Steitz, T. A. (1998) The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *EMBO J.* 17, 4535–4542.
 47. Kensch, O., Connolly, B. A., Steinhoff, H. J., McGregor, A., Goody, R. S., and Restle, T. (2000) HIV-1 reverse transcriptase-pseudoknot RNA aptamer interaction has a binding affinity in the low picomolar range coupled with high specificity. *J. Biol. Chem.* 275, 18271–18278.
 48. Mely, Y., de Rocquigny, H., Sorinas-Jimeno, M., Keith, G., Roques, B. P., Marquet, R., and Gerard, D. (1995) Binding of the HIV-1 nucleocapsid protein to the primer (tRNA(Lys)), in vitro, is essentially not specific. *J. Biol. Chem.* 270, 1650–1656.
 49. Bernacchi, S., Piemont, E., Potier, N., Dorsselaer, A., and Mely, Y. (2003) Excitonic heterodimer formation in an HIV-1 oligonucleotide labeled with a donor-acceptor pair used for fluorescence resonance energy transfer. *Biophys. J.* 84, 643–654.
 50. Cramer, J., Strerath, M., Marx, A., and Restle, T. (2002) Exploring the effects of active site constraints on HIV-1 reverse transcriptase DNA polymerase fidelity. *J. Biol. Chem.* 277, 43593–43598.

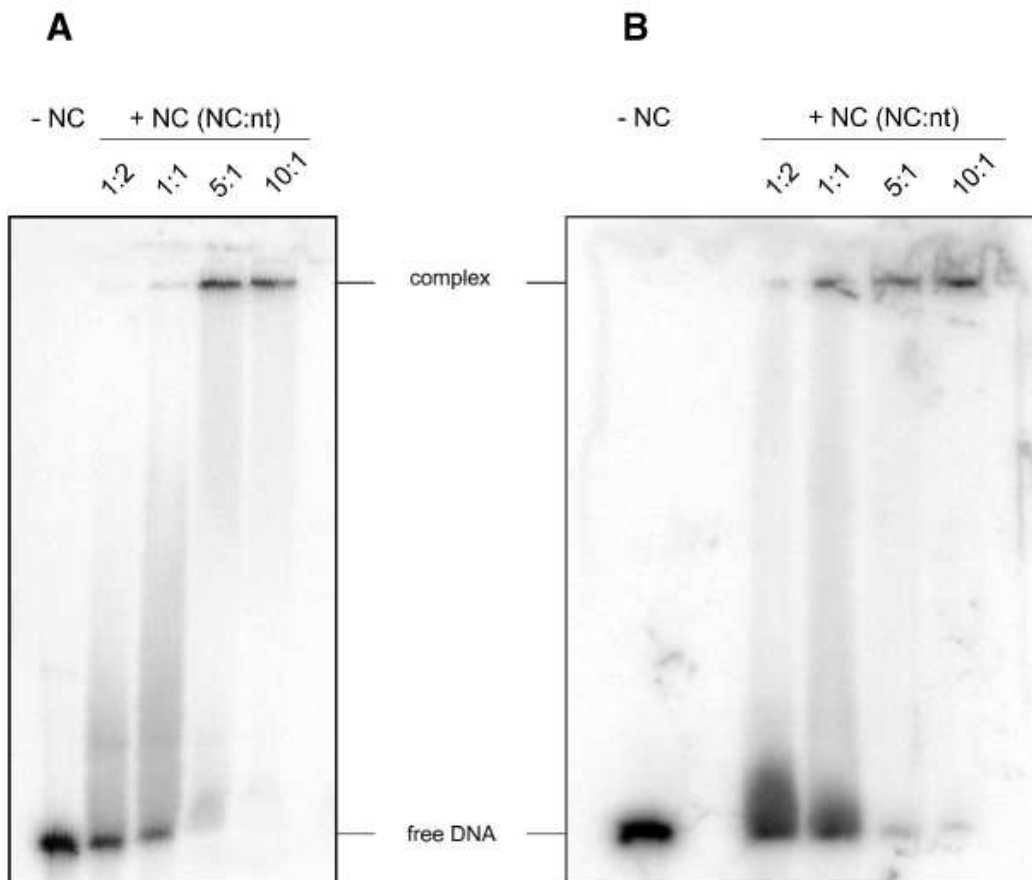
51. Kensch, O. (2000) Untersuchungen zur Konformation und Dynamik der Reversen Transkriptase von HIV-1 durch ESR- und Einzelmolekülfluoreszenzspektroskopie, Thesis, University of Dortmund, Dortmund, Germany.
52. Rothwell, P. J. (2002) Structural Investigations on HIV-1 RT using single pair Fluorescence Resonance Energy Transfer, Thesis, University of Dortmund, Dortmund, Germany.
53. Jarvis, T. C., Ring, D. M., Daube, S. S., and von Hippel, P. H. (1990) "Macromolecular crowding": Thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex. *J. Biol. Chem.* **265**, 15160–15167.
54. Karpel, R. L., Henderson, L. E., and Oroszlan, S. (1987) Interactions of retroviral structural proteins with single-stranded nucleic acids. *J. Biol. Chem.* **262**, 4961–4967.
55. You, J. C., and McHenry, C. S. (1993) HIV nucleocapsid protein. Expression in *Escherichia coli*, purification, and characterization. *J. Biol. Chem.* **268**, 16519–16527.
56. Wu, T., Heilman-Miller, S. L., and Levin, J. G. (2007) Effects of nucleic acid local structure and magnesium ions on minus-strand transfer mediated by the nucleic acid chaperone activity of HIV-1 nucleocapsid protein. *Nucleic Acids Res.* **35**, 3974–3987.
57. Anthony, R. M., and DeStefano, J. J. (2007) In vitro Synthesis of Long DNA Products in Reactions with HIV-RT and Nucleocapsid Protein. *J. Mol. Biol.* **365**, 310–324.
58. Bampi, C., Bibillo, A., Wendeler, M., Divita, G., Gorelick, R. J., Le Grice, S. F., and Darlix, J. L. (2006) Nucleotide excision repair and template-independent addition by HIV-1 reverse transcriptase in the presence of nucleocapsid protein. *J. Biol. Chem.* **281**, 11736–11743.
59. Restle, T., Pawlita, M., Sczakiel, G., Müller, B., and Goody, R. S. (1992) Structure-function relationships of HIV-1 reverse transcriptase determined using monoclonal antibodies. *J. Biol. Chem.* **267**, 14654–14661.
60. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320–6324.
61. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. *Science* **282**, 1669–1675.

BI801386R

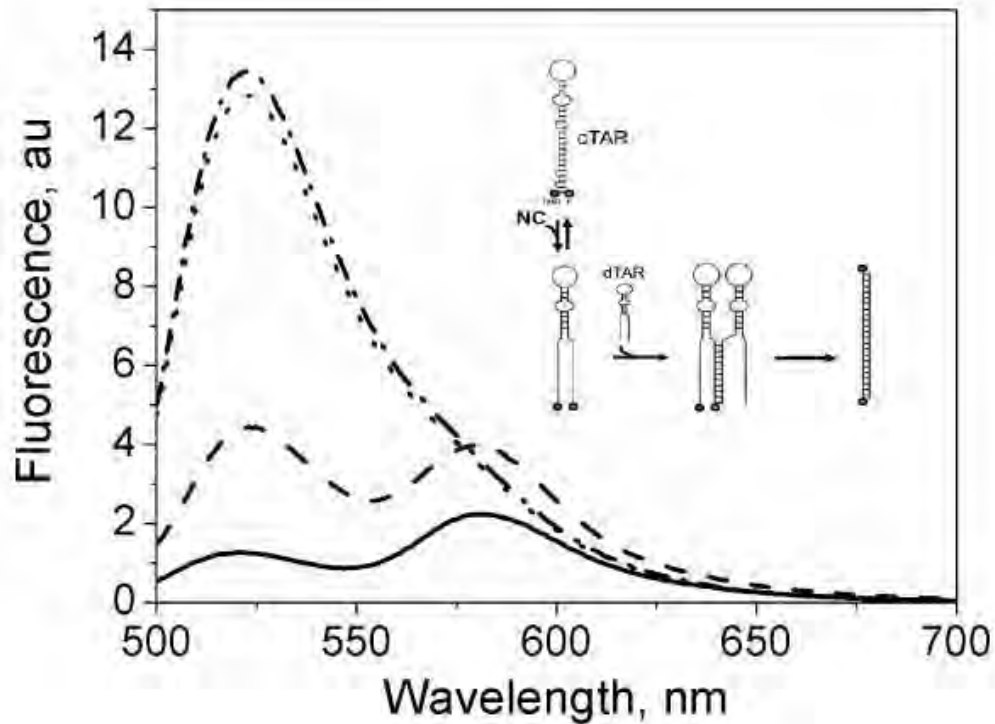
Supporting information

HIV-1 nucleocapsid traps reverse transcriptase on nucleic acid substrates

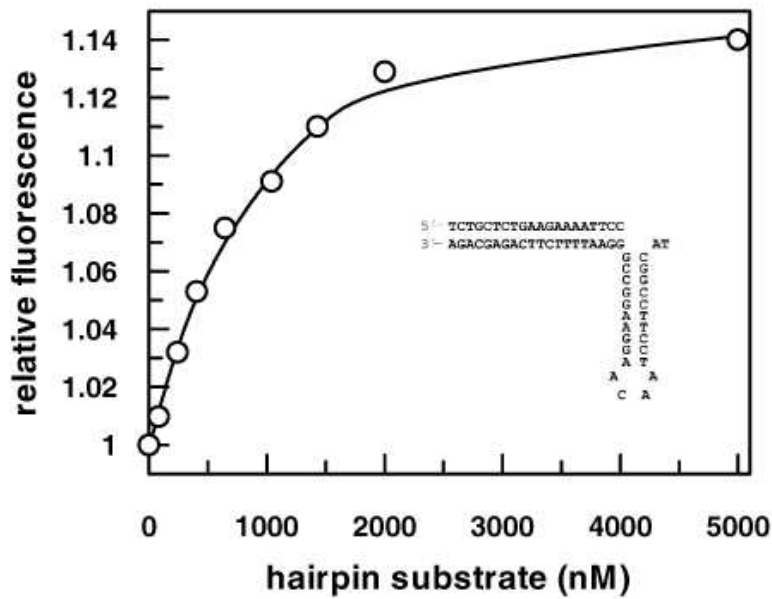
Dina Grohmann, Julien Godet, Yves Mély, Jean-Luc Darlix and Tobias Restle



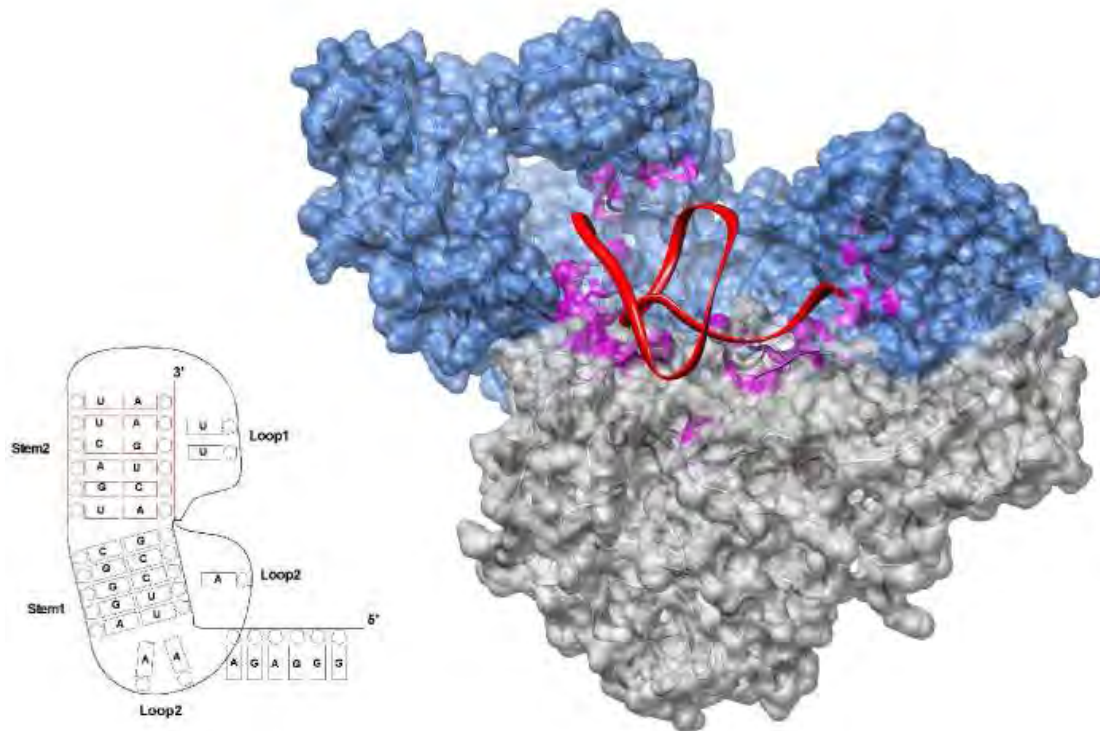
Supplementary Figure 1. Nucleic acid binding properties of HIV-1 NC. 100 nM of a radioactive labelled p/t probe (A: linear substrate, B: hairpin substrate, see Figure 3) was incubated with increasing amounts of HIV-1 NC¹⁻⁵⁵ and the resulting complexes were resolved by non-denaturing gel electrophoresis. Bands corresponding to the free DNA and the DNA-protein complexes were detected by autoradiography. The NC concentrations were in case of linear p/t 2.7 μ M, 5.4 μ M, 27 μ M, and 54 μ M and in case of hairpin p/t 3.4 μ M, 6.8 μ M, 34 μ M, and 68 μ M. Experiments were performed in standard buffer (50 mM Tris/HCl pH 8.0, 50 mM KCl and 10 mM MgCl₂).



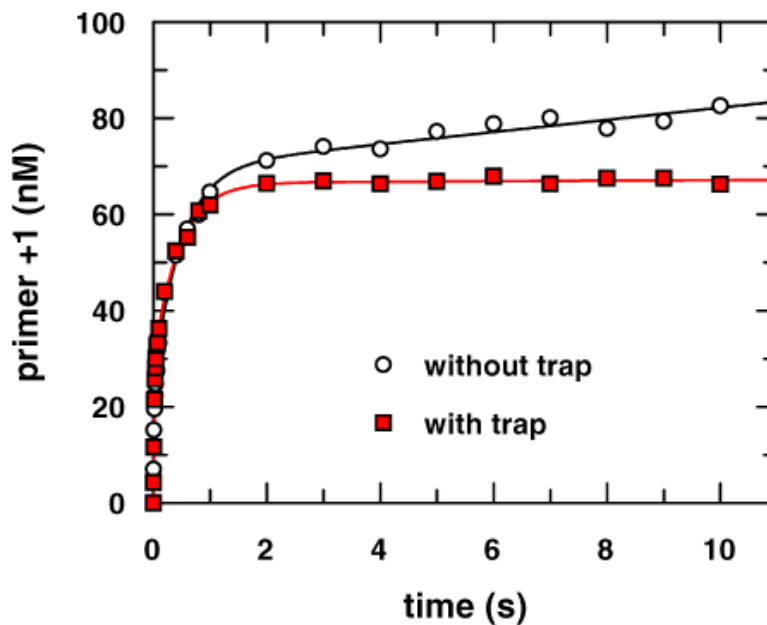
Supplementary Figure 2. Emission spectra of TMR-5'-cTAR-3'-FI. The emission spectra of 0.1 μM TMR-5'-cTAR-3'-FI were recorded in 50 mM Tris/HCl pH 8.0, 50 mM KCl and 0.2 mM MgCl_2 with NC added at a protein to nucleotide ratio of 2, either in the absence (solid line) or presence of 2 μM dTAR, after completion of the annealing reaction (dash). The spectrum recorded after completion of the annealing reaction in the same buffer, but at 10 mM Mg^{2+} (dot) superimposes with the spectrum of heat-promoted cTAR/dTAR duplexes (dash-dot).



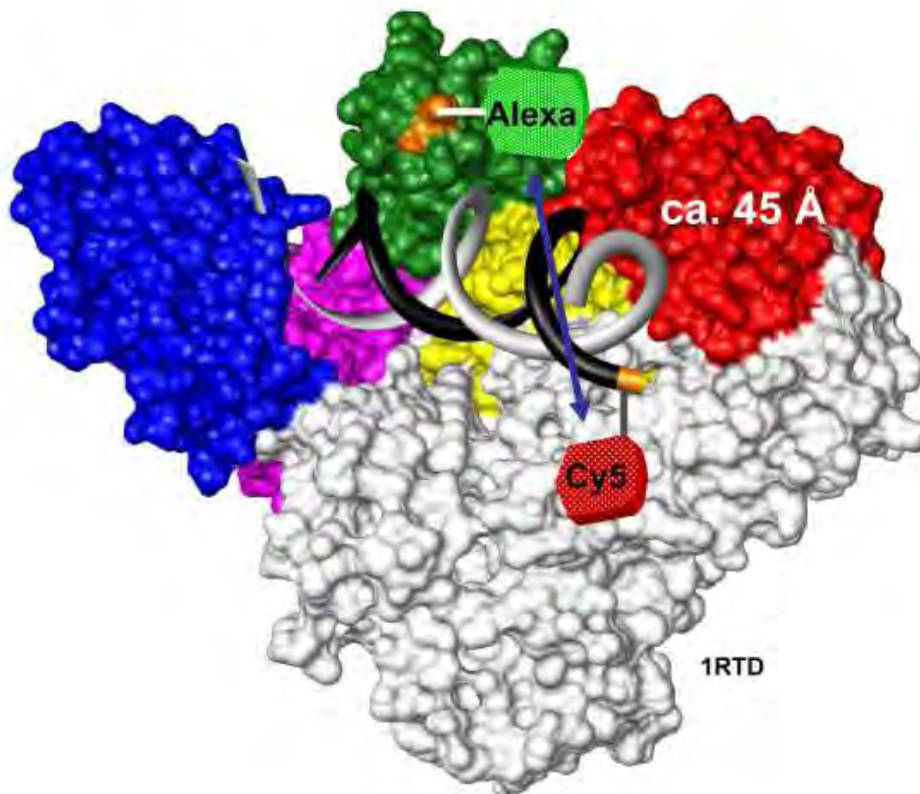
Supplementary Figure 3. Displacement titration of fluorescent p/t bound to RT with the hairpin substrate (see inset). A complex of FAM-labelled 18/36-mer DNA/DNA p/t (25 nM) and RT (45 nM) was titrated with increasing amounts of competitor, i.e. the hairpin substrate. The curve shows the best fit by least-squares fitting to a model describing the two binding equilibria from which a dissociation constant of 83 (± 14) nM for the RT/ hairpin substrate interaction was obtained.



Supplementary Figure 4. Structure of the HIV-1 RT heterodimer complexed with a 33 nucleotide pseudoknot RNA aptamer (PDB file: 1hvu) used for the trap experiments and schematic representation of the base pairs formed in the pseudoknot RNA. The color coding is as follows: p66 (blue), p51 (grey) and RNA (red). Surface residues < 5 Å away from the pseudoknot inhibitor are colored in magenta.



Supplementary Figure 5. Biphasic single turnover, single nucleotide incorporation kinetics (rapid quench experiments) of HIV-1 RT with a 18/36mer DNA/DNA p/t using a pseudoknot RNA aptamer competitor substrate (Supplementary Figure 4). A preformed complex of 200 nM RT and 100 nM p/t was rapidly mixed with a solution containing 100 μ M dGTP without or with 10 μ M of pseudoknot RNA aptamer as a trap. Data were fitted to a double exponential equation plus slope. The addition of the trap leads to a complete disappearance of the linear phase. This phase is due to faulty bound nucleic acid substrate which has to dissociate and properly rebind before nucleotide incorporation occurs (see Scheme 1 for a graphical illustration). Neither the amplitudes of the burst phases nor the burst rates are affected by the RNA aptamer.



Supplementary Figure 6. FRET-based system used for RT-nucleic acid complex dissociation experiments. The p51 subunit is shown in gray and the DNA/DNA primer/template substrate in black and gray, respectively. The p66 subunit is color coded as follows: fingers (blue), palm (magenta), thumb (green), connection (yellow) and RNase H (red). The positions of the two dyes which are approximately 45 Å apart (including the linkers) are indicated (according to PDB 1RTD). The Förster radius of this dye pair is 49 Å.

Étude des propriétés chaperonnes de NCp7 à
l'aide de la 2-Aminopurine

Étude des propriétés chaperonnes de la NCp7 à l'aide de la 2-Aminopurine

Les variations locales de la structure ou de la dynamique de l'ADN sont supposées guider les interactions ADN-protéines au cours de nombreux processus cellulaires. De même, la liaison protéique sur l'ADN induit généralement un remaniement de celle-ci. Il est donc intéressant de pouvoir sonder localement la structure et la dynamique de l'ADN. La 2-Ap, ou 2'-deoxyribosyl-2-aminopurine (Figure 5.1.), est une sonde structurale fluorescente pouvant se substituer aux adénines naturelles dans des acides nucléiques synthétiques. La 2-Ap permet d'accéder à des données structurales et dynamiques de la séquence d'acide nucléique dans laquelle elle est insérée. La 2Ap est une sonde sensible à son microenvironnement et plus particulièrement à l'empilement (*stacking*) avec ses bases voisines. Pour cette raison, elle fournit des informations locales au sein d'un acide nucléique.

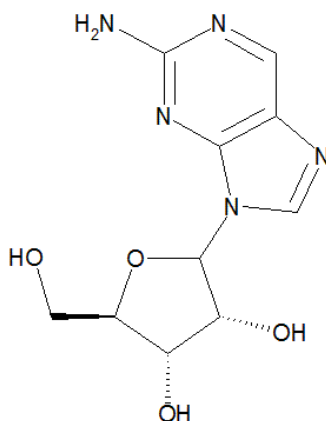


FIGURE 5.1 – Structure de la 2-Ap riboside

5.1 Caractérisation site-spécifique de la liaison de NCp7 sur un oligonucléotide à deux sites de liaison

Les propriétés de la 2-aminopurine ont été exploitées dans le but de décrire la nature des interactions entre la NCp7 et les oligonucléotides. Faisant suite aux travaux inauguraux au sein du laboratoire sur la caractérisation de l'interaction de la NCp7 avec un hexanucléotide [399], la première partie de ce travail [256] s'est portée sur la caractérisation de la liaison de la NCp7 sur des séquences oligonucléotidiques (ODN) de 12 bases possédant deux sites de liaison potentiels (Figure 5.2).

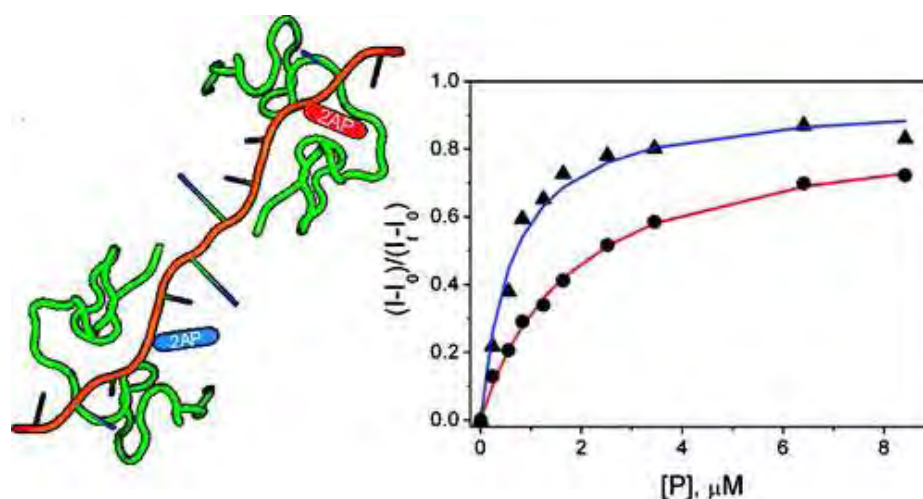


FIGURE 5.2 – Illustration de la liaison de deux NCp7 sur un dodécanucléotide marqué en position 3 (rouge) ou en position 9 (bleu) par une 2-Ap . Le signal d'émission de fluorescence diffère d'un site à l'autre et permet de déterminer les constantes d'affinité de chacun des sites.

La caractérisation des dodécanucléotides marqués (Figure 5.3) a montré une grande similarité des signaux de la fluorescence de la 2-Ap, à la fois en absence et en présence de NCp7, par rapport à ceux rapportés pour la séquence AACGCC utilisée dans des travaux précédents.

En absence de NCp7

Au sein de l'ODN, la fluorescence de la 2-Ap est très largement quenchée, avec un maximum d'émission au voisinage de 370 nm. Le déclin de fluorescence est multiple et nécessite

5.1 Caractérisation site spécifique de la liaison de NCp7

possible de déterminer les constantes d'équilibre microscopiques. Celles-ci sont déterminées selon un modèle d'interaction où deux NCp7 peuvent se lier à l'ODN selon le schéma séquentiel 5.4.

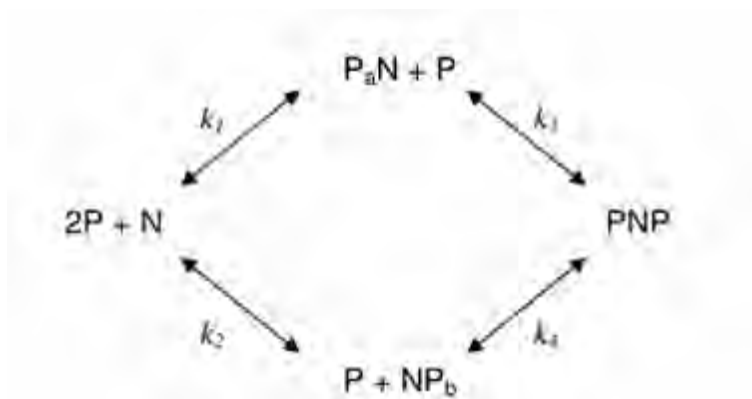


FIGURE 5.4 – Modèle de l'interaction d'un dodécanucléotide (N) avec deux NCp7 (N)

Nous avons observé des différences substantielles d'affinité entre les sites, y compris au sein d'un même ODN (Figure 5.2). De plus, la liaison de la NCp7 sur le motif TG (qui définit le site préférentiel de liaison) est contexte dépendante, puisque l'affinité de NCp7 pour le motif TG est modulée par la position de celui-ci au sein de la séquence. La position relative de la 2-Ap par rapport au motif TG qui constitue le site préférentiel de liaison de la NCp7 définit également la capacité à suivre sélectivement la liaison sur un seul site. Enfin, nos données ne mettent pas en évidence de coopérativité dans le mode de liaison de la NCp7. Cette étude constitue à notre connaissance la première étude de la liaison de NCp7 de manière site-spécifique sur un ADN présentant plusieurs sites de liaison. Ces résultats ont été publiés dans *Biochemistry* en 2009.

5.2

Publication V

Site-Specific Characterization of HIV-1 Nucleocapsid Protein Binding to Oligonucleotides with Two Binding Sites[†]

Sergiy V. Avilov,^{‡,§} Julien Godet,[‡] Etienne Piémont,[‡] and Yves Mély^{*,‡}

Laboratoire Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74 route du Rhin, 67401 Illkirch, France, and Palladin Institute of Biochemistry, 9, Leontovich street, 01030 Kiev, Ukraine

Received December 8, 2008; Revised Manuscript Received January 30, 2009

ABSTRACT: The nucleocapsid protein (NC) of HIV-1 is a highly conserved protein essential for the virus life cycle that constitutes an attractive target for new antiviral agents. Most NC functions rely on its binding to the HIV-1 genomic RNA and its DNA copies that contain multiple and possibly interdependent binding sites. Therefore, a detailed understanding of NC binding requires a site-specific experimental approach. We have recently shown that 2-aminopurine (2Ap), a fluorescent adenine analogue, can site-selectively probe the binding of NC. Here, we introduced 2Ap at various positions of model single-stranded dodecanucleotides containing two TG motifs which constitute putative specific binding sites. Steady-state and time-resolved fluorescence experiments indicated that NC binding strongly increased the fluorescence quantum yield of 2AP by reducing the dynamic quenching of 2Ap by its close neighbors and slowing the picosecond to nanosecond conformational fluctuations of the oligonucleotides. The dodecanucleotides were found to bind two NC molecules at physiological salt concentrations, confirming the preferential binding of NC to TG motifs and an occluded binding site size for NC of five to six bases. Using the NC-induced changes in 2Ap fluorescence, we determined the microscopic affinity constants of the individual binding sites and showed that affinities can significantly differ from one site to another within the same dodecanucleotide, depending on the position of the TG dinucleotide and the nature of its close neighbors. Moreover, our data suggest that binding of NC even to close binding sites shows no strong cooperativity.

HIV-1 nucleocapsid protein (NC),¹ either as a domain of the Gag polyprotein precursor or as a mature protein, is essential for several important steps of the virus life cycle (*1*). For instance, selection of viral genomic RNA for packaging into virions is mediated by binding of the NC domain within Gag to the ψ encapsidation sequence within the untranslated region of the HIV-1 genome (*2*). Further, NC chaperones the annealing of the primer tRNA to the primer binding site, and the two obligatory strand transfers necessary for the synthesis of a complete proviral DNA by reverse transcriptase (*3–5*). Mature NC is a basic protein of 55 amino acids (Figure 1) containing two highly conserved CX₂CX₄HX₂C zinc fingers [residues 15–28 and 36–49, termed F1 and F2, respectively (*6*)] with very high affinity for Zn²⁺ ions (*7*). Binding of Zn²⁺ to the fingers results in a highly constrained folded structure in which the two CCHC motifs of NCp7 are spatially close and weakly interacting

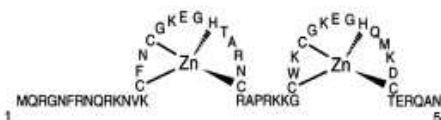


FIGURE 1: Sequence of the HIV-1 nucleocapsid protein.

with one another (*8–10*). Mutations of amino acids involved in zinc binding lead to imperfectly folded protein and noninfectious viruses (*11–20*). The NC mutant virions contain small amounts of genomic RNA and have defects in reverse transcription. NC is thus an attractive target for the development of new antiviral agents (*1, 21*).

The mechanism of the recognition by NC of its nucleic acid targets is not fully understood. Experiments with NC and short oligonucleotides (ODNs) revealed that at least five bases are needed for stable NC binding (*22–24*). Moreover, it was found that NC binds to poly-G better than to any other homopolymer sequence, but much more stable binding was observed for alternating base sequence (TG)_n (*22*). Detailed studies of NC binding to hexanucleotides with systematically varied bases revealed moderate association constants ($\approx 10^5–10^6$ M⁻¹) and a key role of the TG motif for specific binding (*23*). NC binding to its natural targets, namely the stem-loop regions of viral genomic RNA, was studied as well, confirming the preference of NC for TG- or GXC-containing single-stranded sequences (*23, 25–27*). NMR structures of several NC–ODN complexes clearly demonstrated that guanine residues play a key role in NC binding,

[†] The study was supported by Agence Nationale de Recherches sur le SIDA (ANRS) and the European consortium "Targeting Replication and Integration of HIV". S.V.A. was a fellow of Fondation pour la Recherche Médicale, France.

* To whom correspondence should be addressed. Telephone: +33 (0)3 90 24 42 63. Fax: +33 (0)3 90 24 43 12. E-mail: mely@pharma.u-strasbg.fr.

[‡] UMR 7213 CNRS.

[§] Palladin Institute of Biochemistry.

¹ Abbreviations: 2Ap, 2-aminopurine; HIV-1, human immunodeficiency virus type 1; NC, HIV-1 nucleocapsid protein; ODN, oligonucleotide; NMR, nuclear magnetic resonance; PBS, primer binding site; SD, standard deviation.

through stacking with the Trp37 residue of the C-terminal finger motif (28–32). In fact, Trp37 belongs to a hydrophobic plateau that forms at the top of the folded zinc fingers and is critical for the nucleic acid chaperone properties of NC (33, 34). This hydrophobic plateau also includes the Val13, Phe16, Thr24, and Ala25 residues which form additional contacts with the ODN binding sites (28–32). Moreover, it was found that more than one NC molecule binds to the (TG)₄ sequence and more than one octanucleotide can bind to a single NC molecule (35). However, the binding constant of the additional ODN is much lower than that in the 1:1 complex. Finally, the parameters for binding of NC to longer DNA and RNA sequences were also reported, but using simplifying assumptions such as single binding sites or identical and independent binding sites (2, 33, 36–40).

In fact, the analysis of NC binding to ODNs with multiple binding sites is hampered by the intrinsic limitation of most techniques that allow monitoring of the binding in a global manner only, with no ability to discriminate between the different binding sites. Moreover, appropriate models that describe the binding to overlapping binding sites with different affinities and cooperativity are also missing. To circumvent the first limit, ODNs labeled with site-specific reporters can be used. In this respect, we recently demonstrated that 2-aminopurine (2Ap), an environmentally sensitive fluorescent analogue of adenine, may serve as a site-specific sensitive reporter for NC binding (41). In the absence of NC, the fluorescence of 2Ap is highly quenched, due to collisions with the neighbor bases that lead to stacking interactions. Binding of NC restricted the ODN flexibility on the picosecond to nanosecond time scale, impeding the collisions of 2Ap with its immediate neighbors, which in turn results in a strong fluorescence increase. As a consequence, 2Ap can probe site-specific conformational and dynamic changes upon NC binding.

In this respect, the aim of this work was to investigate NC binding to three 2Ap-labeled dodecanucleotides that constitute the simplest presumably “multivalent” ODNs. To limit the contributions of multiple overlapping sites in the binding process, all ODNs contained two TG motifs that constitute strong preferential binding sites for NC. The three ODNs differed by the position and proximity of the TG motifs within their sequence and were non-self-complementary and unfolded. To obtain site-specific information, 2Ap was introduced at various positions within these ODNs. We found that the three tested ODNs bound two NC proteins with affinities dependent on the position and on the neighbors of the TG motifs.

MATERIALS AND METHODS

Materials. NC(1–55) (Figure 1) was synthesized on an Applied Biosystems A433 peptide synthesizer as described previously (42). Single-stranded DNA dodecanucleotides, unlabeled, covalently labeled with AlexaFluor430 (Invitrogen; further termed “Alexa430”), and 2Ap-substituted, were synthesized and HPLC-purified by IBA. The studied sequences TGACCGTGACCG, AATGACTGAAAC, and AAAGTGAATGAC are termed ODN1, ODN2, and ODN3, respectively; 2Ap substitutions are indicated by adding “ApX” to the acronym, where X is the position of the

2-aminopurine residue. All the dodecanucleotides are unfolded at 20 °C according to the mfold software (43) (<http://frontend.bioinfo.rpi.edu/applications/mfold/>). Absorbance spectra were recorded with a Cary 400 UV–vis spectrophotometer (Varian). Concentrations of the ODNs were calculated from their absorbance using the molar extinction coefficients (ϵ_{260}) specified by the supplier. All experiments were performed at 20 °C in 50 mM Hepes (pH 7.5) either with or without 0.1 M NaCl. All chemicals were purchased from Sigma.

Steady-State Fluorescence Spectroscopy. Fluorescence spectra were recorded on a FluoroMax3 spectrofluorimeter (Jobin Yvon) equipped with a thermostated cell compartment. Fluorescence intensities were corrected for buffer fluorescence and screening effects. Quantum yield was calculated using free 2Ap as a reference [0.68 (44)], with an excitation wavelength of 315 nm. To determine the constant for binding of NC to the ODNs, fixed concentrations of the 2Ap-containing ODNs were titrated with increasing concentrations of NC, in the presence of 0.1 M NaCl. Excitation and emission wavelengths were 315 and 370 nm, respectively. The parameters were recovered from nonlinear fits of eqs 4a and 5 (see Results) to the experimental data with Microcal Origin 7.0.

Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique, as described previously (41, 45). Excitation at 315 and 430 nm for 2Ap and Alexa430 was provided by a pulse-picked frequency-tripled and -doubled Ti-sapphire laser (Tsunami, Spectra Physics), respectively, pumped by a Millennia X laser (Spectra Physics). Emission was collected through a polarizer set at magic angle and an 8 nm band-pass monochromator (Jobin-Yvon H10) set at 370 and 540 nm for 2Ap and Alexa430, respectively. The single-photon events were detected with a microchannel plate Hamamatsu R3809U photomultiplier coupled to a Philips 6954 pulse preamplifier and recorded on a multichannel analyzer (Ortec 7100), calibrated at 25.5 ps/channel. The instrumental response function was recorded with a polished aluminum reflector, and its full width at half-maximum was 40 ps. Fluorescence intensity decays $I(t)$ were analyzed as a sum of exponentials: $I(t) = \sum \alpha_i \exp(-t/\tau_i)$, where τ_i values are the fluorescence lifetimes and α_i values are the associated amplitudes such that $\sum \alpha_i = 1$. The mean lifetime was calculated according to the relationship $\langle \tau \rangle = \sum \alpha_i \tau_i$. The population, α_0 , of 2Ap dark species within the ODNs was calculated with the relationship $\alpha_0 = 1 - \tau_{\text{free}}/(\tau_{\text{sample}} R_m)$, where τ_{free} is the lifetime of free 2Ap, τ_{sample} is the measured lifetime of 2Ap within a given ODN (either free or bound to NC), and R_m is the ratio of the corresponding steady-state fluorescence intensities. The remaining amplitudes, α_{ic} , were recalculated from the measured amplitudes, α_i , according to the relationship $\alpha_{ic} = \alpha_i(1 - \alpha_0)$.

For time-resolved anisotropy measurements, the fluorescence decay curves were recorded at vertical and horizontal positions of the polarizer, as described previously (41). Time-resolved fluorescence anisotropy decays were analyzed with the following equations:

$$I_{\parallel}(t) = I(t)[1 + 2r(t)]/3$$

$$I_{\perp}(t) = I(t)[1 + r(t)]/3$$

$$r(t) = \frac{I_{\parallel}(t) - GI_{\perp}(t)}{I_{\parallel}(t) + 2GI_{\perp}(t)} = r_0 \sum \beta_i \exp(-t/\varphi_i) \quad (1)$$

where β_i values are the amplitudes of the rotational correlation times φ_i ; I_{\parallel} and I_{\perp} are the intensities collected at emission polarizations parallel and perpendicular, respectively, to the polarization axis of the excitation beam, and G is the geometry factor at the emission wavelength, determined in independent experiments. The fundamental anisotropy r_0 value (0.33) was determined independently for 2Ap in 77% glycerol (v/v) (41) and was fixed in the analysis of the time-resolved anisotropy data. The theoretical values of the rotational correlation times for the ODNs and their complexes were calculated assuming spherical shapes, by

$$\phi = \frac{\eta M(v+h)}{RT} \quad (2)$$

where η is the viscosity, T is the temperature, v is the specific volume of the particle, h is the degree of hydration, and R is the molar gas constant. Values of 0.78 and 0.4 mL/g were taken for the specific volume and the degree of hydration, respectively.

Time-resolved intensity and anisotropy data were treated with a nonlinear least-squares analysis using a homemade program (kindly provided by G. Krishnamoorthy). In all cases, the χ^2 values were close to 1, and the weighted residuals were distributed randomly around zero, indicating an optimal fit.

RESULTS

Characterization of the Dodecanucleotides and Their Complexes with NC by 2Ap Fluorescence. 2Ap was introduced into various positions within three single-stranded DNA dodecanucleotide sequences (ODN1, ODN2, and ODN3) containing two TG motifs at different positions (Table 1). We observed that the fluorescence spectra of the 2Ap-labeled dodecanucleotides, like free 2Ap and hexanucleotides, exhibit an emission maximum around 370 nm (not shown). The quantum yield of 2Ap in these ODNs was quite low (Table 1), indicating a strong quenching of 2Ap fluorescence by the neighbor bases (for a review, see ref 46). Quenching is the strongest (quantum yields in the range of 0.007–0.019) when 2Ap is flanked by a guanine (ODN1Ap3, ODN1Ap9, ODN2Ap5, ODN3Ap7, and ODN3Ap11), consistent with the fact that G is the most efficient quencher of 2Ap among the natural bases (47). The remaining sequences (ODN2Ap2, ODN2Ap10, and ODN3Ap3) exhibit nearly

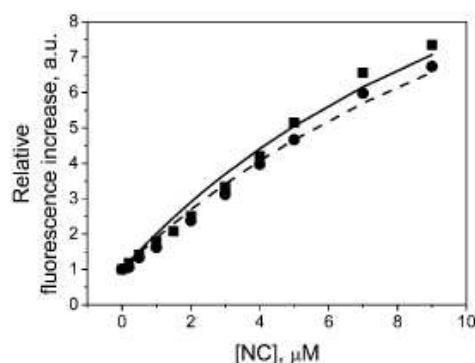


FIGURE 2: Influence of 2Ap substitution on the binding to NC. The 2Ap-labeled ODN2Ap5 sequence was titrated with NC either in the absence (■, solid line) or in the presence of an equimolar concentration of the corresponding unlabeled ODN (●, dashed line). The total ODN concentration was 2 μ M in both cases.

identical quantum yields (0.045–0.051), though the nature of the 3'-flanking residue of 2Ap differs in these three sequences. This indicated that the nature of this residue does not markedly influence the 2Ap quantum yield. The time-resolved intensity decays of the dodecanucleotides (Table 1 of the Supporting Information) showed multiple lifetimes, attributed to picosecond to nanosecond conformational fluctuations of the 2Ap-labeled ODNs in the excited state that drive 2Ap from unstacked to stacked conformations.

Probing NC–dodecanucleotide interactions by 2Ap is only valid if the substitution of a natural base by 2Ap does not affect the interaction. We have shown previously that 2Ap substitution does not significantly influence NC binding to AATGCC and AACGCC sequences (41). To demonstrate that 2Ap substitution does not also affect NC binding to the studied dodecanucleotides, the titration curves of 2 μ M 2Ap-labeled dodecanucleotides with NC were compared to the titration curves obtained using an equimolar mixture of labeled and unlabeled sequences at the same concentration. The data with ODN1Ap3, taken as a representative example, are illustrated in Figure 2. The binding curves were found to be very similar, indicating that 2Ap-labeled and native sequences exhibit comparable affinities. Thus, substitution of A with 2Ap does not significantly influence NC binding. Similar results were obtained for all substituted positions (data not shown).

When NC binds, the fluorescence quantum yield of 2Ap within the dodecanucleotides significantly increases (Table 1), but the position of the emission maximum does not change (data not shown), indicating that NC binding

Table 1: Quantum Yields and NC Binding Constants of the 2Ap-Substituted Dodecanucleotides

| sequence | quantum yield, free ODN ^a | quantum yield, ODN + NC ^a | k_1 ($\times 10^{-6}$ M ⁻¹) ^b | k_2 ($\times 10^{-6}$ M ⁻¹) ^b |
|----------|--------------------------------------|--------------------------------------|---|---|
| ODN1Ap3 | TGApCCGTGACCG | 0.007 | 0.078 | 0.41 |
| ODN1Ap9 | TGACCGTGApCCG | 0.009 | 0.135 | 1.40 |
| ODN2Ap2 | AApTGACTGAAAC | 0.046 | 0.202 | 0.87 |
| ODN2Ap5 | AATGApCTGAAAC | 0.012 | 0.194 | 0.81 |
| ODN2Ap10 | AATGACTGAApAC | 0.051 | 0.105 | 1.80 |
| ODN3Ap3 | AAApCTGAATGAC | 0.045 | 0.173 | 0.66 |
| ODN3Ap7 | AAACTGApATGAC | 0.015 | 0.198 | 0.30 |
| ODN3Ap11 | AAACTGAATGApC | 0.019 | 0.167 | 0.44 |

^a Quantum yields were calculated assuming a quantum yield of 0.68 for free 2Ap (44). The standard error of the mean was less than 15%, for at least three measurements. ^b The microscopic ("site-specific") equilibrium binding constants were obtained by fitting the data of Figure 3 to eqs 4a and 4b, with the exception of the ODN3Ap7 data which were fitted to eq 5. The standard error of the mean for three measurements was less than 20%. Excitation was at 315 nm. Emission was at 370 nm and was corrected for inner-filter effects and buffer fluorescence.

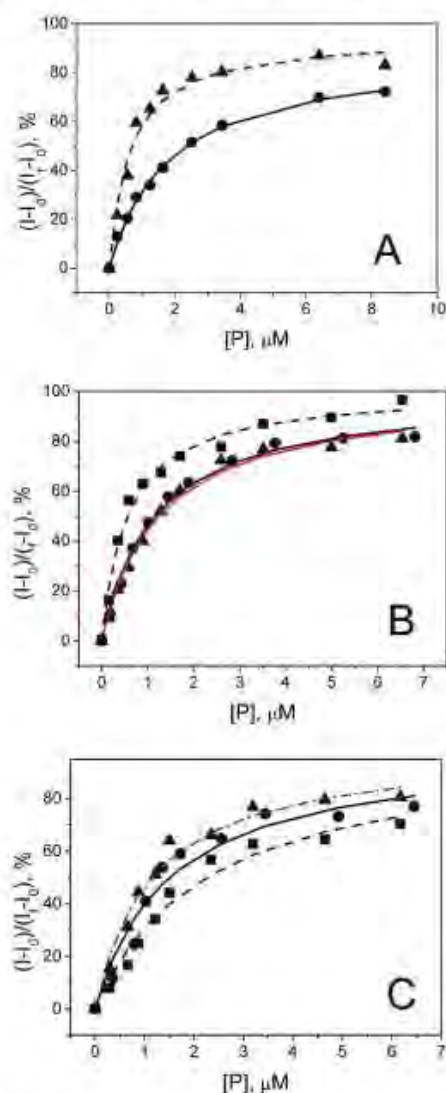


FIGURE 3: Titration curves of 2Ap-substituted dodecanucleotides with NC. The relative 2Ap fluorescence intensity changes $(I - I_0)/(I - I_0)_0$ were plotted as a function of the free NC concentration calculated with eq 3. Excitation was at 315 nm and emission at 370 nm. The ODN concentration was 2 μ M in 50 mM HEPES (pH 7.5) and 0.1 M NaCl. Titrations were performed with (A) ODN1 labeled by 2Ap at position 3 (\bullet , solid line) or 9 (\blacktriangle , dashed line), (B) ODN2 labeled by 2Ap at position 2 (\bullet , black solid line), 5 (\blacktriangle , red solid line), or 10 (\blacksquare , dashed line), and (C) ODN3 labeled at position 3 (\bullet , solid line), 7 (\blacktriangle , blue dashed-dotted line), or 11 (\blacksquare , dashed line). The solid and dotted lines correspond to the fits of the experimental points with eqs 4a and 4b, respectively. The blue dashed-dotted line in Figure 3C corresponds to the fit of ODN3Ap7 with eq 5.

significantly reduces the level of quenching of 2Ap fluorescence by its neighbor bases but does not change the polarity of its surroundings. The NC-induced increase in the fluorescence quantum yield was found to depend on the 2Ap position and the ODN sequence, ranging from 2.1-fold for ODN2Ap10 to 16-fold for ODN2Ap5. The strongest fluorescence increase was observed when 2Ap is next to a G residue. Interestingly, substantial differences in the quantum yields occurred in the bound forms of ODN1Ap3, ODN1Ap9, and ODN2Ap5, though in all these sequences 2Ap is

surrounded by a TG motif and a C residue. This suggests that NC does not bind similarly to the TGApC motifs in these ODNs. Fluorescence lifetime experiments (Table 1 of the Supporting Information) further revealed that as previously reported with hexanucleotides (41), NC restricts the collisions of 2Ap with its neighbor bases and strongly decreases the population of the fully stacked conformations of the probe.

Site-Specific Determination of NC–Dodecanucleotide Binding Parameters. (i) **Binding Stoichiometry.** To determine the stoichiometry of the NC–ODN complexes, time-resolved anisotropy decay experiments were performed to characterize the different rotational motions of the labeled species. The rotational correlation time associated with the tumbling of the labeled species is of special interest, since it depends on the hydrodynamic radius and, thus, on the molecular mass of the labeled species (48). In most free ODNs, the quantum yield of 2Ap was too low to obtain reliable time-resolved anisotropy parameters. In contrast, the higher quantum yields of the NC–ODN complexes allowed us to analyze their anisotropy decay curves. Three rotational correlation times were needed for good fitting of the anisotropy decay curves (Table 2), in contrast to the NC–hexanucleotide complexes which needed two components only (41). This increased complexity is probably associated with a more complicated set of local motions in the longer sequences. The two shorter components (~ 0.1 – 0.2 and 1 – 1.5 ns, respectively) presumably correspond to local rotations of the dye and/or segments of the ODNs, while the longest component (φ_3) is attributed to the tumbling of the whole complex. φ_3 values ranged from 7.5 to 10.7 ns. These values are significantly higher than the theoretical correlation time of 4.7 ns calculated for the tumbling of a sphere with the molecular mass of a 1:1 NC–ODN complex but are consistent with the theoretical correlation time of 7.5 ns calculated for a 2:1 NC–ODN complex. Deviations of the measured φ_3 values from this last value are likely due to the nonspherical shape of the complexes, as observed previously with the 1:1 complexes of NC with hexanucleotides (41). Thus, all tested ODNs were able to bind two NC molecules. To confirm the stoichiometry of the complexes, time-resolved anisotropy experiments were repeated using the same ODNs labeled at their 5'-terminus with Alexa430, a highly fluorescent dye. For the free ODNs, two rotational correlation times were determined (Table 3). The faster component ($\varphi_1 = 0.2$ ns) describes the local rotation of the dye. The slower component (1.1–1.4 ns) is somewhat below the theoretical rotational correlation time (1.8 ns) of a sphere with the molecular mass of the free ODNs, indicating that both tumbling and segmental motions of the ODNs likely contribute to this rotational correlation time (49). Both the large amplitude associated with the local motion and the contributions of segmental motions to the slower component are fully consistent with the high flexibility of the free ODNs, which allows conformational fluctuations in the picosecond to nanosecond time range. Binding of NC was found to strongly modify the time-resolved anisotropy decays of the Alexa430-labeled ODNs. As for the complexes with the 2Ap-labeled ODNs, three rotational correlation times were needed to describe the anisotropy decays of the complexes with Alexa430-labeled ODNs. The values of the slow rotational correlation time ($\varphi_3 = 7.3$ – 8.8 ns) were fully consistent with the values of the corresponding components

Table 2: Time-Resolved Anisotropy Parameters for the 2Ap-Substituted Dodecanucleotides Bound to NCp7^a

| sequence | φ_1 (ns) | β_1 | φ_2 (ns) | β_2 | φ_3 (ns) | β_3 | |
|----------------------|------------------|-----------|------------------|-----------|------------------|-----------|------|
| fre ^c 2Ap | 0.08 | 1.00 | | | | | |
| ODN1Ap3 | TGApCCGTGACCG | 0.32 | 0.23 | 1.1 | 0.22 | 9.6 | 0.55 |
| ODN1Ap9 | TGACCGTGApCCG | 0.12 | 0.21 | 1.0 | 0.25 | 7.5 | 0.54 |
| ODN2Ap2 | AApTGACTGAAAC | 0.15 | 0.40 | 1.6 | 0.13 | 8.8 | 0.47 |
| ODN2Ap5 | AATGApCTGAAAC | 0.10 | 0.29 | 1.3 | 0.15 | 8.5 | 0.56 |
| ODN2Ap10 | AATGACTGAApAC | 0.23 | 0.52 | 1.6 | 0.12 | 9.3 | 0.36 |
| ODN3Ap3 | AAApCTGAAATGAC | 0.13 | 0.14 | 0.9 | 0.29 | 8.4 | 0.57 |
| ODN3Ap7 | AAACTGApATGAC | 0.30 | 0.34 | 1.3 | 0.12 | 9.5 | 0.54 |
| ODN3Ap11 | AAACTGAATGApC | 0.11 | 0.39 | 1.4 | 0.31 | 11 | 0.30 |

^a The fluorescence rotational correlation times (φ_i) and their amplitudes (β_i) were obtained from the time-resolved anisotropy decays, as described in Materials and Methods. Standard deviations for the rotational correlation times and amplitudes are less than 20 and 15%, respectively. Experiments were performed with 0.5–1.0 μ M ODNs and 7.5 μ M NC in 0.05 M HEPES (pH 7.5) in the absence of NaCl. Taking into account the affinities at 0.1 M NaCl (Table 1) and the salt dependence of NC binding to ODNs (23, 40), all binding sites are likely saturated with NC under these conditions. The addition of a third rotational correlation time in the fit to the experimental decay of the NC–ODN complexes decreased the χ^2 values from 2 to 1.1–1.3.

Table 3: Time-Resolved Anisotropy Parameters of Alexa430-Labeled Dodecanucleotides and Their Complexes with NC^a

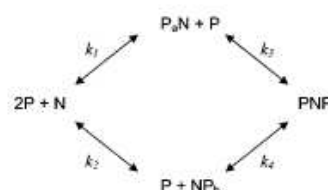
| sequence | NC | φ_1 (ns) | β_1 | φ_2 (ns) | β_2 | φ_3 (ns) | β_3 | |
|---------------|--------------|------------------|-----------|------------------|-----------|------------------|-----------|------|
| Alexa430–ODN1 | TGACCGTGACCG | – | 0.22 | 0.58 | 1.4 | 0.42 | | |
| | + | 0.47 | 0.39 | 3.5 | 0.18 | 7.3 | 0.43 | |
| ODN1–Alexa430 | TGACCGTGACCG | – | 0.25 | 0.50 | 1.4 | 0.50 | | |
| | + | 0.19 | 0.36 | 1.5 | 0.30 | 7.9 | 0.34 | |
| Alexa430–ODN2 | AATGACTGAAAC | – | 0.22 | 0.63 | 1.4 | 0.37 | | |
| | + | 0.39 | 0.34 | 1.7 | 0.21 | 7.5 | 0.45 | |
| Alexa430–ODN3 | AAACTGAATGAC | – | 0.21 | 0.61 | 1.3 | 0.39 | | |
| | + | 0.34 | 0.25 | 1.4 | 0.35 | 8.8 | 0.40 | |
| Alexa430–ODN4 | TGACCGAAACAA | + | 0.09 | 0.06 | 0.7 | 0.29 | 6.1 | 0.65 |
| Alexa430–ODN5 | AATGACCGAAAC | + | 0.14 | 0.30 | 1.0 | 0.30 | 5.9 | 0.40 |

^a Parameter significance and standard deviations are as in Table 2. Experiments were performed with 0.2 μ M ODN in the absence or presence of 2 μ M NC.

in the complexes with the 2Ap-labeled sequences, confirming the 2:1 stoichiometry of the NC–ODN complexes. Moreover, the large amplitude associated with this slow component as well as the absence of any contribution from segmental motions in this component confirmed that NC reduces the flexibility of the ODN and, thus, restricts its picosecond to nanosecond conformational fluctuations. Alexa430 attached to the 3'-terminus yielded the same results (Table 3), suggesting that the position of the dye does not affect the shape and the size of the complexes.

Thus, anisotropy decay data with Alexa430- and 2Ap-labeled ODNs are consistent with a 2:1 stoichiometry for all NC–ODN complexes, as expected from the two high-affinity TG motifs in their sequences (22, 23, 30). To demonstrate the central role of the TG motifs in the binding of the two NC molecules, two mutated Alexa430-labeled ODNs were used. The AATGACCGAAAC sequence (ODN4) corresponds to the ODN1 sequence in which the 3'-terminal TGACCG sequence was substituted with the poorly binding AAACAA sequence (23). The ODN5 sequence (AATGACCGAAAC) corresponded to the ODN2 sequence with a single T⁷ → C mutation that substitutes the central TG motif with a lower-affinity CG motif. With both ODN4 and ODN5 sequences, the slow rotational correlation time dropped to a value of ~6 ns, between the theoretical correlation times of 4.7 and 7.5 ns for spheres with the molecular masses of the 1:1 and 2:1 NC–ODN complexes, respectively. Thus, the more likely explanation is that the complexes of NC with these two ODNs are nonspherical 1:1 complexes, confirming the central role of the TG motifs in the binding sites of the ODN1–ODN3 sequences for the NC molecules.

(ii) *Binding Constants.* As a consequence of the two TG binding sites in the ODN1, ODN2, and ODN3 sequences, a binding scheme with two individual binding sites was used to model the interaction of NC with these sequences:



where P is the free protein, N is the free oligonucleotide, P_aN and NP_b are the 1:1 protein–ODN complexes with the protein bound to the proximal (5'-terminal) and distal (3'-terminal) binding sites, respectively (termed “a” and “b” sites, respectively), PNP is the 2:1 protein–ODN complex, and k_1 – k_4 are the microscopic equilibrium binding constants. Since 2Ap fluorescence within ODNs is sensitive to changes in its immediate vicinity (41, 46, 50, 51), 2Ap should report on NC binding only to the site in which it is included. Using these assumptions, titrations of the various 2Ap-labeled ODNs with NC were used to calculate the concentrations of NC bound to the 2Ap-labeled binding site from the relative increase in 2Ap fluorescence. For instance, for a given ODN labeled at binding site a, the concentration of bound protein at this site is given by

$$\frac{PN_a + PNP}{N_t} = \frac{I_a - I_{a0}}{I_{af} - I_{a0}}$$

where I_a is the 2Ap fluorescence intensity at a given total protein concentration (P_t), I_{a0} and I_{af} are the 2Ap fluorescence intensities in the absence and presence of saturating NC concentrations, respectively, and N_t is the total ODN concentration expressed in species. In contrast, the concentration of free NC, also necessary to draw the binding isotherm and to calculate the binding constants, cannot be determined from a single titration curve, since it also depends on the binding of NC to the nonlabeled binding site. To overcome this, we titrated in parallel equal concentrations of an ODN labeled with 2Ap at the proximal binding site (a) and the same ODN labeled at the distal site (b). Under these conditions, the concentration of free protein P can be calculated by

$$P = P_t - N_t \left(\frac{I_a - I_{a0}}{I_{af} - I_{a0}} + \frac{I_b - I_{b0}}{I_{bf} - I_{b0}} \right) \quad (3)$$

After determination of P values for the various couples of I_a and I_b values, the titration data were replotted as $(I - I_0)/(I_f - I_0)$ versus P (Figure 3) and the microscopic affinity constants for all binding sites were determined by fitting the data to the following equations:

$$\frac{I_a - I_{a0}}{I_{af} - I_{a0}} = \frac{k_1 P + k_1 k_3 P^2}{1 + (k_1 + k_2)P + k_1 k_3 P^2} \quad (4a)$$

$$\frac{I_b - I_{b0}}{I_{bf} - I_{b0}} = \frac{k_2 P + k_1 k_3 P^2}{1 + (k_1 + k_2)P + k_1 k_3 P^2} \quad (4b)$$

where k_1 and k_2 are the microscopic affinity constants for NC binding to sites a and b, respectively. For all sequences, we observed substantial 2Ap position-dependent differences between the titration curves (Figure 3), consistent with our hypothesis that 2Ap reports NC binding at a given site. For the ODN1 sequence, the curve for binding of NC to ODN1Ap9 was significantly shifted toward the low concentrations of the free peptide with respect to the titration curve of ODN1Ap3. The two titration curves were fitted simultaneously to eqs 4a and 4b. Good fits were obtained, assuming that the binding sites were independent, and thus that $k_3 = k_2$ and $k_4 = k_1$. The binding constants in Table 1 were in line with the previously reported binding constants for binding of NC for TG-containing hexanucleotides (23, 41). Moreover, these binding constants show that the b site binds NC with a 3.5-fold higher affinity than the a site even though the ODN1 sequence is a repeat of two TGACCG motifs. This indicates that NC binding to the two TGACCG motifs is context-dependent.

With the ODN2 sequence, the binding curves of ODN2Ap2 and ODN2Ap5 sequences were superimposable (Figure 3B) and clearly shifted to high P concentrations, with respect to the ODN2Ap10 titration curve. This suggests that 2Ap at positions 2 and 5 sensed the binding of NC to the same binding site and that this site had a weaker affinity than the one containing the 2Ap residue at position 10. The fact that the nucleotides at positions 2 and 5 were constitutive of a unique binding site was not surprising since these two residues flanked the central $^3\text{TG}^4$ motif critical for NC binding. In contrast, 2Ap at position 10 is likely associated with the $^7\text{TG}^8$ -containing distal site of the ODN. As in the

case of ODN1, a good fit of the two titration curves was obtained, assuming that the binding sites were independent. The binding constants of the two sites differed by a factor of 2.

For ODN3, analysis of the ODN3Ap3 and ODN3Ap11 binding curves (Figure 3C) revealed that the affinity ($k_1 = 6.6 \times 10^5 \text{ M}^{-1}$) of the proximal binding site with the $^5\text{TG}^6$ motif was 50% higher than the affinity ($k_2 = 4.4 \times 10^5 \text{ M}^{-1}$) of the distal binding site with the $^9\text{TG}^{10}$ motif. Surprisingly, the binding curve of ODN3Ap7 matched with neither of the two preceding curves and was clearly shifted to lower P concentrations, suggesting a higher affinity. Using eq 4a or 4b, an apparent binding constant of $1.1 \times 10^6 \text{ M}^{-1}$, corresponding to the sum of k_1 and k_2 , was obtained. Though 2Ap at position 7 flanks the G^6 residue of the proximal binding site, it is likely sensitive to the binding to both sites, due to its spatial proximity to the distal site. To check this hypothesis, we derived an equation in which we assume that the binding of NC to either site provides the same increase in 2Ap fluorescence and that no further increase occurs when both sites are bound:

$$\frac{I - I_0}{I_f - I_0} = \frac{I_a - I_{a0}}{I_{af} - I_{a0}} + \frac{I_b - I_{b0}}{I_{bf} - I_{b0}} = \frac{(k_1 + k_2)P + k_1 k_3 P^2}{1 + (k_1 + k_2)P + k_1 k_3 P^2} \quad (5)$$

The fit of the ODN3Ap7 binding curve with eq 5 provided k_1 and k_2 values (Table 1) that were in excellent agreement with the values obtained from the fits of the ODN3Ap3 and ODN3Ap11 binding curves, confirming our hypothesis. Thus, the labeling position should be carefully selected to site-selectively monitor the binding of NC.

DISCUSSION

Here we report the first site-specific characterization of NC binding to ODNs with multiple binding sites. An environment-sensitive fluorescent base analogue 2Ap was used for this purpose. In all sequences, 2Ap substituted adenine, which minimally perturbs the ODN properties (Figure 2).

In line with the literature (47, 49–51) and our recent study (41), we observed a strong quenching of 2Ap within the dodecanucleotides, which was dependent on the nature of the flanking bases, the strongest quenching being obtained with G. Interestingly, the quantum yield of 2Ap flanked by a G increases when 2Ap is near the ODN termini. This is likely a consequence of the higher flexibility of the ODN termini, which leads to a less efficient stacking of the bases. As previously shown with hexanucleotides (41), a strong increase in the 2Ap quantum yield (Table 1) occurred upon NC binding. This increase was most pronounced (up to 15-fold) for sequences where 2Ap was flanked by G and was attributed both to NC-induced constraints on ODN flexibility and to specific interactions of NC with the bases flanking the fluorophore, such as the stacking of the Trp37 residue with the flanking guanine (41). Smaller fluorescence increases (<3-fold) were observed when 2Ap was more distant from a TG motif (AAApCTGAATGAC and AATGACT-GAAPAC), in line with the preferential binding of NC to TG motifs (22, 23). Noticeably, NC does not always induce

a 2Ap fluorescence increase on interaction with ODNs, since a 10% fluorescence decrease was reported to accompany NC binding to a bulged 2Ap residue in a stem-loop motif (52). In this case, the 2Ap residue was only poorly stacking with its neighbors in the free ODN, as suggested by its high fluorescence intensity. As a consequence, the NC-induced changes in the ODN flexibility and the local mobility of the bases (including 2Ap) could only marginally affect the stacking and, thus, the fluorescence intensity of this bulged 2Ap.

All studied dodecanucleotides were found to bind two NC molecules at physiological salt concentrations, confirming the preferential binding of NC to the TG motifs and an occluded binding site size of five to six bases (22, 23, 30, 53). Importantly, strong similarities in the binding curves were observed when 2Ap residues were at close positions in a given ODN (for instance, at positions 2 and 5 in ODN2). In contrast, significant shifts in the binding curves occurred when 2Ap residues were distant (for instance, at positions 2 and 9 in ODN2). This clearly suggests that 2Ap at a given position responds mainly to the binding of NC to its closest TG motif, enabling site-specific investigations of NC binding to ODNs with multiple binding sites. Using two parallel titrations of the same ODN labeled with 2Ap at distant positions and assuming independent binding sites, we obtained the microscopic binding constants of the two TG-containing binding motifs. To the best of our knowledge, this constitutes the first direct characterization of two NC binding sites on the same ODN. This approach appears to be an interesting alternative to NMR titrations that require much higher ODN concentrations and thus can only be used for systems with rather low affinities. However, proper selection of the labeled position is critical in the proposed methodology since in the case of ODN3, the 2Ap label at position 7, between the two close TG motifs, was found to be sensitive to the binding of NC to both sites.

Using this approach based on 2Ap labeling at different positions, we found that the affinities of the TG-containing motifs range from 0.4×10^6 to 1.8×10^6 M⁻¹, suggesting that the binding of NC to these motifs is context-dependent. In this respect, ODN1 composed of two TGACCG motifs is of special interest. We were expecting the same microscopic affinities for the two binding sites. In addition, the lifetime distributions for 2Ap at positions 3 and 9 were very similar, being consistent with the same local microenvironment. However, the affinity constant for the proximal binding site is substantially lower than that for the distal binding site. This difference may be explained by the involvement of two guanines (G⁶ and G⁸) in NC binding to the distal binding site. Indeed, two alternative modes of NC binding to target ODNs are known from NMR structures. The first one, found for d(ACGCC) or (-)PBS, involves the interaction of both zinc fingers (F1 and F2) with a single G residue and its upstream base (28, 30). In the second mode found for SL3 (29), SL2 (32), or an U5 ODN from the 5'-untranslated region of HIV-1 (31), F1 and F2 interact through stacking of their aromatic amino acid (Phe16 in F1 and Trp37 in F2) with the two guanines of a GXG motif. This last type of binding is possible only for the distal motif of ODN1 where the G⁶ and G⁸ residues are properly spaced to allow their interaction with the two NC fingers. In contrast, only one guanine (G²) is available for the binding of NC to the

proximal binding site, which may explain its lower affinity, in line with the 1–2 order of magnitude lower affinities of NC for AATGCC (23, 41), the TG-containing loop of (-)PBS (28, 37), or SL3 and U5 mutants, where one G of the GXG motif has been substituted (26, 31) as compared to the affinity of NC for TGTGCC, (TG)₄, or the GXG binding motif of the U5 sequence and the SL2 and SL3 loops (23, 27, 31, 32, 35). In addition, the stability of the complex of NC with the proximal binding site of ODN1 may be further limited by the 5'-position of the TG motif.

For ODN2 and ODN3, the differences between the binding affinities of the proximal and distal binding sites were less pronounced, which does not allow us to draw clear conclusions. A larger set of ODNs with systematically varied bases would be needed to gain further information about the molecular determinants that modulate the affinity of NC for the individual binding sites. Nevertheless, since the affinities of the two sites differ by a factor of less than 4 in all three ODNs and were close to the affinities of the corresponding hexanucleotides, this excludes strong positive or negative cooperativity for the binding of NC to these dodecanucleotides. Such a limited cooperativity was already reported for the binding of NC to (TG)₄ (35), as well as to longer sequences, such as polyA (38), tRNA^{Lys3} (39), or (-)PBS (28). Since the TG binding motifs in ODN2 and ODN3 are separated by only two bases, the two bound NC molecules are necessarily very close, like on the short (TG)₄ sequence. In this respect, the weakly cooperative binding of NC suggests that no strong attraction or repulsion occurs between the bound NC molecules. Both the absence of strong negative cooperativity and the possibility of NC binding at very close binding sites are fully consistent with the extensive coating of the genomic RNA and the proviral DNA by NC to protect them against nucleases (54–57). In addition, the absence of strong positive cooperativity is in line with the required ability of NC to rapidly dissociate from the nucleic acids to exert its chaperone properties (58). Another consequence of our data is that ODNs able to bind several NC molecules with high affinity could be envisioned as therapeutic tools for dissociating the NC molecules bound to the HIV-1 genomic RNA and proviral DNA. Due to the critical role of NC in the HIV-1 life cycle, this dissociation should lead to a strong decrease in HIV-1 infectivity.

In conclusion, NC binding to the individual binding sites of dodecanucleotides was site-specifically characterized by using 2Ap fluorescence. This approach allowed us to determine the NC-induced changes in the local and overall ODN dynamics on the picosecond to nanosecond time scale. Furthermore, it allowed us to determine the microscopic binding constants of the individual binding sites. This last point constitutes a substantial advantage of this 2Ap-based approach over other commonly used techniques such as surface plasmon resonance, fluorescence anisotropy, isothermal titration calorimetry, mass spectroscopy, and ELISA, which do not respond site-specifically. This approach may be further extended to investigate the binding of NC (or any nucleic acid binding protein) to ODNs containing more than two binding sites, provided that the latter are not overlapping.

ACKNOWLEDGMENT

We thank H. de Rocquigny for peptide synthesis and G. Krishnamoorthy for providing the software for analysis of time-resolved fluorescence and anisotropy decay data.

SUPPORTING INFORMATION AVAILABLE

Fluorescence lifetime experiments with 2Ap-labeled ODNs and time-resolved intensity parameters for the 2Ap-substituted dodecanucleotides (Table 1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Darlix, J. L., Cristofari, G., Rau, M., Pechoux, C., Berthoux, L., and Roques, B. (2000) Nucleocapsid protein of human immunodeficiency virus as a model protein with chaperoning functions and as a target for antiviral drugs. *Adv. Pharmacol.* **48**, 345–372.
- Clever, J., Sasseti, C., and Parslow, T. G. (1995) RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* **69**, 2101–2109.
- Bampi, C., Jacquenet, S., Lener, D., Decimo, D., and Darlix, J. L. (2004) The chaperoning and assistance roles of the HIV-1 nucleocapsid protein in proviral DNA synthesis and maintenance. *Int. J. Biochem. Cell Biol.* **36**, 1668–1686.
- Levin, J. G., Guo, J., Rouzina, I., and Musier-Forsyth, K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: Critical role in reverse transcription and molecular mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **80**, 217–286.
- Rein, A., Henderson, L. E., and Levin, J. G. (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: Significance for viral replication. *Trends Biochem. Sci.* **23**, 297–301.
- D'Souza, V., and Summers, M. F. (2005) How retroviruses select their genomes. *Nat. Rev. Microbiol.* **3**, 643–655.
- Mely, Y., De Rocquigny, H., Morellet, N., Roques, B. P., and Gerard, D. (1996) Zinc binding to the HIV-1 nucleocapsid protein: A thermodynamic investigation by fluorescence spectroscopy. *Biochemistry* **35**, 5175–5182.
- Mely, Y., Jullian, N., Morellet, N., De Rocquigny, H., Dong, C. Z., Piemont, E., Roques, B. P., and Gerard, D. (1994) Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer. *Biochemistry* **33**, 12085–12091.
- Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P. (1992) Determination of the structure of the nucleocapsid protein NcP7 from the human immunodeficiency virus type 1 by IH NMR. *EMBO J.* **11**, 3059–3065.
- Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., Jr., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R. C., and Hare, D. R. (1992) Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. *Protein Sci.* **1**, 563–574.
- Aldovini, A., and Young, R. A. (1990) Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* **64**, 1920–1926.
- Demene, H., Dong, C. Z., Ottmann, M., Rouyez, M. C., Jullian, N., Morellet, N., Mely, Y., Darlix, J. L., Fournie-Zaluski, M. C., Saragosti, S., et al. (1994) ¹H NMR structure and biological studies of the His23 → Cys mutant nucleocapsid protein of HIV-1 indicate that the conformation of the first zinc finger is critical for virus infectivity. *Biochemistry* **33**, 11707–11716.
- Dorfman, T., Luban, J., Goff, S. P., Haseltine, W. A., and Gottlinger, H. G. (1993) Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* **67**, 6159–6169.
- Gorelick, R. J., Gagliardi, T. D., Bosche, W. J., Wiltrout, T. A., Coren, L. V., Chabot, D. J., Lifson, J. D., Henderson, L. E., and Arthur, L. O. (1999) Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology* **256**, 92–104.
- Gorelick, R. J., Nigida, S. M., Jr., Bess, J. W., Jr., Arthur, L. O., Henderson, L. E., and Rein, A. (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* **64**, 3207–3211.
- Guo, J., Wu, T., Anderson, J., Kane, B. F., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Levin, J. G. (2000) Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J. Virol.* **74**, 8980–8988.
- Ottmann, M., Gabus, C., and Darlix, J. L. (1995) The central globular domain of the nucleocapsid protein of human immunodeficiency virus type 1 is critical for virion structure and infectivity. *J. Virol.* **69**, 1778–1784.
- Ramboarina, S., Morellet, N., Fournie-Zaluski, M. C., and Roques, B. P. (1999) Structural investigation on the requirement of CCHH zinc finger type in nucleocapsid protein of human immunodeficiency virus 1. *Biochemistry* **38**, 9600–9607.
- Stote, R. H., Kellenberger, E., Muller, H., Bombarda, E., Roques, B. P., Kieffer, B., and Mely, Y. (2004) Structure of the His44 → Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: A combined NMR, molecular dynamics simulation, and fluorescence study. *Biochemistry* **43**, 7687–7697.
- Thomas, J. A., and Gorelick, R. J. (2008) Nucleocapsid protein function in early infection processes. *Virus Res.* **134**, 39–63.
- de Rocquigny, H., Shvadchak, V., Avilov, S., Dong, C. Z., Dietrich, U., Darlix, J. L., and Mely, Y. (2008) Targeting the viral nucleocapsid protein in anti-HIV-1 therapy. *Mini-Rev. Med. Chem.* **8**, 24–35.
- Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J. Virol.* **72**, 1902–1909.
- Vuilleumier, C., Bombarda, E., Morellet, N., Gerard, D., Roques, B. P., and Mely, Y. (1999) Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NcP7: A fluorescence study. *Biochemistry* **38**, 16816–16825.
- You, J. C., and McHenry, C. S. (1993) HIV nucleocapsid protein. Expression in *Escherichia coli*, purification, and characterization. *J. Biol. Chem.* **268**, 16519–16527.
- Hagan, N., and Fabris, D. (2003) Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry* **42**, 10736–10745.
- Paoletti, A. C., Shubsda, M. F., Hudson, B. S., and Borer, P. N. (2002) Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry* **41**, 15423–15428.
- Shubsda, M. F., Paoletti, A. C., Hudson, B. S., and Borer, P. N. (2002) Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry* **41**, 5276–5282.
- Bourbigot, S., Ramalanjaona, N., Boudier, C., Salgado, G. F., Roques, B. P., Mely, Y., Bouaziz, S., and Morellet, N. (2008) How the HIV-1 Nucleocapsid Protein Binds and Destabilises the (–)Primer Binding Site During Reverse Transcription. *J. Mol. Biol.* **383**, 1112–1128.
- De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science* **279**, 384–388.
- Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M. C., and Roques, B. P. (1998) Structure of the complex between the HIV-1 nucleocapsid protein NcP7 and the single-stranded pentanucleotide d(ACGCC). *J. Mol. Biol.* **283**, 419–434.
- Spriggs, S., Garyu, L., Connor, R., and Summers, M. F. (2008) Potential Intra- and Intermolecular Interactions Involving the Unique-5' Region of the HIV-1 5'-UTR. *Biochemistry* **47**, 13064–13073.
- Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K. J., Wu, Z. R., and Summers, M. F. (2000) NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J. Mol. Biol.* **301**, 491–511.
- Beltz, H., Clauss, C., Piemont, E., Ficheux, D., Gorelick, R. J., Roques, B., Gabus, C., Darlix, J. L., de Rocquigny, H., and Mely, Y. (2005) Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J. Mol. Biol.* **348**, 1113–1126.

34. Egele, C., Piemont, E., Didier, P., Ficheux, D., Roques, B., Darlix, J. L., de Rocquigny, H., and Mely, Y. (2007) The single-finger nucleocapsid protein of moloney murine leukemia virus binds and destabilizes the TAR sequences of HIV-1 but does not promote efficiently their annealing. *Biochemistry* 46, 14650–14662.
35. Fisher, R. J., Fivash, M. J., Stephen, A. G., Hagan, N. A., Shenoy, S. R., Medaglia, M. V., Smith, L. R., Worthy, K. M., Simpson, J. T., Shoemaker, R., McNitt, K. L., Johnson, D. G., Hixson, C. V., Gorelick, R. J., Fabris, D., Henderson, L. E., and Rein, A. (2006) Complex interactions of HIV-1 nucleocapsid protein with oligonucleotides. *Nucleic Acids Res.* 34, 472–484.
36. Beltz, H., Piemont, E., Schaub, E., Ficheux, D., Roques, B., Darlix, J. L., and Mely, Y. (2004) Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J. Mol. Biol.* 338, 711–723.
37. Egele, C., Schaub, E., Ramalanjaona, N., Piemont, E., Ficheux, D., Roques, B., Darlix, J. L., and Mely, Y. (2004) HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences and chaperones their kissing interactions. *J. Mol. Biol.* 342, 453–466.
38. Khan, R., and Giedroc, D. P. (1994) Nucleic acid binding properties of recombinant Zn²⁺ HIV-1 nucleocapsid protein are modulated by COOH-terminal processing. *J. Biol. Chem.* 269, 22538–22546.
39. Mely, Y., de Rocquigny, H., Sorinas-Jimeno, M., Keith, G., Roques, B. P., Marquet, R., and Gerard, D. (1995) Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3Lys), in vitro, is essentially not specific. *J. Biol. Chem.* 270, 1650–1656.
40. Urbaneja, M. A., Kane, B. P., Johnson, D. G., Gorelick, R. J., Henderson, L. E., and Casas-Finet, J. R. (1999) Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: Elucidation of the structural determinants for function. *J. Mol. Biol.* 287, 59–75.
41. Avilov, S. V., Piemont, E., Shvadehak, V., de Rocquigny, H., and Mely, Y. (2008) Probing dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-aminopurine. *Nucleic Acids Res.* 36, 885–896.
42. de Rocquigny, H., Ficheux, D., Gabus, C., Fournie-Zaluski, M. C., Darlix, J. L., and Roques, B. P. (1991) First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem. Biophys. Res. Commun.* 180, 1010–1018.
43. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415.
44. Ward, D. C., Reich, E., and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* 244, 1228–1237.
45. Ben Gaied, N., Glasser, N., Ramalanjaona, N., Beltz, H., Wolff, P., Marquet, R., Burger, A., and Mely, Y. (2005) 8-Vinyldeoxyadenosine, an alternative fluorescent nucleoside analog to 2'-deoxyribosyl-2-aminopurine with improved properties. *Nucleic Acids Res.* 33, 1031–1039.
46. Nag, N., Ramreddy, T., Kombrabail, M., Krishna Mohan, P. M., D'Souza, J., Rao, B. J., Duportail, G., Mely, Y., and Krishnamoorthy, G. (2006) in *Reviews in Fluorescence* (Geddes, C., and Lakowicz, J. R., Eds.) pp 311–340, Springer Science, New York.
47. Guest, C. R., Hochstrasser, R. A., Sowers, L. C., and Millar, D. P. (1991) Dynamics of mismatched base pairs in DNA. *Biochemistry* 30, 3271–3279.
48. Lakowicz, J. R. (1999) *Principles of Fluorescence Spectroscopy*, 2nd ed., Kluwer Academic/Plenum Publishers, Dordrecht, The Netherlands.
49. Ramreddy, T., Rao, B. J., and Krishnamoorthy, G. (2007) Site-Specific Dynamics of Strands in ss- and dsDNA As Revealed by Time-Domain Fluorescence of 2-Aminopurine. *J. Phys. Chem. B* 111, 5757–5766.
50. Jean, J. M., and Hall, K. B. (2004) Stacking-unstacking dynamics of oligodeoxynucleotide trimers. *Biochemistry* 43, 10277–10284.
51. Larsen, O. F., van Stokkum, I. H., Gobets, B., van Grondelle, R., and van Amerongen, H. (2001) Probing the structure and dynamics of a DNA hairpin by ultrafast quenching and fluorescence depolarization. *Biophys. J.* 81, 1115–1126.
52. Rist, M. J., and Marino, J. P. (2002) Mechanism of nucleocapsid protein catalyzed structural isomerization of the dimerization initiation site of HIV-1. *Biochemistry* 41, 14762–14770.
53. Karpel, R. L., Henderson, L. E., and Oroszlan, S. (1987) Interactions of retroviral structural proteins with single-stranded nucleic acids. *J. Biol. Chem.* 262, 4961–4967.
54. Buckman, J. S., Bosche, W. J., and Gorelick, R. J. (2003) Human immunodeficiency virus type 1 nucleocapsid Zn²⁺ fingers are required for efficient reverse transcription, initial integration processes, and protection of newly synthesized viral DNA. *J. Virol.* 77, 1469–1480.
55. Gao, K., Gorelick, R. J., Johnson, D. G., and Bushman, F. (2003) Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro. *J. Virol.* 77, 1598–1603.
56. Krishnamoorthy, G., Roques, B., Darlix, J. L., and Mely, Y. (2003) DNA condensation by the nucleocapsid protein of HIV-1: A mechanism ensuring DNA protection. *Nucleic Acids Res.* 31, 5425–5432.
57. Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J. L. (1995) Formation of stable and functional HIV-1 nucleoprotein complexes in vitro. *J. Mol. Biol.* 252, 563–571.
58. Cruceanu, M., Gorelick, R. J., Musier-Forsyth, K., Rouzina, L., and Williams, M. C. (2006) Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. *J. Mol. Biol.* 363, 867–877.
59. Xu, D., Evans, K. O., and Nordlund, T. M. (1994) Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry* 33, 9592–9599.

BI8022366

Supporting Information

Fluorescence lifetime experiments with 2Ap-labeled ODNs.

Good fittings ($\chi^2 < 1.4$) of the intensity decay curves of the 2Ap-labeled ODNs required four components ranging from ≈ 100 ps to 7.3 ns, showing a large conformational heterogeneity of 2Ap the studied ODNs (Table 1 in Supporting information). Multiple lifetimes are commonly observed 2Ap-labeled ODNs and are attributed to different extents of stacking of 2Ap with its neighbours (47, 50, 59). By analogy with other single-stranded 2Ap-labeled ODNs (41, 49, 50), these lifetimes could arise from picosecond-nanosecond conformational fluctuations of the 2Ap-labeled dodecanucleotides in the excited state that drive 2Ap from an unstacked to a stacked conformation. Moreover, the decrease of the mean lifetime with respect to free 2Ap was far less than the 11- to >100-fold decrease in the corresponding quantum yields (Table 1), indicating a significant subpopulation “dark” species with apparently null lifetime (α_0). Since high viscosity (77% glycerol) increased the mean lifetime value by 2-3 fold and strongly decreased the α_0 value (data not shown), the dark species was attributed to ultra-fast dynamic quenching.

NC binding strongly decreased the population α_0 of the dark species, indicating that NC prevents stacking of 2Ap with its neighbors. Comparison of the various ODNs further revealed that the differences in the quantum yields with NC are mainly due to the α_0 values. Two major classes of complexes could be distinguished. The first class is composed of complexes in which binding of NC resulted in the complete disappearance of the dark species and a strong shift towards longer lived conformations. In the second, the ultra-fast quenching component remains after NC binding but is strongly reduced. All together, the steady-state and time-resolved data for the dodecanucleotides suggest that NC slows down the collisions of 2Ap with its neighbor bases and strongly decreases the population of the fully stacked conformations of the probe.

Table 1. Time-resolved intensity parameters for the 2Ap-substituted dodecanucleotides¹.

| Sequence | α_0 | τ_1 , ns | α_1 | τ_2 , ns | α_2 | τ_3 , ns | α_3 | τ_4 , ns | α_4 | $\langle \tau \rangle$, ns |
|---|------------|---------------|------------|---------------|------------|---------------|------------|---------------|------------|-----------------------------|
| free 2Ap | | | | | | | | 10.2 | 1.00 | 10.2 |
| ODN1Ap3 TG Ap CCGTGACCG | 0.90 | 0.07 | 0.05 | 1.0 | 0.02 | 2.8 | 0.02 | 7.3 | 0.01 | 1.1 |
| TG Ap CCGTGACCG + NC | 0.55 | 0.15 | 0.20 | 1.2 | 0.05 | 2.9 | 0.09 | 7.8 | 0.11 | 2.6 |
| ODN1Ap9 TGACCGTG Ap CCG | 0.86 | 0.08 | 0.07 | 0.26 | 0.04 | 2.1 | 0.02 | 6.6 | 0.01 | 0.95 |
| TGACCGTG Ap CCG + NC | 0.31 | 0.12 | 0.24 | 0.9 | 0.13 | 3.3 | 0.15 | 8.1 | 0.17 | 2.9 |
| ODN2Ap2 A Ap TGACTGAAAC | 0.24 | 0.10 | 0.40 | 0.7 | 0.23 | 2.6 | 0.09 | 6.6 | 0.04 | 0.9 |
| A Ap TGACTGAAAC + NC | 0.00 | 0.20 | 0.40 | 1.0 | 0.29 | 3.2 | 0.17 | 7.9 | 0.14 | 2.7 |
| ODN2Ap5 AATG Ap CTGAAAC | 0.74 | 0.06 | 0.18 | 0.50 | 0.04 | 2.2 | 0.03 | 5.8 | 0.01 | 0.67 |
| AATG Ap CTGAAAC + NC | 0.03 | 0.13 | 0.40 | 0.9 | 0.15 | 3.7 | 0.18 | 8.4 | 0.24 | 3.0 |
| ODN2Ap10 AATGACTGA Ap AC | 0.39 | 0.15 | 0.14 | 1.0 | 0.31 | 2.1 | 0.13 | 5.8 | 0.03 | 1.3 |
| AATGACTGA Ap AC + NC | 0.26 | 0.14 | 0.18 | 1.0 | 0.23 | 2.7 | 0.24 | 7.4 | 0.09 | 2.1 |
| ODN3Ap3 AA Ap CTGAATGAC | 0.24 | 0.17 | 0.39 | 1.1 | 0.28 | 2.7 | 0.07 | 5.9 | 0.02 | 0.9 |
| AA Ap CTGAATGAC + NC | 0.00 | 0.15 | 0.22 | 1.0 | 0.34 | 3.0 | 0.25 | 7.8 | 0.19 | 2.6 |
| ODN3Ap7 AA Ap CTGA Ap ATGAC | 0.85 | 0.12 | 0.04 | 1.3 | 0.07 | 2.7 | 0.03 | 6.6 | 0.01 | 1.5 |
| AA Ap CTGA Ap ATGAC + NC | 0.05 | 0.05 | 0.50 | 1.6 | 0.11 | 4.0 | 0.13 | 8.9 | 0.21 | 2.7 |
| ODN3Ap11 AA Ap CTGAATG Ap C | 0.76 | 0.09 | 0.13 | 0.9 | 0.04 | 2.6 | 0.05 | 5.5 | 0.02 | 1.2 |
| AA Ap CTGAATG Ap C + NC | 0.09 | 0.14 | 0.29 | 1.4 | 0.17 | 3.0 | 0.25 | 6.7 | 0.20 | 2.7 |

5.3 Rôle des doigts de zinc de la NCp7 dans le mécanisme de l'hybridation de (+)/(-)PBS au cours du second saut de brin

Après avoir caractérisé de manière site-spécifique la liaison de la NCp7 à l'aide de la 2-Ap, nous nous sommes intéressés dans un second temps à l'effet de la NCp7 sur la boucle du *Primer Binding Site* (PBS) dans le cadre du second saut de brin.

PBS est une séquence de 18 nucléotides se structurant en tige-boucle et intervenant à la fois au cours de l'initiation de la transcription inverse et du second transfert de brin. Lors de cette dernière étape, les séquences complémentaires de PBS situées à l'extrémité 3' des brins (-) et (+) de l'ADN proviral en cours de synthèse, respectivement nommées (-)PBS et (+)PBS, vont s'hybrider afin de permettre la poursuite de la transcription inverse. La NCp7 est nécessaire à cette étape puisqu'elle facilite l'hybridation de (-)PBS avec son complémentaire [281]. La NCp7 induit un *switch* dans le mécanisme d'hybridation de ces deux séquences. En effet, alors que l'hybridation de (+)/(-)PBS en absence de protéine implique une étape de nucléation au niveau de l'extrémité protrudente simple brin du PBS, la NCp7 promeut une hybridation *via* les boucles des PBS complémentaires (figure 5.5).

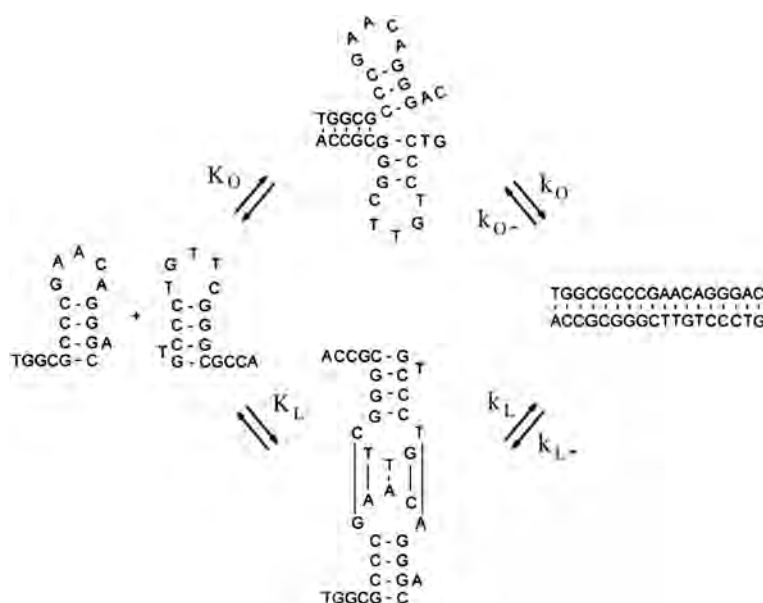


FIGURE 5.5 – Illustration des deux voies d'hybridation de (+)/(-)PBS [281]

5.3 Rôle des doigts de zinc de NCp7 dans le second saut de brin

La résolution de la structure RMN du complexe $\Delta P(-)PBS$ avec la NC(11-55) [227] a montré que la NC induit un changement de la conformation de la boucle, notamment en étirant la boucle et en augmentant l'exposition des bases de celle-ci vers le solvant. Ce remodelage de la boucle de PBS pourrait favoriser les interactions boucle-boucle entre les séquences PBS complémentaires.

Nous avons étudié les paramètres de fluorescence résolue en temps d'un $\Delta P(-)PBS$, une séquence PBS délétée de son extrémité 3'_{term} simple-brin, qui a été modifiée en position 6, 8 et 10 de la boucle avec une 2-aminopurine (2-AP). Nous avons mis en relation les informations sur la dynamique des bases de la boucle de PBS avec des données mécanistiques, cinétiques et thermodynamiques de la réaction d'hybridation des (+)/(-)PBS. Le *switch* induit par NCp7 est associée à une restriction importante de la dynamique locale des bases de la boucle du PBS à l'échelle de la picoseconde-nanoseconde, ce qui indique que la NCp7 est en mesure de "geler" les conformations compétentes pour une réaction d'hybridation par les boucles. En utilisant différents mutants NCp7, nous avons constaté que les modifications de la structure et de la dynamique de la boucle de PBS qui régissent le mécanisme de l'hybridation sont strictement dépendantes de l'intégrité de la plateforme hydrophobe qui se forme en surface des deux doigts de zinc repliés. Ces observations montrent au niveau moléculaire que les propriétés d'ARN chaperonne de la NCp7 portées par les deux doigts de zinc sont nécessaires pour assurer la spécificité de la réaction d'hybridation impliquée dans le deuxième transfert de brin. Elles soulignent encore davantage le rôle clé de la NCp7 dans l'organisation temporelle de la transcription inverse et ses conséquences quant à la production d'un ADN proviral stable.

Ces travaux sont actuellement soumis à publication.

5.4

Publication VI

The HIV-1 nucleocapsid protein zinc-fingers are mandatory to induce specificity in the annealing of the (+)/(-) copies of the primer binding site during the second strand transfer.

Julien Godet^{1,2,#}, Nick Ramalanjaona^{1,#}, Kamal K Sharma¹, Ludovic Richert¹, Hugues de Rocquigny¹, Guy Duportail¹, Yves Mély^{1*}

¹ Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74 route du Rhin, 67401 Illkirch, France

² Medical Information and Biostatistics Department, Strasbourg University Hospital, 67000 Strasbourg, France

NR and JG contributed equally to this work

* Corresponding author

Tel : +33 (0)3 68 85 42 63 Fax : +33 (0)3 68 85 43 12

e-mail: yves.mely@unistra.fr

Abstract

Conversion of the human immunodeficiency virus type I (HIV-1) genomic RNA into the proviral DNA by reverse transcriptase involves two obligatory strand transfers that are chaperoned by the nucleocapsid protein (NCp7). The second strand transfer relies on the annealing of the (-) and (+) DNA copies of the primer binding site, (-)PBS and (+)PBS, which fold into complementary stem-loops. NCp7 increases the (-)PBS/(+)PBS annealing kinetics and modifies the mechanistic pathway of the annealing reaction by strongly activating the loop-loop kissing pathway that poorly contributes in the absence of NCp7. This switch in the mechanistic pathway likely results from the NMR-described conformational changes of the loop in response to NCp7 binding. To site-selectively characterize the dynamics of the loop bases within NCp7/PBS complexes, we investigated the time-resolved fluorescence parameters of a $\Delta P(-)PBS$, a sequence lacking the 3' single strand overhang of (-)PBS, which has been modified in position 6, 8 and 10 of the loop with 2-aminopurine(2-Ap), a fluorescent nucleoside analogue sensitive to its environment. We further correlated the dynamic information of the PBS loop bases to the kinetic and thermodynamic data of the (+)/(-) annealing reaction. The NCp7-induced switch of (+)/(-)PBS annealing towards the loop pathway was associated to a dramatic restriction of the local dynamics of the PBS loop bases at the picosecond and nanosecond timescale, indicating that NCp7 was able to “freeze” the conformations competent for an annealing reaction through the loops. Using various NCp7 mutants, we found that the modifications of the PBS loop structure and dynamics that govern the annealing mechanism were strictly dependent on the integrity of the hydrophobic platform at the top of the two folded zinc fingers. These observations evidenced at the molecular level that the chaperone properties of NCp7 carried by the two zinc fingers are required to ensure the specificity of the annealing reaction involved in the second strand transfer. This further underlines the key role of NCp7 for appropriate reverse transcription timing and production of proviral DNA.

Introduction

Human immunodeficiency virus type 1 (HIV-1) viral DNA synthesis is a complex multi-step process catalyzed by the viral reverse transcriptase (RT). Conversion of the genomic RNA into proviral DNA requires two obligatory DNA transfer reactions. During the first strand transfer, the minus-strand strong-stop DNA ((-)ssDNA) is translocated to the 3' end of the viral RNA genome, in a reaction mediated by base-pairing of the repeat sequences at the 3' ends of the RNA and DNA reactants, to allow the reaction to resume up to the 5' end of the primer binding site, (-) PBS. The second strand transfer relies on (-)PBS annealing to the (+)PBS sequence located in the plus strong stop DNA (ssDNA(+)), the copy of the minus strand synthesized by RT (1,2). The annealing of the two complementary PBS DNA stem-loops enables the RT to resume and complete the DNA synthesis. These two obligatory strand transfers are chaperoned by the HIV-1 nucleocapsid protein (NCp7), a potent nucleic acid chaperone that plays a major role in the viral replication cycle (3-6). The mature NCp7 is a 55 amino-acids protein encoded by the HIV-1 Gag polyprotein. It contains two highly conserved CCHC zinc fingers that coordinate zinc ions with high affinity (7). The two folded zinc fingers are connected by a highly conserved basic sequence and are flanked by N- and C-terminal basic domains. NCp7 catalyzes the annealing reactions of (+)/(-)PBS complementary sequences (8,9) and of (+)PBS RNA with tRNA^{Lys-3} (10-12), in line with its ability to chaperone the rearrangement of nucleic acids into their most thermodynamically stable conformations (6,13). In the absence of NCp7, (+)PBS can spontaneously anneal to (-)PBS (8) *in vitro*. This annealing reaction proceeds mainly through the single-strand overhangs of the PBS sequences while nucleation of the final duplex through loop-loop interaction appears marginal. NC(11-55), a mutant that corresponds to the zinc finger domain of NCp7, increases (-)/(+)PBS annealing kinetics by about 6-fold, while the native NCp7 further increases this reaction by tenfold. Interestingly, while NCp7 accelerates the annealing reaction by only a factor of about two for the pathway via the single-stranded PBS overhangs; it accelerates the pathway through the loops, by several orders of magnitude. As a consequence, the annealing of (-)/(+)PBS through the loops becomes the major pathway in the presence of NCp7. This switch in the mechanism of the annealing reaction likely results from significant changes of the loop structure and/or dynamics in response to NCp7 binding. Information on the structural changes of the loop was obtained by solving with 1H NMR the structure of NCp7 and NCp7(12-55) peptides complexed to $\Delta P(-)PBS$, a (-)PBS derivative without the 3' protruding sequence (14). The two peptides were shown to bind in a slow exchange equilibrium to the 5' end of the loop (5-CTG-7) of $\Delta P(-)PBS$. This binding is mediated by the hydrophobic plateau, involving the Val13, Phe16, Thr24, Ala25, Trp37, Gln45 and Met46 residues that forms on the top of the folded and spatially close zinc fingers (15-17). The Phe16 and Trp37 residues insert between T6 and G7 bases, allowing a strong stacking of the Trp37 with G7 that

stabilizes the complex (18) and directs nucleotides T6 and G7 toward the exterior of the loop. NCp7(12-55) was also shown to bind $\Delta P(-)PBS$ on an additional site (10-CGG-12) in an intermediate-exchange equilibrium on the NMR timescale. The binding of NCp7 on the 5' end of $\Delta P(-)PBS$ loop induces the stretching of the loop, which increases the accessibility of the 8-TTC-10 nucleobases and slightly destabilizes the upper base pair of the stem. As a consequence, NCp7 is thought to both favor the kissing interactions between $(-)PBS$ and $(+)PBS$ loops and help disrupting the stem to convert the kissing complex into the final extended duplex.

To gain further insight into the molecular mechanism and characterize the protein determinants responsible for the NCp7-promoted annealing of $(-)/(+)PBS$, we correlated in this study the annealing mechanism investigated by a PBS mutational analysis with thermodynamic and kinetic information obtained in the presence of various NCp7 mutants. To site-specifically investigate the dynamic changes of both $(-)PBS$ and $(+)PBS$ loops in response to NCp7 binding, the loops were modified with 2-aminopurine (2-Ap), a fluorescent adenine analogue sensitive to its environment (19), previously used to show that binding of NCp7 strongly decreases the dynamics of small linear oligonucleotides (ODN) (20,21). In line with its weak destabilizing activity on the PBS stem (22,23), we found that NCp7 does not modify the activation energy of the $(+)(-)PBS$ annealing reaction, which was thus associated to the melting of the stem base pairs. Moreover, the aggregating properties of NCp7 and its mutants were shown to strongly influence the annealing rates but only the mutants with intact zinc fingers were able to modify the structure and restrict the dynamics of both PBS loops and, switch the annealing mechanism towards the loop-loop kissing pathway. Our data evidenced that the hydrophobic plateau at the top of the folded zinc fingers is needed to direct the formation of competent PBS loop conformations for annealing, ensuring the specificity of the annealing reaction involved in the second strand transfer.

Materials and Methods

Materials

The NC(1-55), NC(11-55), (SSHS)₂NC(11-55), L₃₇NC(11-55) and A₁₆NC(11-55) peptides were synthesized on a Applied Biosystems A433 peptide synthesizer, as described (24,25). Unmodified or labelled ODNs were synthesized and HPLC- or PAGE-purified by IBA GmbH Nucleic Acids Product Supply (Germany). A 2'-deoxyribosyl-2-aminopurine (2-Ap) was selectively introduced at different positions (6, 7 or 10) within the $\Delta P(-)$ PBS loop and at position 11 of $\Delta P(+)$ PBS, substituting the corresponding natural base. By monitoring the Trp37 fluorescence quenching on ODN binding (26,27), this base substitution was found to not affect the NCp7 binding parameters (see supplementary materials). Doubly-labelled (+)PBS sequences were modified at their 5' terminus with 6-carboxyrhodamine (Rh6G) and their 3' terminus with 4-(4'-dimethylaminophenylazo)benzoic acid (DABCYL), via an amino linker with a six carbon spacer arm. All experiments were performed in 25 mM TRIS-HCl, pH 7.5, 30 mM NaCl and 0.2 mM MgCl₂ at 20°C, unless specified otherwise.

Steady-state fluorescence spectroscopy

Fluorescence emission spectra were recorded with a Fluorolog or a Fluoromax-3 spectrofluorimeter (Jobin Yvon) equipped with thermostated cell compartments. All fluorescence intensities were corrected for screening effects, buffer emission and lamp fluctuations. The quantum yield was calculated using free 2-Ap riboside as a reference (quantum yield = 0.68 (19)). 2-Ap was excited at 315nm.

Kinetics of (-)/(+)PBS annealing was monitored in real-time using fluorescent doubly-labeled (+)PBS sequences and non-labeled (-)PBS. Excitation and emission wavelengths were at 520 nm and 550 nm, respectively, to monitor the fluorescence restoration of Rh6G resulting from the formation of the (-)/(+)PBS duplex. Concentrations of 5'Rh6G-(+)PBS-3'DABCYL and (-)PBS were 10 nM and 100 nM to 1.1 μ M, respectively, to ensure pseudo-first order conditions. Both reactants (of identical volume) coated by NC peptides, were mixed together to trigger the reaction. The apparent rate constants k_{obs} were determined from the kinetic data, as previously described (8). All fitting procedures were carried out with the Microcal Origin 6.1 software based on non-linear least-squares methods applying the Levenberg-Marquardt algorithm.

Time-resolved fluorescence measurements

Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique. Excitation pulses were generated by a pulse-picked frequency-tripled Ti-sapphire laser (Tsunami, Spectra Physics) pumped by a Millennia X laser (Spectra Physics) (28).

Excitation wavelength was set at 315 nm, with a repetition rate of 4 MHz. The fluorescence emission was collected through a polarizer set at magic angle and a 16 mm band-pass monochromator (Jobin Yvon) at 370 nm. The single-photon events were detected with a micro-channel plate photomultiplier (Hamamatsu) either coupled to a pulse pre-amplifier (Philips) and recorded on a multi-channel analyser (Ortec) calibrated at 25.5 ps/channel or coupled to a pulse pre-amplifier HFAC (Becker-Hickl) and recorded on a SPC-130 board (Becker-Hickl). The instrumental response function (IRF) was recorded using a polished aluminium reflector, and its full-width at half-maximum was ~ 40 ps. The mean lifetime $\langle\tau\rangle$ was calculated from the individual fluorescence lifetimes (τ_i) and their relative amplitudes (α_i) according to $\langle\tau\rangle = \sum \alpha_i \tau_i$. The population, α_0 , of dark species of 2-AP was calculated by: $\alpha_0 = 1 - \tau_{\text{free}} / (\tau_{\text{sample}} R_m)$, where τ_{free} is the lifetime of the free 2-AP, τ_{sample} is the measured mean lifetime of 2-AP within the ODN and R_m is the ratio of their corresponding quantum yields. The remaining amplitudes, α_{ic} were recalculated from the measured amplitudes according to $\alpha_{ic} = \alpha_i / (1 - \alpha_0)$.

Time-resolved anisotropy, resulting from the measurement of the fluorescence decay curves recorded in directions parallel (I_{\parallel}) and perpendicular (I_{\perp}) alternatively, to the excitation beam polarization, was analysed by the following equations:

$$I_{\parallel} = I(t)[1 + 2r(t)]/3 \quad \text{Equation (1)}$$

$$I_{\perp} = I(t)[1 - 2r(t)]/3 \quad \text{Equation (2)}$$

$$r(t) = \frac{I_{\parallel} - G \times I_{\perp}}{I_{\parallel} + 2G \times I_{\perp}} = r_0 \sum_i \beta_i \times \exp(-t/\Phi_i) \quad \text{Equation (3)}$$

where β_i are the amplitudes of the rotational correlation times Φ_i , and G is the geometry factor at the emission wavelength, determined in independent experiments. Theoretical values of the rotational correlation times were calculated from the molecular mass (M) of the complexes, assuming a spherical shape, by:

$$\Phi = \eta M (v + h) / RT$$

where η is the viscosity (assumed to be 1 cP), T is the temperature (293K), v is the specific volume of the particle (assumed to be 0.83 mL/g (29)), h is the hydration degree (assumed to be 0.2 mL/g for proteins) and R is the molar gas constant.

The cone semi-angle θ_0 of the local motion of 2AP (30) was calculated as:

$$\theta_0 = \cos^{-1}(0.5 \times (\sqrt{1 + 8S} - 1))$$

where S is the generalized order parameter

$$S = \left(\frac{\sum_{i=2}^n \beta^i}{\sum_{i=1}^n \beta^i} \right)^{0.5} \quad \text{Equation (4)}$$

in which β'_i are the amplitudes of the rotational correlation times, recalculated such that $\sum \beta'_i = r_0$. Since the shortest lifetime and rotational correlation times are in the range of ~ 0.08 - 0.25 ns, close to both the width of the IRF and the time per channel used, we measured the rotational correlation time (Φ_i) of the free 2-Ap riboside in sucrose-water mixture with increasing viscosities (η) in the range 1-8 cP to check the reliability of these values. The Φ_i values obtained at 20°C were 0.080 ± 0.001 ns, 0.237 ± 0.002 ns, 0.478 ± 0.029 ns and 0.850 ± 0.006 ns for viscosities of respectively 1.0 cP, 1.9 cP, 3.9 cP and 7.8 cP, as expected from the Stokes-Einstein relationship ($\Phi = \eta V / kT$) ($r = 0.994$; $p < 0.01$).

Time-resolved intensity and anisotropy data were treated according to the maximum entropy method (Pulse 5 software) (31,32) or according to a non-linear least-square analysis using a homemade software (kindly provided by G. Krishnamoorthy). In all cases, the χ^2 values were close to 1 and the weighted residuals as well as their autocorrelation were distributed randomly around 0, indicating an optimal fit.

Results

Characterization of the $\Delta P(-)$ PBS sequences labelled with 2-Ap

2-Ap was introduced in three positions of the $\Delta P(-)$ PBS loop: at the 5' end (Ap6) and in the middle of the loop (Ap8), within or close to the major binding site for NCp7, or at the 3' end of the loop (Ap10) within the second binding site of lower affinity for NCp7. Since 2-Ap is able to report locally on the dynamics of the ODN sequence in which it is inserted, we aimed to probe the whole loop with these three substitutions. Fluorescence spectra of the various 2-Ap-labeled $\Delta P(-)$ PBS sequences exhibited a maximum emission wavelength at about 365 nm, an emission wavelength slightly blue-shifted in respect to that of free 2-Ap and single-stranded ODNs labeled with 2-Ap (20,21). The quantum yield of the 2-Ap residues inserted in $\Delta P(-)$ PBS was largely reduced as compared to the free probe (19) (Table 1), indicating a strong fluorescence quenching by the neighbor bases (33). The quenching was the most pronounced when the 2-Ap residue was inserted in position 10, suggesting a strong stacking with G11. This inferred stacking is in full line with the stacking of C10 with G11 observed by NMR (14), indicating that substitution of C10 by 2-Ap likely preserves the local structure of the loop. Taking into account that each 2-Ap was flanked by a guanine, the most efficient quencher of 2-Ap among the natural bases (34), the quantum yield of the 2-Ap at positions 6 and 8 appeared comparatively higher than those observed in single-stranded ODNs (20,21,35). This suggested that the PBS loop is at least partially ordered, so that efficient collisions with neighbour bases, and notably with the G7 base, are restricted. This is again in line with NMR data (14), since the T8 base was found perpendicular to the G7 residue, while the T6 base could stack with C5 but not with G7. Thus, our data strongly suggest that 2-Ap substitutions of the natural bases do only marginally perturb the native folding of the loop.

The time-resolved intensity decays were complex, showing that conformational fluctuations of the loop occurred during the excited state of the 2-Ap. These decays were best fitted with four discrete lifetime components, ranging from 0.1 ns to nearly 9 ns (Table 1), indicating that 2-Ap experiences at least four main conformational states. Moreover, since the mean lifetime was found to be only 4-6.5 times shorter than the 10.2 ns of the free 2-Ap, far from the ratios (25-155) of their corresponding quantum yields, additional conformations associated to ultra-short-lived lifetimes below the time resolution of our set-up had to be considered (36). These dark-species were previously shown to result from ultra-fast dynamic quenching (21).

Dark-species and weakly emitting species were by far the most populated conformations, since they represented more than 90% of the 2-Ap conformations, as it could be seen from the sum of the α_0 and α_i amplitudes (Table 1). This efficient dynamic quenching of 2-Ap by its neighbours can be related with conformational fluctuations of the loop in the picosecond–nanosecond (ps–ns) range that allows quenching of 2Ap through a charge transfer mechanism (36-40). The stacked-

conformations rose to 98% for 2-Ap10, further confirming a strong stacking with G11. The values of the longer-lived lifetime τ_4 , associated to extra-helical or unstacked conformations, were close to the lifetime of free 2-Ap but low-populated, representing from 1 to 4% of the whole conformations (Table 1). The high value of the long-lived τ_4 lifetime confirmed the limited flexibility of the (-)PBS loop, since in short ss-DNAs, the corresponding lifetime did not exceed 5ns (21,41), due to efficient collisions permitted by the high flexibility of these sequences during the probe lifetime. Moreover, the low amplitude α_4 observed for Ap6, Ap8 and Ap10, suggested that stable conformations with bases rejected toward the solvent were marginal, in full agreement with the orientation of the bases towards the interior of the PBS loop (14,42).

The local dynamics of the 2-Ap residues was further explored by time-resolved fluorescence anisotropy which provides information on their rotational dynamics. Fluorescence anisotropy decays were adequately fitted with a three-exponential model (Table 2 and Figure 2). The two shorter components presumably correspond to the local rotation of the dye and the segmental mobility of the loop, respectively. The slowest correlation time of 2.1 ns to 2.6 ns was attributed to the global tumbling of the DNA sequence. Its deviations from the theoretical correlation time (1.75 ns) calculated for a sphere with the same molecular mass than the ODN (29) were related to the non-spherical shape of PBS stem-loop. The relative amplitude of the fastest component allowed to estimate the angular range of the local motion of 2-Ap modelled as the diffusion in a cone semi-angle (θ_0). These angles were found comparable for the 2-Ap residues at the three positions and significantly lower to those observed in small single-stranded ODNs (20,21,35), confirming that the degree of freedom of the $\Delta P(-)$ PBS loop bases was restricted as compared to ssDNAs. The restricted rotation of the 2Ap bases is in full line with the internal orientation of the corresponding bases in the loop, which provides a relatively crowded environment (14,42).

Effect of NCp7 on the dynamics of the $\Delta P(-)$ PBS loop

To characterize the effect of NCp7 on the dynamics of the $\Delta P(-)$ PBS loop through the 2-Ap fluorescence changes, we first checked whether the 2-Ap substitutions did not alter the binding of NCp7 to the $\Delta P(-)$ PBS derivatives. By monitoring the binding of the 2-Ap-labeled ODNs to NCp7 through the quenching of the intrinsic Trp37 residue (26,27), we found that the 2-Ap substitutions did not induce significant changes in the binding parameters (see supplementary data), in line with the aforementioned marginal changes induced by the 2-Ap substitutions in the $\Delta P(-)$ PBS structure. In response to NCp7 added to the 2Ap-labelled ODNs at a ratio of 3:1 to ensure saturation of the two protein binding sites on the loop (14), the maximum fluorescence emission wavelength of 2-Ap, about 365 nm, was not modified, but its fluorescence quantum yield was increased by a factor of 2 to 5 according to its position in $\Delta P(-)$ PBS (Table 1). As a consequence, the binding of NCp7

on $\Delta P(-)$ PBS loop significantly reduces the level of quenching of 2-Ap fluorescence by its neighbor bases but does not change the polarity of its environment. Furthermore, analysis of the time-resolved intensity decays revealed that the increase in the 2-Ap quantum yield was mainly due to a decrease in the populations of the dark species to the benefit of the most emitting species, as well as to an increase in the τ_d value. These changes in the amplitudes and lifetime values suggested that NCp7 severely restricted the stacking and collisions of 2Ap with its neighbors, in line with the previously reported NCp7-induced stretching of the entire loop and directing of the T6 and G7 bases toward the exterior of the loop (14) that markedly increase the distance between the bases. Interestingly, the large changes in the amplitudes and τ_d values observed for 2-Ap at position 8, are likely a consequence of the stacking of the Trp37 residue with G7, which prevents its collisions and stacking with 2Ap8. Noticeably, the large changes in the amplitudes and τ_d values of 2Ap10 at the 3' end of the loop further indicated that NCp7 binding to its second site in the loop induces similar restrictions on the stacking and collisions of 2Ap.

Furthermore, time-resolved anisotropy decays further revealed that binding of NCp7 to the $\Delta P(-)$ PBS loop induced a strong decrease in the amplitude associated with the local motion of the 2-Ap bases, irrespective of their position in the loop. The restriction of the local motion of the bases at positions 6 and 8, is fully consistent with the numerous contacts observed by NMR between NCp7 and these bases, when NCp7 binds to the 5' end of the loop (14). Interestingly, we observed that binding of NCp7 to the 3' end of the loop induced a similar freezing of the local motion of 2-Ap at position 10, suggesting that NCp7 binding induces similar restrictions on the local mobility of the bases in its two binding sites. The binding of at least two NCp7 molecules on the 2Ap-labeled ODNs was clearly confirmed by the 9-10 ns value of the Φ_3 correlation time, which is in line with the theoretical values of 7 ns and 10 ns expected respectively, for a 2:1 complex and a 3:1 complex with a spherical shape. Interestingly, in all $\Delta P(-)$ PBS derivatives, the amplitude associated with the Φ_3 correlation time strongly increases on NCp7 binding, in line with an overall decrease in the loop flexibility, so that the loop and the stem tumble as a whole. Taken all together, our data indicate that binding of NCp7 molecules to its two binding sites on $\Delta P(-)$ PBS loop strongly restricts the picosecond to nanosecond dynamics of the loop, by constraining both the overall flexibility of the loop and the local mobility of the bases. A major consequence of the restricted dynamics of the loop is that it results in base unstacking.

Effect of NCp7 mutants on the dynamics of the $\Delta P(-)$ PBS loop

To identify the protein determinants responsible for the NCp7-induced changes in the $\Delta P(-)$ PBS loop structure and dynamics, we used a series of NCp7 mutants (Figure 1).

The contribution of the N-terminal domain was investigated with NC(11-55), a peptide composed

of the zinc finger domain but lacking the basic N-terminal domain (43). Binding of NC(11-55) to $\Delta P(-)$ PBS labeled by 2-Ap at positions 6, 8 or 10 induced changes in both steady-state and time-resolved fluorescence parameters similar to those of NCp7 (Tables 1 and 2), indicating that the modifications in the structure and dynamics of the $\Delta P(-)$ PBS loop are largely mediated by the zinc finger domain.

Next, we investigated the role of the Zn^{2+} -induced folding of the NCp7 fingers, by using the $(SSHS)_2NC(11-55)$ mutant where all cysteines are substituted for serines, in order to prevent the binding of zinc and thus, the folding of the fingers (44,45). Binding of this mutated peptide to the 2Ap-labeled ODNs was evidenced by the large increase of the Φ_3 correlation time (Table 2), which suggests that at least two peptides bind to the ODN. However, in sharp contrast to NCp7 and NC(11-55), $(SSHS)_2NC(11-55)$ induced only marginal changes in the steady-state and time-resolved parameters of the 2Ap-labeled $\Delta P(-)$ PBS derivatives, indicating that this mutant only marginally restricts the flexibility of the loop and the local mobility of the bases. Thus, the folding of the zinc fingers appears critical for the NCp7-induced changes in the structure and dynamics of the $\Delta P(-)$ PBS loop.

In a next step, the contribution of the hydrophobic plateau at the top of the two zinc fingers was investigated by mutating punctually the two conserved aromatic residues that belong to this plateau and play a critical role in ODN binding (46-48). Phe₁₆ and Trp₃₇ residues were substituted by Ala and Leu in the A₁₆NC(11-55) and L₃₇NC(11-55) mutants, respectively. Since these two amino-acids are not involved in zinc chelation, these mutations were expected to not alter the folding of the zinc fingers (Mély et al, Biochemistry, 1994). Due to their limited affinity for ODNs (21), A₁₆NC(11-55) and L₃₇NC(11-55) peptides were added in large excess (1 peptide per 2 nt and 1 peptide to 1 nt, respectively), to ensure full coating of $\Delta P(-)$ PBS. The extensive coating of the ODNs by the peptides was evidenced by the large ϕ_3 values, comparable to those obtained with NC(11-55). Both A₁₆NC(11-55) and L₃₇NC(11-55) mutants induced only a limited increase in the quantum yield of the different 2Ap-substituted $\Delta P(-)$ PBS sequences, suggesting that these mutations dramatically alter the ability of NCp7 to reduce base stacking and restrict base collisions. This was confirmed by the limited lifetime redistribution towards the less quenched conformations. In addition, the two mutants induced only a moderate restriction of the local mobility of 2-Ap as compared to NC(11-55), demonstrating the critical role of the hydrophobic platform in this process. Moreover, it should be noted that these two mutants show asymmetric effects in respect to the unstacking of 2-Ap in position 6 and 8. Indeed, A₁₆NC(11-55) was more efficient than W₃₇NC(11-55) to reduce the amplitude α_0 associated to dark species (0.70 vs 0.79) and the local mobility of 2Ap in position 8 ($\beta_I = 0.18$ vs 0.26). In contrast, at position 6, L₃₇NC(11-55) appeared more efficient ($\alpha_0 = 0.62$ vs 0.78 and $\beta_I = 0.21$ vs 0.28). This asymmetric effect was consistent with NMR data showing that in

the NC(12-53)/ Δ P(-)PBS complex, F₁₆ interacts with T6 while W₃₇ interacts with G7.

These observations strongly suggest that the ability of NCp7 to unstack and reorient the bases of the loop and to restrict the local and overall dynamics of Δ P(-)PBS are mainly mediated by the hydrophobic plateau at the top of the folded zinc fingers.

Role of the restriction by NCp7 of the (-)PBS loop dynamics in the (-)/(+)PBS annealing reaction

To further evaluate the relevance of the restriction by NCp7 of the dynamics of (-)PBS loop and notably, its contribution to the promotion of (-)/(+)PBS annealing, we investigated the annealing reactions of doubly-labelled (+)PBS with (-)PBS in the presence of NCp7 and its mutants. The initial fluorescence of the folded 5'Rh6G-(+)PBS-3'DABCYL was very low due to its stem-loop structure, which brings the Rh6G fluorophore in close vicinity to the DABCYL group, acting as a quencher. Addition of NCp7 or NC(11-55) to the doubly-labelled (+)PBS at a ratio of one NCp7 per 5 nt induced a small fluorescence increase, in line with the weak destabilizing activity of NCp7 on the PBS stem (23). In contrast, the (SSHS)₂NC(11-55) or L₃₇NC(11-55) mutants did not induce any increase in the Rh6G fluorescence, confirming the critical role the Trp37 residue and the folded zinc fingers in the NCp7 destabilizing activity (46).

Addition of (-)PBS to 5'Rh6G-(+)PBS-3'DABCYL allowed the formation of the extended (+)/(-)PBS duplex, inducing an important increase of the dye-to-dye distance and thus, Rh6G fluorescence restoration. In the absence of peptide, the (+)/(-)PBS hybridization occurred spontaneously, but at a rather slow rate (Table 3)(8). An about two-order of magnitude increase in the rate of the (+)/(-)PBS annealing reaction was observed in the presence of NCp7. The NC(11-55) and L₃₇NC(11-55) mutants also increased the annealing reaction rate, albeit to a lesser extent than NCp7 (5- to 6-fold increase). Interestingly, a much stronger increase (30-fold) in the annealing rate was obtained with the (SSHS)₂NC(11-55) mutant. Since this mutant was shown to primarily promote annealing reactions through its electrostatic nucleic acid aggregating component (49,50), this component appears to play a major role in the (+)/(-)PBS reaction rate.

To further dissect the effect of the NCp7 mutants on the (+)/(-)PBS annealing reaction, we monitored the annealing reaction in the temperature range 5-50°C and plotted the reaction rates through an Arrhenius plot. According to the Arrhenius model, the transition state thermodynamic parameters can be derived from the reaction rates using:

$$k = A \exp (-E_a/RT) \quad \text{Equation (5)}$$

where A is the preexponential factor, E_a is the activation energy, R the gas constant and T the temperature. In line with the Arrhenius model, the logarithm of the bimolecular rate constants was linearly dependent on the inverse of the temperature (Figure 3). Interestingly, the fitting of the

experimental data in the absence and the presence of the different NC mutants resulted in parallel lines (Figure 3), indicating that the activation energy (E_a) needed for the reaction to be productive was poorly affected by NCp7 and its mutants (Table 4). In sharp contrast to the E_a parameter, the rates of annealing k were strongly dependent on the presence and nature of the peptide, giving significant variation of the Gibbs free energy ΔG^* , according to:

$$\Delta G^* = -RT \ln \left(\frac{k \cdot h}{k_B \cdot T} \right) \quad \text{Equation (6)}$$

where h and k_B are the Planck and Boltzmann constants, respectively. As the ΔH^* value was invariant for all NCp7 derivatives, changes in ΔG^* had to be explained by changes in the activation entropy of the reaction (ΔS^*) (Table 3). The highest ΔS^* values were observed with the native NCp7 and the SSHS NC(11-55) derivative, while a much lower value was observed for the NC(11-55) mutant.

Next, we performed a (-)PBS mutational analysis to determine the reaction pathways in the presence of NCp7 and its mutants. The annealing experiments were first performed with the $A_5(-)$ PBS mutant (Figure 1) which prevents the annealing pathways initiated through the loops. Both NCp7 and NC(11-55) were previously shown to prevent the annealing of (+)PBS to $A_5(-)$ PBS by dissociating the duplexes that spontaneously form in the absence of peptide (8). In sharp contrast, both $L_{37}NC(11-55)$ or $(SSHHS)_2NC(11-55)$ strongly promoted the annealing of (+)PBS to $A_5(-)$ PBS, with annealing rates close to those observed with the native (+)/(-)PBS sequences, suggesting that the (-)PBS loop does not play a critical role in the reaction pathway promoted by these mutants.

We next investigated the annealing reaction of (+)PBS with a $T_4(-)$ PBS mutant, in which the 15-GCCA-18 sequence was substituted by a T_4 sequence to prevent the nucleation through the PBS stem. In the presence of NCp7 or NC(11-55), the rate constants with the $T_4(-)$ PBS mutant were similar to those with the native (-)PBS, suggesting that the (-)PBS single strand overhang is not critical for the main reaction pathway promoted by NCp7 and NC(11-55). In contrast, a 15-fold and a 4-fold decrease in the annealing rates was observed when the corresponding annealing reactions were compared in the presence of $(SSHHS)_2NC(11-55)$ or $L_{37}NC(11-55)$, respectively (Table 3). This clearly indicated that the (-)PBS overhang plays a central role in the $(SSHHS)_2NC(11-55)$ -promoted (+)/(-)PBS annealing reaction, and to a lesser extent in the $L_{37}NC(11-55)$ -promoted (+)/(-)PBS annealing reaction.

Thus, our data strongly suggest that $(SSHHS)_2NC(11-55)$ or $L_{37}NC(11-55)$ promote the (+)/(-)PBS annealing reaction, mainly through the single-strand overhangs at the bottom of the PBS stems, while, NCp7 and NC(11-55) promote this reaction mainly through the loops. As a consequence, the activity of NCp7 mutants on the loop dynamics correlates well with their ability to promote the (+)/(-)PBS annealing reaction through the loops. Therefore, NCp7-induced restriction of the local

and overall dynamics of the PBS loop likely constitutes a molecular requisite for promoting the (+)PBS/(-)PBS annealing reaction through the loop-loop pathway.

Discussion

In this study, we investigated the mechanism of the NCp7-promoted (+)/(-)PBS annealing reaction that occurs during the second strand transfer. To this end, we used $\Delta P(-)$ PBS and $\Delta P(+)$ PBS stem-loops substituted with 2-aminopurine (2-Ap) at different positions in the loop to characterize the structure and dynamic changes in the loop promoted by NCp7 and various NCp7 mutants. Then, these observed changes were correlated with the thermodynamic and kinetic parameters of the annealing reactions obtained in the presence of the same proteins.

The substitution of natural bases by 2-Ap in the PBS loops was shown to marginally affect the folding of the loop and its binding parameters with NCp7. In the absence of protein, the 2Ap residues in both (-)PBS and (+)PBS loops were found to experience multiple conformations, but in a large majority (> 90%) of them, the 2-Ap could be efficiently quenched through collisions with its neighbors. These data were fully consistent with an orientation of the bases toward the interior of the loop (42). Moreover, the value of the long-lived lifetime observed irrespective of the position of 2-Ap in the loop confirmed the partial order within the loop observed by NMR (14,42).

NCp7 induced position-dependent changes of the 2Ap fluorescence, indicating that NCp7 restricts the picosecond to nanosecond dynamics of the PBS loops, by constraining both the overall flexibility of the loops and the local mobility of the bases. Similar dynamic changes were observed with small flexible single-stranded ODNs (20,21), suggesting that they correspond to a general feature of the NCp7/ODN interaction. Using various NCp7 mutants, the dynamic changes in the PBS loops were found to be mainly promoted by the folded finger motifs, and more specifically by the hydrophobic plateau at the top of the fingers that supports the destabilizing component of the chaperone activity of NCp7. Most interestingly, the NCp7-induced changes in the dynamics in the PBS loops were found to strongly correlate with the ability of NCp7 to switch the (-)/(+) PBS annealing reaction through a loop-loop kissing pathway that poorly contributes to the annealing reaction in the absence of NCp7. Single point mutations of either of the critical F₁₆ and W₃₇ aromatic aminoacids in the two zinc fingers resulted in very limited mobility restriction, spatially confined to the binding site of the non mutated finger on the loop. Since these NCp7 mutants marginally induce the annealing pathway switch and since in spite of its short size, the loop can accommodate two NCp7 proteins, one can speculate that the dynamics of the whole PBS loop should be restricted to obtain the mechanism switch. This conclusion is in line with the observation that optimal chaperoning of nucleic acid sequences requires their saturation by NCp7 (3,5,28,51).

From the temperature dependence of the annealing rate, it was established that NCp7 did not lower the activation energy E_a of the annealing reaction. In fact, the corresponding positive enthalpy ($\Delta H^* \sim 77$ kJ/mol) (Table 4) was consistent with the melting of about 3-4 base-pairs (49,52), indicating that the melting of the PBS stem likely constitutes the rate-limiting step of the annealing

reaction (8). Since the ΔH^* value was not affected by NCp7 or its mutants, the transition state of the reaction likely remains the same whatever the annealing pathway and thus constitutes a saddle point in the energy landscape. Furthermore, the absence of NCp7-induced changes in the ΔH^* value was fully consistent with the previously reported inability of NCp7 to melt the stable (-)PBS stem (22,23).

Comparison of the annealing rates of NCp7 and of the NC(11-55) mutant strongly suggested that the aggregating properties of the N-terminal domain of NCp7 (43,50,53) play a key role in increasing the annealing reaction rates. This “aggregation” effect, due to the additional screening of the negatively charged ODNs by the basic N_{terminal} domain, is thought to strongly increase the ΔS^* value and the rate at which collisions occur between the peptide coated ODNs, facilitating the diffusional search for the complementary sequences (51). The ΔS^* value and the annealing rate were also high in the presence of (SSHS)₂NC(11-55). As a consequence of its unfolded structure, the (SSHS)₂NC(11-55) can better fit to the ODN structure than the folded NC(11-55) peptide and thus, more efficiently screen the repulsion forces between the phosphate charges within the ODN, creating short range attractions between ODNs, even at sub-aggregating concentrations (5). A similar improved performance of a SSHS mutant as compared to the native NCp7 was also observed in the annealing reaction of tRNA^{Lys3} to the PBS (49). Since in addition to be less efficient than (SSHS)₂NC(11-55), NC(11-55) increased the rate of the (+)/(-)PBS annealing reaction to the same extent than the L₃₇NC(11-55) derivative, the annealing rates appeared disconnected from the reaction pathway and the ability of the mutants to restrict the mobility of the PBS loop. Moreover, since the unfolding of the zinc fingers also abolished the nucleic acid destabilizing activity of NCp7 (46,50,54), the annealing reaction rates appear also poorly related to the destabilizing activity of the protein.

Furthermore, our data provide insights into the understanding of the annealing mechanism switch at the molecular level. The observed changes in the loop dynamics and the resulting base unstacking clearly suggested that NCp7 stabilizes conformations where the loop bases of both (+) and (-)PBS species are “stretched”, in line with an induced-fit model. This model was also previously proposed to describe the effect of NCp7 on the AACGCC hexanucleotide (21). The NMR structure of the ACGCC/NC(12-53) and $\Delta P(-)PBS/NC(11-55)$ complexes show that the hydrophobic plateau at the top of the two zinc fingers is the major determinant for the binding of NCp7 to the two ODNs (14,55). However, the structures of these two complexes markedly differ in the sense that NCp7 binds at the surface of the 5' end of the (-)PBS loop, and increases the exposure of the loop bases, while NCp7 wraps the ACGCC sequence and decreases the base exposure (Figure 4C). Interestingly, the ACGCC sequence is strongly similar to the highly flexible CGCCA 3' overhang of the (-)PBS sequence, described as a putative binding site for NCp7 within PBS (23,42) (Figure 4).

One can thus speculate that the binding of NCp7 to the 3' overhang of the (-)PBS sequence results, as for ACGCC, in a dramatic restriction of the base exposure with a NCp7 molecule that wraps the DNA strand. This situation is opposite to that in the absence of NCp7, where the bases of the loop are pointing to the interior of the loop (14), while the bases of the unfolded 3' overhang are largely exposed. As a consequence, the differences in the base exposure between the loop and the 3' overhang provide a rational explanation for the preferential annealing through the PBS overhangs in the absence of protein, and through the PBS loops in the presence of NCp7. Being unable to change the structure and dynamics of the PBS loops as well as those of unfolded oligonucleotides (21), the (SSHS)₂NC(11-55) and L₃₇NC(11-55) mutants will thus promote the annealing through the more exposed PBS overhangs. It is also worth to notice that the binding of NCp7 is highly dynamic in nature (5,56), with a lifetime of the NC(11-55)/PBS complex of about 350ms (14). One has thus to picture a process where NCp7 rapidly and transiently capture the competent conformations for annealing through the loop before dissociating (57). As a consequence, NCp7 insures the specificity of the annealing pathway, through its hydrophobic platform. Our data on NCp7 and PBS mutants also suggest a strong evolutionary relationship between the NCp7 structure and the specific annealing pathway. In this context, the speed of reaction becomes a secondary outcome since rates and extend of the reactions do not directly translate the chaperone capacities of the NCp7 protein. Indeed, a strict comparison of the rates of the different mutants used in this study would indeed lead to misleading conclusions if the annealing pathway is not considered.

Our data clearly evidenced that the pathways of the nucleic acids rearrangements that must occur during reverse transcription to generate the proviral DNA are multiple. This may explain contradictory observations between *in vitro* observations where the folded zinc fingers of NCp7 appeared not mandatory in the promotion of annealing reactions (58-60), while even subtle mutations within the zinc fingers lead to non infectious particle *in vivo* (61-63). Indeed, mutations of the NCp7 zinc fingers were reported to modify the reverse transcription timing, so that reverse transcription can occur before the virion release, giving DNA-containing viral particles (64-67). This "late reverse transcription" likely results from the facilitation of the reverse transcription by the aspecific polycations or proteins in the physiological environment. The loss of the NCp7 destabilizing activity induced by mutations in the zinc fingers (46,54) seems thus to result in an increased efficiency of reverse transcription which is detrimental to virus replication. The question of how the NCp7 nucleic acid chaperoning can prevent premature reverse transcription finds some answers in our observations. Since zinc finger mutants do not modify the ODN dynamics and structure, they will thus poorly modify the energy landscape of the ODN structures and they will stimulate their rearrangements through already available existing pathways. In contrast, the destabilizing activity of NCp7 reshapes the energy landscape and funnels the ODN rearrangements

towards specific routes that ensure appropriate timing of the reverse transcription and formation of appropriate genomic DNA product. The inhibition of the annealing of imperfect complementary sequences (see mutant A₅(-)PBS in Table 3) and the change in the annealing mechanism pathways constitutes two examples of how NCp7 directs the ODN to specific routes.

It clearly appeared in this study that the restriction of the base mobility constitute a key feature of the chaperone activity of NCp7. It could be interesting to determine whether our observations can be extended to ODNs such as cTAR that are destabilized by NCp7 (28,68-72) and for which a NCp7-induced mechanistic switch in the annealing pathway was also shown (51). This work is under progress.

References

1. Basu, V.P., Song, M., Gao, L., Rigby, S.T., Hanson, M.N. and Bambara, R.A. (2008) Strand transfer events during HIV-1 reverse transcription. *Virus Res*, **134**, 19-38.
2. Darlix, J.L., Lapadat-Tapolsky, M., de Rocquigny, H. and Roques, B.P. (1995) First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J Mol Biol*, **254**, 523-537.
3. Cristofari, G. and Darlix, J.L. (2002) The ubiquitous nature of RNA chaperone proteins. *Prog Nucleic Acid Res Mol Biol*, **72**, 223-268.
4. Darlix, J.L., Cristofari, G., Rau, M., Péchoux, C., Berthoux, L. and Roques, B. (2000) Nucleocapsid protein of human immunodeficiency virus as a model protein with chaperoning functions and as a target for antiviral drugs. *Adv Pharmacol*, **48**, 345-372.
5. Levin, J.G., Guo, J., Rouzina, I. and Musier-Forsyth, K. (2005) Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. *Prog Nucleic Acid Res Mol Biol*, **80**, 217-286.
6. Rein, A., Henderson, L.E. and Levin, J.G. (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication. *Trends Biochem Sci*, **23**, 297-301.
7. Mély, Y., Rocquigny, H.D., Morellet, N., Roques, B.P. and Gérard, D. (1996) Zinc binding to the HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence spectroscopy. *Biochemistry*, **35**, 5175-5182.
8. Ramalanjaona, N., de Rocquigny, H., Millet, A., Ficheux, D., Darlix, J.-L. and Mély, Y. (2007) Investigating the mechanism of the nucleocapsid protein chaperoning of the second strand transfer during HIV-1 DNA synthesis. *J Mol Biol*, **374**, 1041-1053.
9. Wu, T., Guo, J., Bess, J., Henderson, L.E. and Levin, J.G. (1999) Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer: analysis in reconstituted and endogenous reverse transcription systems. *J Virol*, **73**, 4794-4805.
10. Chan, B. and Musier-Forsyth, K. (1997) The nucleocapsid protein specifically anneals tRNA^{Lys-3} onto a noncomplementary primer binding site within the HIV-1 RNA genome in vitro. *Proc Natl Acad Sci U S A*, **94**, 13530-13535.
11. Li, X., Quan, Y., Arts, E.J., Li, Z., Preston, B.D., de Rocquigny, H., Roques, B.P., Darlix, J.L., Kleiman, L., Parniak, M.A. et al. (1996) Human immunodeficiency virus Type 1 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis primed by human tRNA(Lys3) in vitro: studies of viral RNA molecules mutated in regions that flank the primer binding site. *J Virol*, **70**, 4996-5004.
12. Rong, L., Liang, C., Hsu, M., Kleiman, L., Petitjean, P., de Rocquigny, H., Roques, B.P. and Wainberg, M.A. (1998) Roles of the human immunodeficiency virus type 1 nucleocapsid protein in annealing and initiation versus elongation in reverse transcription of viral negative-strand strong-stop DNA. *J Virol*, **72**, 9353-9358.
13. Herschlag, D. (1995) RNA chaperones and the RNA folding problem. *J Biol Chem*, **270**, 20871-20874.
14. Bourbigot, S., Ramalanjaona, N., Boudier, C., Salgado, G.F.J., Roques, B.P., Mély, Y., Bouaziz, S. and Morellet, N. (2008) How the HIV-1 nucleocapsid protein binds and destabilises the (-)primer binding site during reverse transcription. *J Mol Biol*, **383**, 1112-1128.
15. Lee, B.M., Guzman, R.N.D., Turner, B.G., Tjandra, N. and Summers, M.F. (1998) Dynamical behavior of the HIV-1 nucleocapsid protein. *J Mol Biol*, **279**, 633-649.
16. Mély, Y., Jullian, N., Morellet, N., Rocquigny, H.D., Dong, C.Z., Piémont, E., Roques, B.P. and Gérard, D. (1994) Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer. *Biochemistry*, **33**, 12085-12091.
17. Morellet, N., de Rocquigny, H., Mély, Y., Jullian, N., Déméné, H., Ottmann, M., Gérard, D., Darlix, J.L., Fournie-Zaluski, M.C. and Roques, B.P. (1994) Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. *J Mol Biol*, **235**, 287-301.
18. Mori, M., Dietrich, U., Manetti, F. and Botta, M. (2010) Molecular Dynamics and DFT Study on HIV-1 Nucleocapsid Protein-7 in Complex with Viral Genome. *J Chem Inf Model*.
19. Ward, D.C., Reich, E. and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J Biol Chem*, **244**, 1228-1237.
20. Avilov, S., Godet, J., Piémont, E. and Ly, Y.M. (2009) Site-specific characterization of HIV-1 nucleocapsid protein binding to oligonucleotides with two binding sites. *Biochemistry*.
21. Avilov, S.V., Piémont, E., Shvadchak, V., de Rocquigny, H. and Mély, Y. (2008) Probing dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-aminopurine. *Nucleic Acids Res*, **36**, 885-896.
22. Egelé, C., Schaub, E., Piémont, E., de Rocquigny, H. and Mély, Y. (2005) Investigation by fluorescence correlation spectroscopy of the chaperoning interactions of HIV-1 nucleocapsid protein with the viral DNA initiation sequences. *C R Biol*, **328**, 1041-1051.
23. Egelé, C., Schaub, E., Ramalanjaona, N., Piémont, E., Ficheux, D., Roques, B., Darlix, J.-L. and Mély, Y. (2004) HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences and chaperones their kissing interactions. *J Mol Biol*, **342**, 453-466.
24. de Rocquigny, H., Ficheux, D., Gabus, C., Fournie-Zaluski, M.C., Darlix, J.L. and Roques, B.P. (1991) First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem Biophys Res Commun*, **180**, 1010-1018.
25. Shvadchak, V.V., Klymchenko, A.S., de Rocquigny, H. and Mély, Y. (2009) Sensing peptide-oligonucleotide interactions by a two-color fluorescence label: application to the HIV-1 nucleocapsid protein. *Nucleic Acids Res*.
26. Mély, Y., de Rocquigny, H., Piémont, E., Déméné, H., Jullian, N., Fournie-Zaluski, M.C., Roques, B. and Gérard, D. (1993) Influence of the N- and C-terminal chains on the zinc-binding and conformational properties of the central zinc-finger structure of Moloney murine leukaemia virus nucleocapsid protein: a steady-state and time-resolved fluorescence study.

- Biochim Biophys Acta*, **1161**, 6–18.
27. Vuilleumier, C., Bombarda, E., Morellet, N., Gérard, D., Roques, B.P. and Mély, Y. (1999) Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study. *Biochemistry*, **38**, 16816–16825.
 28. Bernacchi, S., Stoylov, S., Piémont, E., Ficheux, D., Roques, B.P., Darlix, J.L. and Mély, Y. (2002) HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J Mol Biol*, **317**, 385–399.
 29. Lakowicz, J.R. (1999) *Principles of Fluorescence Spectroscopy*. 2nd ed. Kluwer Academic/Plenum Publishers.
 30. Lipari, G. and Szabo, A. (1980) Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes. *Biophys J*, **30**, 489–506.
 31. Brochon, J.C. (1994) Maximum entropy method of data analysis in time-resolved spectroscopy. *Methods Enzymol*, **240**, 262–311.
 32. Livesey, A.K. and Brochon, J.C. (1987) Analyzing the Distribution of Decay Constants in Pulse-Fluorimetry Using the Maximum Entropy Method. *Biophys J*, **52**, 693–706.
 33. Wan, C., Xia, T., Becker, H.-C. and Zewail, A.H. (2005) Ultrafast unequilibrated charge transfer: A new channel in the quenching of fluorescent biological probes. *Chemical Physics Letters*, **412**, 158–163.
 34. Guest, C.R., Hochstrasser, R.A., Sowers, L.C. and Millar, D.P. (1991) Dynamics of mismatched base pairs in DNA. *Biochemistry*, **30**, 3271–3279.
 35. Jean, J.M. and Hall, K.B. (2002) 2-Aminopurine electronic structure and fluorescence properties in DNA. *Biochemistry*, **41**, 13152–13161.
 36. Larsen, O.F., van Stokkum, I.H., Gobets, B., van Grondelle, R. and van Amerongen, H. (2001) Probing the structure and dynamics of a DNA hairpin by ultrafast quenching and fluorescence depolarization. *Biophys J*, **81**, 1115–1126.
 37. Kelley, R.E., Namon, R.A., Mantelle, L.L. and Chang, J.Y. (1993) Sensitivity and specificity of transcranial Doppler ultrasonography in the detection of high-grade carotid stenosis. *Neurology*, **43**, 1187–1191.
 38. Wan, C., Fiebig, T., Schiemann, O., Barton, J.K. and Zewail, A.H. (2000) Femtosecond direct observation of charge transfer between bases in DNA. *Proc Natl Acad Sci U S A*, **97**, 14052–14055.
 39. Fiebig, T., Wan, C. and Zewail, A.H. (2002) Femtosecond charge transfer dynamics of a modified DNA base: 2-aminopurine in complexes with nucleotides. *Chemphyschem*, **3**, 781–788.
 40. O'Neill, M.A. and Barton, J.K. (2002) 2-Aminopurine: a probe of structural dynamics and charge transfer in DNA and DNA:RNA hybrids. *J Am Chem Soc*, **124**, 13053–13066.
 41. Jean, J.M. and Hall, K.B. (2004) Stacking-unstacking dynamics of oligodeoxynucleotide trimers. *Biochemistry*, **43**, 10277–10284.
 42. Johnson, P.E., Turner, R.B., Wu, Z.R., Hairston, L., Guo, J., Levin, J.G. and Summers, M.F. (2000) A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry*, **39**, 9084–9091.
 43. Stoylov, S.P., Vuilleumier, C., Stoylova, E., De Rocquigny, H., Roques, B.P., Gérard, D. and Mély, Y. (1997) Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. *Biopolymers*, **41**, 301–312.
 44. Green, L.M. and Berg, J.M. (1990) Retroviral nucleocapsid protein-metal ion interactions: folding and sequence variants. *Proc Natl Acad Sci U S A*, **87**, 6403–6407.
 45. Guo, J., Wu, T., Anderson, J., Kane, B.F., Johnson, D.G., Gorelick, R.J., Henderson, L.E. and Levin, J.G. (2000) Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J Virol*, **74**, 8980–8988.
 46. Beltz, H., Clauss, C., Piémont, E., Ficheux, D., Gorelick, R.J., Roques, B., Gabus, C., Darlix, J.-L., de Rocquigny, H. and Mély, Y. (2005) Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J Mol Biol*, **348**, 1113–1126.
 47. Fisher, R.J., Fivash, M.J., Stephen, A.G., Hagan, N.A., Shenoy, S.R., Medaglia, M.V., Smith, L.R., Worthy, K.M., Simpson, J.T., Shoemaker, R. *et al.* (2006) Complex interactions of HIV-1 nucleocapsid protein with oligonucleotides. *Nucleic Acids Res*, **34**, 472–484.
 48. Gorelick, R.J., Gagliardi, T.D., Bosche, W.J., Wiltout, T.A., Coren, L.V., Chabot, D.J., Lifson, J.D., Henderson, L.E. and Arthur, L.O. (1999) Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology*, **256**, 92–104.
 49. Hargittai, M.R.S., Gorelick, R.J., Rouzina, I. and Musier-Forsyth, K. (2004) Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J Mol Biol*, **337**, 951–968.
 50. Williams, M.C., Rouzina, I., Wenner, J.R., Gorelick, R.J., Musier-Forsyth, K. and Bloomfield, V.A. (2001) Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc Natl Acad Sci U S A*, **98**, 6121–6126.
 51. Vo, M.-N., Barany, G., Rouzina, I. and Musier-Forsyth, K. (2009) HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper". *J Mol Biol*, **386**, 789–801.
 52. Rouzina, I. and Bloomfield, V.A. (1999) Heat capacity effects on the melting of DNA. I. General aspects. *Biophys J*, **77**, 3242–3251.
 53. Le Cam, E., Coulaud, D., Delain, E., Petitjean, P., Roques, B.P., Gérard, D., Stoylova, E., Vuilleumier, C., Stoylov, S.P. and Mély, Y. (1998) Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron microscopy investigation. *Biopolymers*, **45**, 217–229.
 54. Williams, M.C., Gorelick, R.J. and Musier-Forsyth, K. (2002) Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc Natl Acad Sci U S A*, **99**, 8614–8619.
 55. Morellet, N., Déméné, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournié-Zaluski, M.C. and Roques, B.P. (1998) Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J Mol Biol*, **283**, 419–434.
 56. Cruceanu, M., Gorelick, R.J., Musier-Forsyth, K., Rouzina, I. and Williams, M.C. (2006) Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. *J Mol Biol*, **363**,

- 867--877.
57. Leulliot, N. and Varani, G. (2001) Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry*, **40**, 7947-7956.
 58. De Rocquigny, H., Gabus, C., Vincent, A., Fournie-Zaluski, M.C., Roques, B. and Darlix, J.L. (1992) Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc Natl Acad Sci U S A*, **89**, 6472-6476.
 59. Lapadat-Tapolsky, M., Gabus, C., Rau, M. and Darlix, J.L. (1997) Possible roles of HIV-1 nucleocapsid protein in the specificity of proviral DNA synthesis and in its variability. *J Mol Biol*, **268**, 250-260.
 60. Prats, A.C., Housset, V., de Billy, G., Cornille, F., Prats, H., Roques, B. and Darlix, J.L. (1991) Viral RNA annealing activities of the nucleocapsid protein of Moloney murine leukemia virus are zinc independent. *Nucleic Acids Res*, **19**, 3533-3541.
 61. Dorfman, T., Luban, J., Goff, S.P., Haseltine, W.A. and Gottlinger, H.G. (1993) Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol*, **67**, 6159-6169.
 62. Gorelick, R.J., Chabot, D.J., Rein, A., Henderson, L.E. and Arthur, L.O. (1993) The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J Virol*, **67**, 4027-4036.
 63. Gorelick, R.J., Gagliardi, T.D., Bosche, W.J., Wiltout, T.A., Coren, L.V., Chabot, D.J., Lifson, J.D., Henderson, L.E. and Arthur, L.O. (1999) Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes: characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology*, **256**, 92-104.
 64. Didierlaurent, L., Houzet, L., Morichaud, Z., Darlix, J.-L. and Mougel, M. (2008) The conserved N-terminal basic residues and zinc-finger motifs of HIV-1 nucleocapsid restrict the viral cDNA synthesis during virus formation and maturation. *Nucleic Acids Res*, **36**, 4745-4753.
 65. Houzet, L., Morichaud, Z., Didierlaurent, L., Muriaux, D., Darlix, J.-L. and Mougel, M. (2008) Nucleocapsid mutations turn HIV-1 into a DNA-containing virus. *Nucleic Acids Res*, **36**, 2311-2319.
 66. Mougel, M., Houzet, L. and Darlix, J.-L. (2009) When is it time for reverse transcription to start and go? *Retrovirology*, **6**, 24.
 67. Thomas, J.A., Bosche, W.J., Shatzer, T.L., Johnson, D.G. and Gorelick, R.J. (2008) Mutations in human immunodeficiency virus type 1 nucleocapsid protein zinc fingers cause premature reverse transcription. *J Virol*, **82**, 9318-9328.
 68. Azoulay, J., Clamme, J.P., Darlix, J.L., Roques, B.P. and Mély, Y. (2003) Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J Mol Biol*, **326**, 691-700.
 69. Beltz, H., Azoulay, J., Bernacchi, S., Clamme, J.-P., Ficheux, D., Roques, B., Darlix, J.-L. and Mély, Y. (2003) Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol*, **328**, 95-108.
 70. Beltz, H., Piémont, E., Schaub, E., Ficheux, D., Roques, B., Darlix, J.-L. and Mély, Y. (2004) Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol*, **338**, 711-723.
 71. Cosa, G., Harbron, E.J., Zeng, Y., Liu, H.-W., O'Connor, D.B., Eta-Hosokawa, C., Musier-Forsyth, K. and Barbara, P.F. (2004) Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys J*, **87**, 2759-2767.
 72. Cosa, G., Zeng, Y., Liu, H.-W., Landes, C.F., Makarov, D.E., Musier-Forsyth, K. and Barbara, P.F. (2006) Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned opening of DNA hairpins. *J Phys Chem B*, **110**, 2419-2426.

Table 1. Steady-state and time-resolved fluorescence parameters of 2Ap-substituted $\Delta P(-)$ PBS

| | Quantum yield | α_0 | τ_1 | α_1 | τ_2 | α_2 | τ_3 | α_3 | τ_4 | α_4 | $\langle \tau \rangle$ |
|---|--------------------|------------|----------|------------|----------|------------|----------|------------|----------|------------|------------------------|
| Free 2Ap riboside | 0.680 ^a | - | - | - | - | - | - | - | 10.2 | 1.00 | 10.2 |
| $\Delta P(-)$PBS 6AP | 0.024 | 0.81 | 0.10 | 0.13 | 0.66 | 0.02 | 3.60 | 0.02 | 8.00 | 0.03 | 1.81 |
| $\Delta P(-)$PBS 6AP + NCp7 | 0.044 | 0.73 | 0.12 | 0.10 | 0.60 | 0.07 | 2.80 | 0.06 | 8.30 | 0.05 | 2.43 |
| $\Delta P(-)$PBS 6AP + NC(11-55) | 0.040 | 0.74 | 0.12 | 0.10 | 0.60 | 0.06 | 3.40 | 0.04 | 7.50 | 0.05 | 2.33 |
| $\Delta P(-)$PBS 6AP + A₁₆NC(11-55) | 0.026 | 0.78 | 0.11 | 0.12 | 0.70 | 0.04 | 3.10 | 0.04 | 8.30 | 0.03 | 1.77 |
| $\Delta P(-)$PBS 6AP + L₃₇NC(11-55) | 0.035 | 0.62 | 0.10 | 0.20 | 0.70 | 0.08 | 3.30 | 0.06 | 8.75 | 0.05 | 1.76 |
| $\Delta P(-)$PBS 6AP + SSHS₂NC(11-55) | 0.024 | 0.81 | 0.09 | 0.08 | 0.60 | 0.04 | 2.90 | 0.06 | 7.50 | 0.02 | 1.95 |
| $\Delta P(-)$PBS 8AP | 0.028 | 0.83 | 0.09 | 0.08 | 0.70 | 0.03 | 3.10 | 0.03 | 8.30 | 0.04 | 2.10 |
| $\Delta P(-)$PBS 8AP + NCp7 | 0.087 | 0.60 | 0.09 | 0.13 | 0.60 | 0.08 | 3.80 | 0.08 | 8.90 | 0.11 | 3.28 |
| $\Delta P(-)$PBS 8AP + NC(11-55) | 0.081 | 0.60 | 0.14 | 0.13 | 0.60 | 0.10 | 3.80 | 0.07 | 8.50 | 0.10 | 3.04 |
| $\Delta P(-)$PBS 8AP + A₁₆NC(11-55) | 0.033 | 0.70 | 0.07 | 0.16 | 0.70 | 0.05 | 3.20 | 0.04 | 8.20 | 0.04 | 2.00 |
| $\Delta P(-)$PBS 8AP + L₃₇NC(11-55) | 0.035 | 0.79 | 0.10 | 0.09 | 0.90 | 0.04 | 3.60 | 0.04 | 8.90 | 0.04 | 2.51 |
| $\Delta P(-)$PBS 8AP + SSHS₂NC(11-55) | 0.032 | 0.80 | 0.10 | 0.09 | 0.70 | 0.04 | 3.20 | 0.04 | 8.20 | 0.03 | 2.10 |
| $\Delta P(-)$PBS 10AP | 0.006 | 0.95 | 0.13 | 0.03 | 0.80 | 0.01 | 3.50 | 0.01 | 8.20 | 0.01 | 2.10 |
| $\Delta P(-)$PBS 10AP + NCp7 | 0.028 | 0.83 | 0.13 | 0.08 | 0.70 | 0.03 | 4.30 | 0.04 | 8.80 | 0.02 | 2.43 |
| $\Delta P(-)$PBS 10AP + NC(11-55) | 0.021 | 0.85 | 0.15 | 0.07 | 0.70 | 0.03 | 3.90 | 0.02 | 8.00 | 0.02 | 2.13 |
| $\Delta P(-)$PBS 10AP + A₁₆NC(11-55) | 0.008 | 0.95 | 0.07 | 0.02 | 0.60 | 0.01 | 2.85 | 0.01 | 7.50 | 0.01 | 2.10 |
| $\Delta P(-)$PBS 10AP + L₃₇NC(11-55) | 0.009 | 0.95 | 0.09 | 0.02 | 0.80 | 0.01 | 3.10 | 0.01 | 8.20 | 0.01 | 2.64 |
| $\Delta P(-)$PBS 10AP + SSHS₂NC(11-55) | 0.007 | 0.92 | 0.10 | 0.05 | 0.80 | 0.02 | 3.90 | 0.02 | 8.30 | 0.01 | 2.00 |
| $\Delta P(+)$PBS 11AP | 0.039 | 0.52 | 0.10 | 0.32 | 0.80 | 0.09 | 4.00 | 0.03 | 8.40 | 0.04 | 1.33 |
| $\Delta P(+)$PBS 11AP + NCp7 | 0.064 | 0.51 | 0.10 | 0.22 | 0.85 | 0.12 | 3.40 | 0.06 | 8.10 | 0.07 | 1.91 |
| $\Delta P(+)$PBS 11AP + SSHS₂NC(11-55) | 0.039 | 0.51 | 0.10 | 0.30 | 0.90 | 0.10 | 3.50 | 0.05 | 8.40 | 0.03 | 1.20 |

^a data from (19) τ_i (ns) are the fluorescence lifetimes, α their amplitudes. The amplitude values are corrected for the dark species, as described in the Materials and Methods section. $\langle \tau \rangle$ is the mean fluorescence lifetime.

Table 2. Fluorescence anisotropy decay parameters of 2Ap-substituted $\Delta P(-)$ PBS

| | Φ_1 | β_1 | Φ_2 | β_2 | Φ_3 | β_3 | S | Θ_0 |
|---|----------|-----------|----------|-----------|----------|-----------|------|------------|
| Free 2-Ap | 0.08 | 1.00 | - | - | - | - | 0.00 | 90 |
| $\Delta P(-)$PBS 6AP | 0.17 | 0.38 | 0.74 | 0.34 | 2.20 | 0.28 | 0.79 | 32 |
| $\Delta P(-)$PBS 6AP + NCp7 | 0.20 | 0.16 | 1.48 | 0.22 | 9.70 | 0.63 | 0.88 | 19 |
| $\Delta P(-)$PBS 6AP + NC(11-55) | 0.17 | 0.18 | 1.15 | 0.28 | 7.40 | 0.54 | 0.91 | 21 |
| $\Delta P(-)$PBS 6AP + A¹⁶NC(11-55) | 0.08 | 0.28 | 1.22 | 0.30 | 9.40 | 0.41 | 0.85 | 27 |
| $\Delta P(-)$PBS 6AP + L³⁷NC(11-55) | 0.14 | 0.21 | 1.22 | 0.33 | 9.00 | 0.46 | 0.89 | 23 |
| $\Delta P(-)$PBS 6AP + SSHS₂NC(11-55) | 0.18 | 0.32 | 1.20 | 0.21 | 9.50 | 0.47 | 0.82 | 29 |
| $\Delta P(-)$PBS 8AP | 0.22 | 0.36 | 0.68 | 0.11 | 2.14 | 0.53 | 0.79 | 31 |
| $\Delta P(-)$PBS 8AP + NCp7 | 0.13 | 0.13 | 1.28 | 0.25 | 10.20 | 0.62 | 0.91 | 17 |
| $\Delta P(-)$PBS 8AP + NC(11-55) | 0.13 | 0.13 | 0.89 | 0.32 | 7.68 | 0.55 | 0.93 | 17 |
| $\Delta P(-)$PBS 8AP + A¹⁶NC(11-55) | 0.10 | 0.18 | 0.60 | 0.24 | 7.80 | 0.58 | 0.87 | 21 |
| $\Delta P(-)$PBS 8AP + L³⁷NC(11-55) | 0.20 | 0.26 | 1.20 | 0.24 | 9.10 | 0.50 | 0.83 | 25 |
| $\Delta P(-)$PBS 8AP + SSHS₂NC(11-55) | 0.18 | 0.30 | 1.10 | 0.27 | 9.30 | 0.43 | 0.84 | 28 |
| $\Delta P(-)$PBS 10AP | 0.18 | 0.47 | 0.85 | 0.27 | 2.60 | 0.27 | 0.73 | 36 |
| $\Delta P(-)$PBS 10AP + NCp7 | 0.22 | 0.16 | 0.76 | 0.24 | 9.20 | 0.60 | 0.91 | 19 |
| $\Delta P(-)$PBS 10AP + NC(11-55) | 0.20 | 0.21 | 0.90 | 0.20 | 7.50 | 0.59 | 0.89 | 23 |
| $\Delta P(-)$PBS 10AP + A¹⁶NC(11-55) | 0.10 | 0.36 | 0.78 | 0.24 | 7.30 | 0.41 | 0.80 | 31 |
| $\Delta P(-)$PBS 10AP + L³⁷NC(11-55) | 0.15 | 0.32 | 1.16 | 0.33 | 9.10 | 0.35 | 0.82 | 29 |
| $\Delta P(-)$PBS 10AP + SSHS₂NC(11-55) | 0.16 | 0.38 | 0.90 | 0.20 | 9.10 | 0.42 | 0.79 | 32 |

Φ_i (ns) are the rotational correlation times, β_i their amplitudes, S is the generalized order parameter and Θ_0 the cone semi-angle for 2Ap local motion (calculated as described in Materials and Methods section)

Table 3. Kinetic parameters of the annealing of (+)PBS to (-)PBS mutants in the absence and in the presence of NC derivatives

| Labeled sequence | Complementary sequence | k ($M^{-1} s^{-1}$) | | | | |
|------------------|------------------------|-----------------------------|----------------------------------|--------------------------------|---------------------------|------------------------------|
| | | - | NCp7 | NC(11-55) | L ₃₇ NC(11-55) | (SSH) ₂ NC(11-55) |
| (+)PBS | (-)PBS | 3,800 ^a ± 100 | 180,000 ^a ± 10,000 | 18,400 ^a ± 1,200 | 22,600 ± 2,000 | 99,000 ± 6,500 |
| | A5(-)PBS | 3,200 ^a ± 100 | - ^b | - ^b | 18,000 ± 3,000 | 64,000 ± 4,000 |
| | T ₄ (-) PBS | ~20 ^a | 210,000 ± 20,000 | 11,100 ^a ± 200 | 6,000 ± 1,000 | 7,000 ± 500 |

a. Data from Ramalanjaona & al. (8)

b. No rate constant could be defined since NCp7 promotes the dissociation of the (-)/(+)PBS duplex

Table 4. Thermodynamic Parameters of the NC-promoted (+)/(-)PBS Annealing at 25°C

| | k^a ($M^{-1} s^{-1}$) | Ea^b (kJ/mol) | ΔG^{*c} (kJ/mol) | ΔH^{*d} (kJ/mol) | ΔS^{*e} (eu) |
|-----------------------------|------------------------------|--------------------|-----------------------------|-----------------------------|-------------------------|
| No peptide | $4.3 \pm 1.0 \times 10^3$ | 80 ± 8 | 57 | 78 ± 5 | 17 ± 3 |
| NCp7 | $3.2 \pm 1.0 \times 10^5$ | 78 ± 10 | 45 | 76 ± 8 | 26 ± 3 |
| NC(11-55) | $3.1 \pm 0.6 \times 10^4$ | 78 ± 12 | 52 | 76 ± 9 | 19 ± 2 |
| A ₁₆ NC(11-55) | $9.4 \pm 0.7 \times 10^3$ | 80 ± 8 | 55 | 78 ± 6 | 17 ± 2 |
| L ₃₇ NC(11-55) | $3.2 \pm 0.8 \times 10^4$ | 77 ± 11 | 48 | 74 ± 8 | 18 ± 2 |
| SSHS ₂ NC(11-55) | $1.8 \pm 0.6 \times 10^5$ | 81 ± 9 | 48 | 79 ± 7 | 25 ± 3 |

- a. Determined at 25°C
- b. Determined from the Arrhenius plot (Figure 3)
- c. Calculated from the annealing rate constant at 25°C according to equation (7)
- d. ΔH^* is given by $\Delta H^* = Ea - RT$ for T=25°C.
- e. Calculated according to $\Delta G^* = \Delta H^* - T\Delta S^*$. eu corresponds to ($cal mol^{-1} K^{-1}$)

Figures legends:

Figure 1: Sequences of the oligonucleotides and peptides used in this study.

Figure 2: Experimental anisotropy decay curve of the $\Delta P(-)$ PBS AP6 in the absence (black) and in the presence of NCp7 (green). The continuous fitted lines (red and blue) correspond to the fit of the data with the parameters in table 2.

Figure 3: Arrhenius analysis of the annealing of $(-)$ PBS to $(+)$ PBS in the presence of NCp7 derivatives. The experiments were performed in the absence (squares) or in the presence of NCp7 (open squares) and its different mutants, NC(11-55) (circles), S₂NC(11-55) (triangles), A₁₆NC(11-55) (open circles), and W₃₇NC(11-55) (open triangles).

Figure 4:

- A. Superimposition of 20 refined NMR structures of the $(-)$ PBS (PDB 1EN1) (42). The bulged nucleotide T2, the bases of the loop and the 3'end termini single-stranded segment appeared clearly disordered.
- B. Superimposition of 19 and 10 refined NMR structures of the $\Delta P(-)$ PBS (upper, PDB 2EXF)) and ACGCC pentanucleotide (lower, PDB 1BJ6), respectively, in the presence of NCp7 (14,55).
- C. Three dimensional structure of the $\Delta P(-)$ PBS/(11-55)NCp7 (upper) and ACGCC/(12-53)NCp7 (lower) complexes. The ODN and NCp7 atoms are colored in grey and green respectively. (11-55)NCp7 binds on one side of the $\Delta P(-)$ PBS loop, leaving the 8-TTG-10 bases not involved in the interaction with NC easily accessible for annealing the complementary sequence. The binding of (12-53)NCp7 to the ACGCC pentanucleotide results in a deep wrapping of the protein around the nucleic acid sequence.

Figure 1:

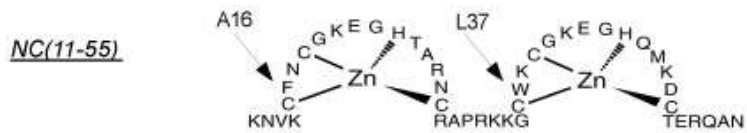
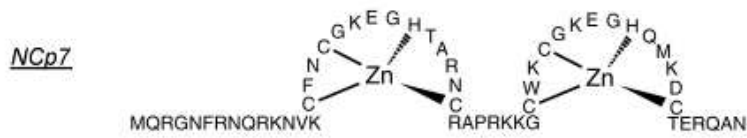
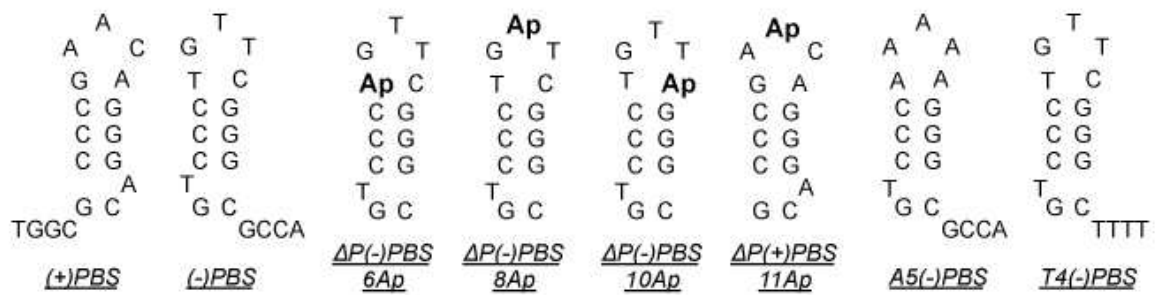


Figure 2:

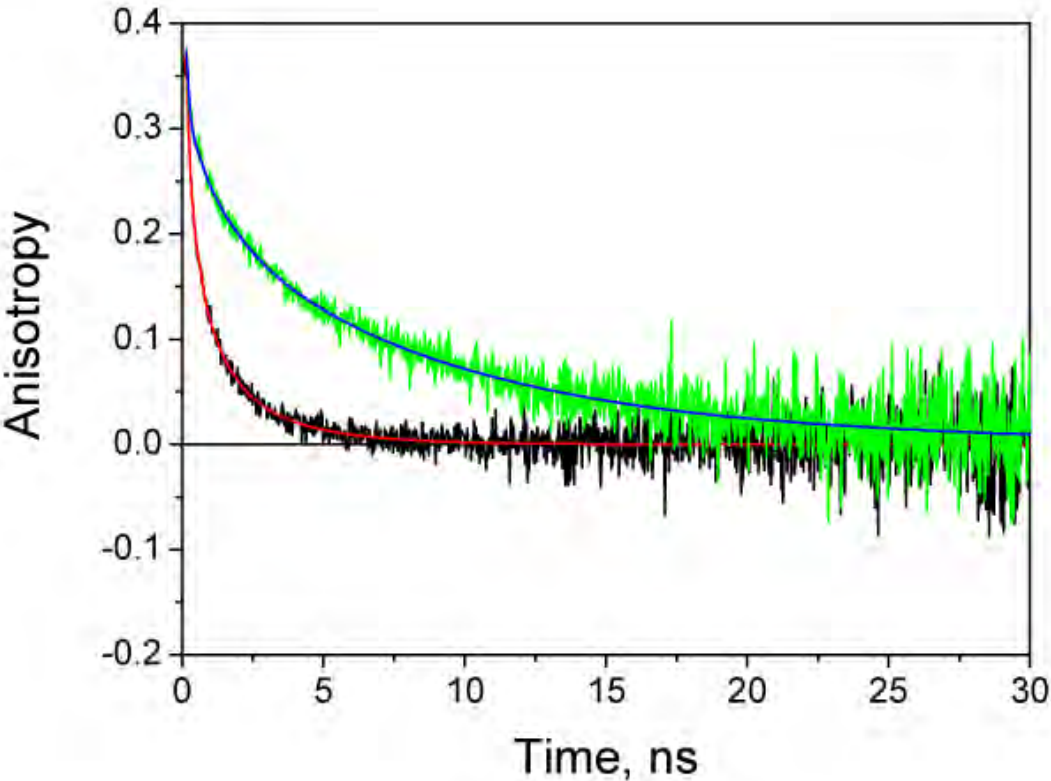


Figure 3:

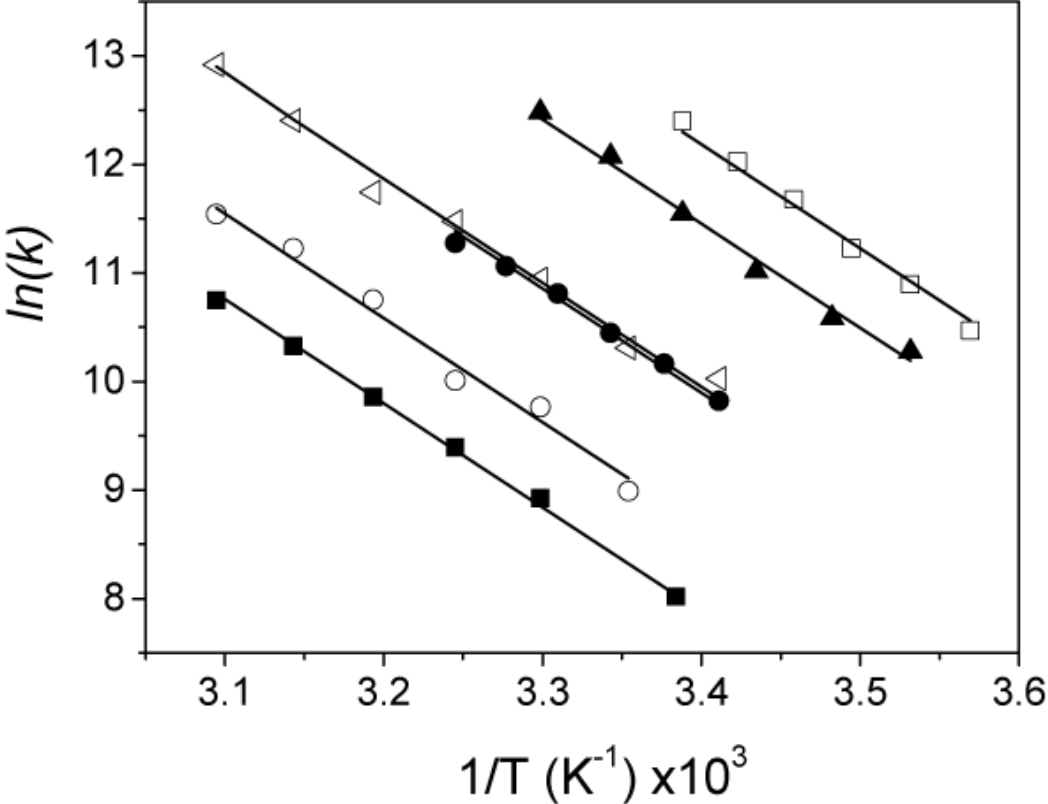
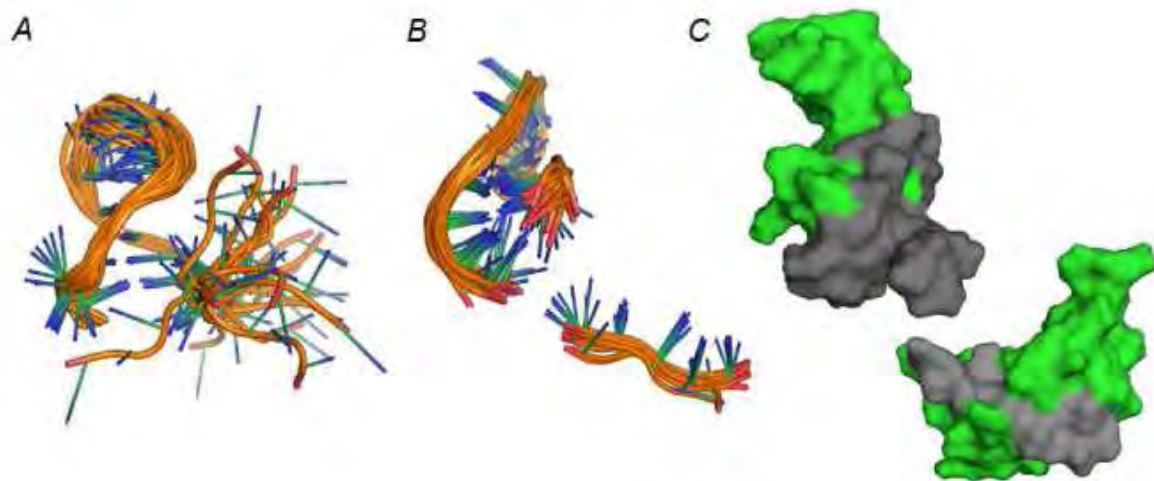


Figure 4:



Supplementary materials

Binding affinities of NC(11-55) for ΔP (-)PBS sequences substituted at positions 6, 8 or 10.

| Sequence | K_{app} | n | I_f/I_0 |
|------------------------|-------------------------|---------------|-----------------|
| ΔP (-)PBS | $8 \pm 2 \times 10^6$ | 1.9 ± 0.2 | 0.14 ± 0.02 |
| ΔP (-)PBS 6AP | $7 \pm 2 \times 10^6$ | 2.2 ± 0.2 | 0.13 ± 0.01 |
| ΔP (-)PBS 8AP | $11 \pm 3 \times 10^6$ | 2.3 ± 0.1 | 0.14 ± 0.02 |
| ΔP (-)PBS 10AP | $8.5 \pm 2 \times 10^6$ | 2.3 ± 0.3 | 0.14 ± 0.02 |

Binding constants were determined by Trp quenching experiments. Peptide concentration was 1 μ M in 25 mM Tris-HCl, pH 7.5, 30 mM NaCl, and 0.2 mM MgCl₂. Excitation and emission wavelengths were 295 and 350 nm, respectively.

5.5 Caractérisation de la déstabilisation locale de cTAR par la NCp7

Les propriétés de déstabilisation des acides nucléiques par la NCp7 ont été très largement étudiées (pour une revue voir article I section 1.9). La déstabilisation de la séquence cTAR a été particulièrement étudiée. Des expériences de spectroscopie d'absorption ont montré que la NCp7 est capable de fondre environ 7 à 8 paires de bases au sein de cTAR. La caractérisation de la fluorescence d'une séquence de cTAR doublement marquée à ses extrémités avec un donneur et un accepteur de fluorescence a permis de montrer que cette déstabilisation est principalement localisée dans la partie basse de la tige de cTAR [393]. La NCp7 agit en activant les fluctuations conformationnelles du bas de la tige de cTAR [267, 400]. De plus, la capacité de déstabilisation de la NCp7 est largement dépendante de la présence de *bulges* (et notamment du G_{52}) au sein de la séquence [268]. En ligne avec sa faible activité de déstabilisation, la stabilité intrinsèque des acides nucléiques gouverne largement la déstabilisation induite par NCp7. Ainsi, la partie supérieure de cTAR, plus stable que la partie basse, est beaucoup moins destabilisée [269], tout comme la séquence PBS, également plus stable [386, 348]. Cette capacité de déstabilisation dépend de l'intégrité du plateau hydrophobe qui se forme à la surface des doigts de zinc correctement repliés [226]. Les mécanismes de déstabilisation apparaissent complexes, impliquant des fluctuations de la structure secondaire de cTAR sur diverses échelles de temps [401].

Malgré toutes ces caractérisations, les travaux décrivant précisément l'activité de déstabilisation de la NCp7 au niveau des paires de bases sont peu nombreux et se résument pour l'essentiel à des approches par RMN [227, 260, 264, 265]. La possibilité offerte par la 2-Ap de sonder localement les dynamiques des bases devrait permettre d'étudier de manière site-spécifique la liaison de la NCp7 ainsi que son activité de déstabilisation au sein de la séquence cTAR. Dans cette optique, nous avons marqué la séquence cTAR par l'insertion sélective de 2-Ap en différentes positions. La caractérisation de la fluorescence des 2-AP au sein de cTAR montre que l'environnement de la 2-Ap est hétérogène et dynamique. La présence de NCp7 induit de manière site-spécifique des restrictions de la mobilité de la sonde et des mouvements des segments d'acides nucléiques où la sonde est insérée. De manière

5.5 Caractérisation de la déstabilisation locale de cTAR par la NCp7

intéressante, deux sites de liaisons préférentiels ont été observés. Ceux-ci sont positionnés “face à face” sur chacun des brins au coeur de la tige de cTAR et flanqués de part et d’autre par deux paires de bases non-appariées. Nos observations suggèrent que la déstabilisation de cTAR résulte initialement d’une capture conformationnelle par la NCp7 des espèces ouvertes, puis la déstabilisation de cTAR est ensuite étendue, possiblement de manière coopérative, par la liaison de NCp7 supplémentaires résultant en un domaine de déstabilisation évoluant sur de multiples échelles de temps.

Ces travaux sont présentés sous forme d’un article en cours d’élaboration.

5.6

Publication VII

Site-selective probing of the binding of the nucleocapsid protein of HIV-1 with 2-aminopurine-labelled cTAR.

Julien Godet^{1,2}, Cyril Kenfack^{1,3}, Ludovic Richert¹, Etienne Piemont¹, Guy Duportail¹, Yves Mély^{1*}

¹ Biophotonic and Pharmacology Laboratory, Faculty of Pharmacy, UMR 7213 CNRS, University of Strasbourg, 67401 Illkirch, France

² Medical Information and Biostatistics Department, Strasbourg University Hospital, 67000 Strasbourg, France

³ Centre de Physique Atomique Moléculaire et Optique Quantique, Faculté des Sciences, Université de Douala, BP 8580, Douala, Cameroon

* Corresponding author

Tel : +33 (0)3 90 24 42 63 Fax : +33 (0)3 90 24 43 12

E-mail: yves.mely@unistra.fr

Running title: Binding of the nucleocapsid protein NCp7 on cTAR

Abstract:

Double-stranded DNA is a dynamic structure that undergoes conformational changes in response to protein binding. Fluorescence spectroscopy of 2'-deoxyribosyl-2-aminopurine (2-AP), a fluorescent nucleosid analogue highly sensitive to its environment, offers an elegant opportunity to probe specifically the DNA structure and dynamic changes induced by DNA-protein interactions. The 57 DNA nucleotide cTAR, the DNA complementary sequence of TAR in HIV-1, has been selectively modified with 2-AP in different positions within its stem-loop structure. During the reverse transcription, cTAR is strongly destabilized by the nucleocapsid protein of HIV-1, a chaperone protein that strongly stimulates the TAR/cTAR annealing during the first strand transfer. The cTAR destabilizing properties of (11-55)NCp7, a truncated version of that lacks aggregating properties, has been evaluated through the time-resolved fluorescence intensity and anisotropy decays of 2-AP. In the absence of protein, time-resolved measurements of 2-AP:cTAR revealed an heterogeneous environment with a distribution of partially stacked structures. Perturbation in the base stacking induced by the binding of NCp7 led to a fluorescence increase due to the shift towards less-stacked conformations and to a decrease of the local mobility of the DNA segment where (11-55)NCp7 is bound. Using the site-specific 2Ap probe, we propose for the first time a statistical map of the preferential binding sites for NCp7 on the cTAR sequence.

Abbreviations: 2AP, 2'-deoxyribosyl-2-aminopurine; HIV-1, human immunodeficiency virus type 1; NC, nucleocapsid protein; RT, reverse transcriptase; TAR, transactivation response element; cTAR, complementary DNA sequence to TAR; dTAR, DNA analogue to TAR; SL, stem loop; LTR, long terminal repeat; FRET, fluorescence resonance energy transfer; ssDNA, strong-stop DNA; PBS, primer binding site

Key words: time-resolved fluorescence, 2-aminopurine, HIV NCp7, TAR, binding

Introduction

The nucleocapsid protein (NCp7) of the Human immunodeficiency virus type 1 (HIV-1) is a small basic protein generated by the viral protease-mediated cleavage of the Gag polyprotein precursor. NCp7 is formed of two highly-conserved CCHC zinc finger motifs connected by a short flexible linker and flanked by poorly-folded N- and C-terminal basic domains [1, 2]. NCp7 is critically involved during the early and late steps of HIV-1 viral life cycle [3, 4]. Most NCp7 functions rely on its nucleic chaperone chaperoning properties that direct the rearrangement of nucleic acid molecules into their most thermodynamically stable structure [5, 6]. These chaperone properties are thought to play notably a critical role in reverse transcription [7-11]. For instance, during the first strand-transfer of the reverse transcription, NCp7 likely favours the transfer of the newly synthesized minus strong-stop DNA ((-) ssDNA) to the 3' end of the viral RNA genome, by promoting the base pairing of the complementary repeat (R) sequences at the 3'ends of the (-)ssDNA and genomic RNA, respectively. In vitro, the stable stem-loop (SL) of the Transactivation response element (TAR) sequence within the R domain of the genomic RNA does not readily anneal with the complementary cTAR sequence of the (-)ssDNA in the absence of NCp7 [9, 11-13]. In contrast, addition of NCp7 strongly increases the rate and extent of the cTAR/TAR annealing reaction [7, 9, 11-14]. Two distinct components of NCp7 chaperone properties are involved in the promotion of the first strand transfer. The first one corresponds to the duplex destabilizing activity of NCp7 that activates the transient opening (fraying) of the lower half of the cTAR stem [15-20]. NCp7 also melts TAR but less efficiently than cTAR. This NCp7 activity is mediated by the hydrophobic platform formed by the Val₁₃, Phe₁₆, Thr₂₄, Ala₂₅, Trp₃₇, Gln₄₅ and Met₄₆ residues at the surface of the folded finger domain [1, 2, 21]. The second component of the NCp7 chaperone activities consists in the activation of the annealing of the complementary sequences to generate the thermodynamically favoured cTAR/TAR duplex [12, 13, 20, 22, 23]. This activity likely relies on the numerous basic residues of NCp7 that can decrease the electrostatic repulsion between the oligonucleotides by neutralizing the negatively charged phosphate groups [4, 13, 24]. These NCp7 chaperone properties require a complex and highly dynamic binding mode to ensure the fraying of the stem loop ends (destabilization effect) [25] and to promote the productive encounter between the complementary sequences (hybridization effect) [4]. NCp7 binding to oligonucleotides likely results from both specific [26, 27] and nonspecific interactions [7, 28, 29]. Though NCp7 can bind to nearly any sequence of five to seven nucleotides length [27, 28, 30, 31], NCp7 binds preferentially with TG- or UG-rich sequences [27, 31], as well as with GXG- containing sequences [31-33]. The hydrophobic platform on the top of the two folded zinc fingers motifs has been shown to be critical for binding [21, 34]. More specifically, the Trp₃₇ residue plays a critical role, through

stacking with G residues [27, 31, 34-36]. Kinetic differences also appear between the binding sites. While preferential binding sites gave stable complexes on the NMR time scale [34-36], the binding of NCp7 with tRNA^{Lys,3} was found to be in fast exchange on the NMR time timescale, suggesting that NCp7 protein could “slide” from one position to another, in a submillisecond time scale [37]. More recently, 2'-deoxyribosyl-2-aminopurine (2Ap), a fluorescent analogue of adenine, has been used to site-specifically characterize the NCp7-induced dynamic changes on 2Ap-labelled single-stranded sequences [38, 39]. Time-resolved fluorescence studies of 2Ap showed that NCp7 restricted the ODN flexibility and the base local mobility on a picoseconds-nanosecond time scale. The distortion of the ODN structure and the local “freezing” of the bases were assumed to constitute the initial events for both the nucleic acid destabilization and annealing components of NCp7 chaperone activity.

To further relate the dependence of the chaperone properties of NCp7 on its binding and promoted changes in the local dynamics of selected regions of target ODNs, we site-selectively characterized in the present study the interaction of NCp7 with 2Ap-labeled cTAR sequences (Fig. 1). Contrary to usual techniques which are able to monitor global structural changes, the exquisite sensibility of 2-AP fluorescence to its local environment allows regio-specific characterization of the structural modifications within the cTAR hairpin sequence. The cTAR sequence of the Lai strain was preferred to that of the MAL strain, due to its higher content in A residues that enables to conservatively substitute these positions with 2Ap. Since cTAR is an imperfect stem-loop characterized by a 3 nt internal loop as well as by several bulged nucleotides and mismatches, strategic positions within the sequences were selected for labelling: position 9 forms a mismatch with C within the lower part of the stem, positions 28 and 35 are within the central and the internal loop, respectively, positions 53 and 55 are flanking the G54 bulge and positions 17, 21, 45 and 49 are distributed all over the stem within double stranded segments. Using time-resolved fluorescence techniques, NCp7 was found to destabilize preferentially the lower internal segment of the cTAR stem and to locally decrease the base mobility in this region, in line with the critical role of NCp7 to promote the cTAR/TAR annealing through a zipping mechanism.

Results

Site-specific characterization of the 2Ap-labeled cTAR sequences in the absence of NCp7

We first characterized the fluorescence properties of the 2Ap-labeled cTAR sequences in the absence of NCp7. The fluorescence spectra of the different 2AP-labelled cTAR exhibited a maximum emission wavelength at around 370nm (Fig 2), similarly to the free 2Ap. In contrast, the quantum yield (QY) of 2Ap in cTAR was strongly reduced in comparison to the free probe. Large diminution of the QY is always observed when 2Ap is incorporated in ODNs and results from both static and dynamic quenching of 2Ap by its close neighbors [40-44]. Though quantum yields were low in all labelled cTAR sequences, clear differences appeared as a function of the labelled position and the 2AP-substituted cTAR sequences could be classified in three groups according to their quantum yield. The lowest quantum yields (< 0.01) were observed for positions AP17, AP21 and AP49. These three positions are located within the double-stranded regions in line with the fact that the regular and stable structure of the double strands is known to favour the stacking of 2Ap with its close neighbours, and thus quenching of its fluorescence [41]. The second cluster, pooling together positions AP9, AP45, AP53 and AP55, corresponds to quantum yield in the range of 0.01 to 0.02. All these locations but one corresponded to imperfect DNA helix, either because the 2AP was mismatchly base-paired with a cytosine (AP9) or because the 2AP was flanking the G52-bulged residue, in both case resulting in a decrease of the base stacking. Positions AP45 appeared to be ill-assorted since there were no obvious reasons for the G10-G14 and C43-C47 residues not to form a regular helix. In fact, this increase of the quantum yield in position AP45 compared to positions AP17, AP21 or AP49 is likely due to the absence of guanine in the close vicinity of the 2AP. Because guanine was described to be the best quencher by charge transfer among bases [40, 45], the quantum yield of 2AP in position 45 was significantly higher compared to the others 2Ap flanked by one guanine. It is interesting to notice that the fluorescence of the 2AP in position AP9 is relatively low in respect to the others positions in this group. This lowering of the quantum yield is likely due to the higher ability of 2Ap residues as compared to As to base pair with the opposite C [46-50], so that the double stranded segment containing the 2Ap-C mismatch will be only weakly destabilized and the 2Ap will stack efficiently with its neighbours. The third cluster was constituted by 2Ap located in single-stranded segments. A rather low quantum yield was however associated with 2Ap at position 28 in the loop, suggesting that 2Ap is efficiently quenched by the G29 flanking residue. This is in line NMR data showing that the loop of cTAR is highly flexible [51] and could thus, through fast conformational fluctuations drive 2Ap from an unstacked to a stacked conformation. By far, the highest quantum yield (≈ 0.08) was observed with the 35 position in the internal loop. This can be partly attributed to the presence of only A neighbours that are not efficient quenchers of 2AP fluorescence, but also and probably mainly due to the rather constrained structure

of the internal loop that prevent the collisions of the central 2Ap base with its neighbours.

In order to further characterize the local environment of the probe in the different 2AP-substituted cTARs, temperature-induced melting transitions were performed. Thermal denaturation curves of cTAR stem-loop allowed comparing the 2AP fluorescence emission of the folded cTAR, in respect to its corresponding single-stranded sequence. The effects of the local conformations and dynamics of the DNA, rather than the nature of the flanking bases which strongly governs the photophysics of the 2AP, are therefore investigated. Two types of transitions were observed. When the 2AP was located in ds-segments (AP17, AP21, AP45, AP49), the fluorescence emission of the 2AP strongly increased with the duplex melting (Figure 3.A). This can be ascribed to the disruption of stable stacking between bases in the duplex while melting. AP9 presented a similar transition pattern, confirming that 2AP in this position was base-paired with the opposite C48 (as well as the actual A9, if we consider that AP49 could probe the C48 conformations). On the contrary, down-transitions were observed for AP53 and AP55 (Figure 3.B). Both AP53 and AP55 presented a melting transition from weaker-stacked conformations in the double-strand to stronger ones in the single-strand. The existence of such two-state transition clearly indicated that the last ds-segments of the cTAR stem bottom were formed at 20°C. The transition of AP55 was however reached out over a broader scale of temperature. It signed a weaker stability of the very last ds-segment of cTAR and suggested a significant probability for transient melting at 20°C. Moreover, the low fluorescence quenching of the 2AP in the folded AP53 and AP55 probably resulted from the extrahelical position of the G54 [52]. AP28 also underwent a down melting transition. This was quite surprising since the 2AP was incorporated in the single-stranded loop of cTAR, and thus, as for AP35 (Figure 3.C), no melting transition was expected for AP28. This transition obviously signed the disappearance of the structural constraints within the apical loop. 2AP in position AP28 was thus probing the melting of cTAR stem, providing evidence that, under some conditions, 2AP can probe DNA structure over a longer range than just its directly flanking bases.

To give further insight into the heterogeneity of the probe environment and its associated dynamics, we performed time-resolved experiments of the 2AP-substituted cTAR sequences. In contrast, to the fluorescence decay of free 2-AP which is monoexponential with a lifetime of 10.2ns [44], the fluorescence decay of the 2Ap-labelled cTAR sequences were multiexponential and a satisfactory fitting procedure was obtained with 4 lifetime components, ranging from about 100ps to nearly 10ns, demonstrating a large conformational heterogeneity of the 2Ap residues in the cTAR derivatives. Multiple lifetimes are frequently observed in 2Ap labelled ODNs and are usually associated with various degrees of 2Ap stacking with its neighbours [40-42, 50, 53]. The long-lived component τ_4 is between 6.1 to 9.6 ns, and likely corresponds to conformations where 2Ap is either unstacked or extrahelical [40, 54]. Interestingly, the lowest α_4 values were associated to 2Ap

positions within the double stranded segments, indicating that 2Ap residues have more limited freedom than in the loops and cannot adopt fully extrahelical conformations. The relatively longer lifetime τ_4 of the position 21 suggested that the G20:C33 basepair was more likely to dissociate. This was in line with the destabilizing effect of the internal loop on the top half sequence of cTAR [55]. The remaining lifetime values are about 0.1 ns, 0.5-1 ns and 2.6-3.6 ns, respectively. These ps-ns time regime likely correspond to conformations in which nanosecond conformational fluctuations can lead to 2AP collisions with its neighbours, leading to a dynamic charge transfer and a lifetime shortening [41, 53, 56]. Furthermore, comparison of the mean lifetime values with the quantum yields suggested the presence of additional short-lived lifetimes beyond the time-resolution of our setup (<30ps). Indeed, the mean lifetime of 2Ap in the different positions in cTAR (0.91-1.33 ns) was found to be 5- to 15-fold shorter than that of the free dye (Table 1). These differences in lifetimes were far less than the 10 to 80 ratios in the corresponding quantum yields (Table 1: R_m values), suggesting the presence of a significant subpopulation of “dark” species (α_0) (i.e. species with a lifetime shorter than the detection limit of our equipment (≈ 30 ps)). Such dark species, with 10 ps lifetimes were already observed in DNA hairpins [57] as well as in small hexa-oligonucleotides [39] and were associated to stacked conformations with ultra-fast dynamic quenching through hole transfer from 2Ap to nearby bases.

The strongly-stacked 2-AP conformations, associated with the dark species and τ_1 lifetime were largely dominant for all the 2Ap positions within cTAR (except positions 35, see below), representing at least 90%, as can be seen from the sum of α_0 and α_1 . Strongly-stacked conformations were also largely dominant (90%) for 2Ap at position 28 in the loop. Similar high percentages were observed for hexanucleotides, when 2Ap was next to a G residue [39]. Thus, the cTAR loop likely exhibits high conformational flexibility that allows efficient quenching of 2Ap28 by the G29 residue. In sharp contrast, a more limited population (54%) of strongly stacked conformations was observed with 2Ap at position 35 in the internal loop. For this position, rather large populations of conformations were associated with the τ_2 and τ_3 lifetimes, probably due to constraints in the internal loop that prevent efficient quenching by the neighbour bases. Interestingly, the values of the α_4 amplitudes associated with the long-lived τ_4 lifetimes were very small ($\leq 1\%$) in the double stranded segments, indicating that the partly extrahelical conformations are prevented by the stable structure of the double stranded segments. Significantly higher α_4 values (4%) were seen for the single stranded loops. It indicated that extrahelical 2Ap conformations are more energetically favoured in this case.

Rotational dynamics of 2-AP monitored through time-resolved fluorescence anisotropy provides information on the local structural aspects of the cTAR sequence. The anisotropy decays were satisfactorily fitted with a three exponential model (only two components were sufficient for AP55,

see below). The initial anisotropy r_0 varies from 0.31 to 0.35, close to the 0.33 value of the initial anisotropy for 2AP in 77% glycerol [39]. The short correlation time (Φ_1), 0.1 to 0.45 ns, was assigned to the local motion of the 2-AP with respect to the strand backbone (see the experimental section for the validation of the measurement of the short correlation time). The contribution of any individual rotational diffusion motion to the whole fluorescence depolarisation process is proportional to its corresponding amplitude. This amplitude for the short rotational correlation time in AP9, and AP49 where 2AP is located in the stem region (42 and 41% respectively) was comparable to that of AP28 (47%) in the apical loop or of AP53 (44%), and was much more higher than the other positions where 2AP was in a regular ds-DNA segment of cTAR. This showed that the inter-strand interaction in the stem at this peculiar position was weaker and that the photophysics of 2AP is governed by the proximal G. The lowest amplitude of the short rotational correlation is obtained for AP35, considering the higher value of the long lifetime (τ_4) amplitude in the fluorescence decay obtained for this location, this striking result confirmed that 2AP behaviour in position 35 is peculiar. The cone semi-angle model used as an estimate of the excursion angle (θ_0) of the geometries explored by the 2-AP within the DNA sequence summarized these observations (Table 2), exhibiting for the 2AP a θ value of 22° , confirming the high constraints in this position. Both the high values of θ_0 of AP53 and AP55 (35° and 49° , respectively) showed that 2AP in this two positions was much less constrained than in a regular DNA helix, suggesting that the G54 is likely to be extra-helical.

The long rotational correlation time (Φ_3), usually associated to the global tumbling of the molecule, was scattered, ranging from about 4 to 9 ns. Φ_3 was $\sim 8-9$ ns when 2AP was in the upper half-part of the hairpin loop (i.e AP17, AP21, AP28 and AP35). These correlation times were probably reflecting the whole tumbling of the sequence. Indeed, these values are close to the theoretical correlation times calculated for the tumbling motion of a cylinder rod of length 95 Å and diameter 20 Å [58]. The expected tumbling correlation times for such a structure, where the transition dipole is not collinear with the long axis of the DNA molecule, are 7.70 ns (75% amplitude) and 34 ns (25 % amplitude) [58], a value substantially higher than the $\sim 4-5$ ns measured in the bottom of the cTAR sequence (AP45, AP49, AP53, AP55). By taking into account that only the site located in the lower part of stem undergoes long correlation rotational time shortenings, the Φ_3 component described the overall tumbling but also the “fraying” motion of the conformations, in line with the previously described fluctuations associated with the opening and closing of the double-stranded terminal segments of cTAR [15]. Besides the two usually reported rotational correlation times above mentioned, we obtained for all the sites (but AP55) a rotational correlation time Φ_2 (~ 1 ns) with an amplitude almost constant ($\sim 30\%$), possibly corresponding to a segmental motion of the whole cTAR sequence. The disappearance of this third exponential component in AP55 was likely

due to the shortening of the intermediate correlation time Φ_2 , making it impossible to resolve from the short-one. This shortening probably resulted from an increase of the segmental motions of the end-segment of cTAR at the nanosecond timescale.

Fluorescence emission of 2AP substituted cTAR in presence of NCp7

In order to investigate the effect of NCp7 on 2AP-modified cTAR fluorescence, we first checked out that cTAR modifications with 2AP did not alter NCp7 binding. Due to the low quantum yield of the 2-AP once inserted in DNA, high concentrations of labelled ODNs were required (typically 1-5 μM). This impeded us to use the (1-55)NCp7, which aggregates at these concentrations. The effect of NCp7 binding on the 2AP fluorescence was thus investigated using (11-55)NCp7, a mutant that lacks the major determinants of aggregation of the wild type sequence but which is sufficient for binding and destabilisation. No significant differences were induced by the substitution of an adenine by a 2-AP in the cTAR sequence, indicating that 2-AP does not affect NCp7 binding (see supplementary).

Effect of NCp7

Addition of (11-55)NCp7 to 2Ap substituted cTAR solutions led generally to an increase of the fluorescence intensity with a position of the emission maximum not affected, suggesting that NCp7 binding does not change the polarity of the surrounding of the probe but only affect the quenching by the neighbouring bases (Figure 4.A). This fluorescence changes were relying both on the amount of NCp7 added and to the position of the 2AP within the sequence (Table1). We explored these position-dependant changes with incremental addition of one, four, seven, ten and fifteen equivalents of protein on the 2AP-substituted cTAR modified in different positions. The increase of fluorescence observed in steady-state fluorescence, was linked to an increase of the average lifetime, resulting from redistribution towards the long-lived lifetimes (Table 1 and Figure 4.B). Lifetime values were not affected and remained almost constant for any protein ratio, confirming that NCp7 binding only redistribute the existing conformations of cTAR by altering the quenching within the bases. The redistribution in the conformational states of the 2Ap resulted in an increase of the amplitudes of the longer-lived lifetimes which occurred at the expense of the amplitudes associated with the shortest and the dark species (Figure 4.B). Similar observations were previously done for 2Ap in short ss-hexa- or dodecanucleotides [38, 39]. Anisotropy decay parameters were also obviously modified by the binding of (11-55)NCp7 (Figure 4.C). In the presence of (11-55)NCp7, the decay was still composed of three correlation times. The long correlation time Φ_3 was increased progressively as the amount of (11-55)NCp7 increased to reach values in between 26 and

36ns, confirming that about ten (11-55)NCp7 was bound to the oligonucleotide. The progressive increase of (11-55)NCp7 concentrations did not strongly affect Φ_1 and Φ_2 which remained almost constant (Figure 4.D).

Interestingly, the fluorescence changes in response to NCp7 binding were strongly modulated by the position where the 2-AP was inserted. When (11-55)NCp7 was fully coating the cTAR sequence, the quantum yield of the different 2AP was either unchanged (AP28) or increased by a factor of about two to nearly 14 compared to than in the absence of protein (Table 1 and Figure 5), showing that very different conformational or dynamical changes occurred in response to (11-55)NCp7 binding and suggesting different binding modes for (11-55)NCp7. The sum of the α_0 and α_1 components, which represents the stacked conformations of 2Ap, was differently affected by (11-55)NCp7 (Figure 5). When the 2AP was located in the ds-segments of cTAR, (11-55)NCp7 induced a dose-dependent decrease of the stacked conformations. (11-55)NCp7 did not decrease the stacking of the 2-AP located in the loop, suggesting that the base-stacking in the loop was governed by the structural constrains of the DNA and was not modified by (11-55)NCp7. The decrease of the base stacking appeared homogeneous in cTAR AP17, AP21, AP45, AP53 and AP55. Since these positions are distributed all over the sequence, it confirms that (11-55)NCp7 binds all over the cTAR sequence. In positions AP9 and AP49, the decrease of the amplitude associated to the stacked conformations was much more pronounced, suggesting a preferential binding site at these positions. The amplitude of Φ_1 were significantly decreased in AP9 (Figure 4.E), AP28, AP35, AP49, AP53 and AP55 but not AP17, AP21 or AP45. The local motion of the probe, characterized by the excursion angle of the hindered rotation, was thus decreased in presence of (11-55)NCp7, with a reduction of about 7° to 20° (30-50%) of the cone semi-angle in which the probe rotates. As a consequence, the binding of NCp7 reduced in a site-specific manner the degree of freedom of the 2-AP. The shortenings of the longer rotational correlation time responsible for the differences observed in absence of protein between the upper part of the sequence compared to the bottom were in presence of NCp7 shifted towards the very end of the cTAR sequence and limited to the positions AP53 and AP55 when (11-55)NCp7 fully coated the cTAR sequence. The Φ_3 value were indeed higher for all the 2AP positions (except AP53 and AP55) than the 30ns expected for the theoretical rotational correlation time value of a sphere with the same molecular mass (72kDa) [58], suggesting that the global shape of the complex differed from a sphere. This also suggested that the rotational correlation time mainly described the overall tumbling motion of the complex, with marginal contribution of segmental motion. As a consequence, the binding of (11-55)NCp7 on cTAR resulted in a decrease of the local mobility of the DNA segments containing AP9, AP45 and AP49, which were shown to undergo higher segmental motions in the absence of (11-55)NCp7.

Equimolar mixture of NCp7 and cTAR

A special care was given on the equimolar mixture of cTAR and (11-55)NCp7 measurements. Since it corresponds to conditions under which statistically, one NCp7 is bound per cTAR, changes in the fluorescence of the 2Ap should occur for positions where the protein is preferentially bound. The fluorescence of the 2-Ap was differently affected according to the locations of the 2AP and this was more pronounced compared to the situation where the (11-55)NCp7 was fully coating the cTAR. Indeed, for AP45 almost no change of the quantum yield of the 2AP occurred. There were also no changes in the lifetime parameters, although the protein was bound on the cTAR sequence as seen from the increase of the longer rotational correlation time. As a consequence, we can assume that the protein did not bind to cTAR in the surrounding of AP45. On the contrary, the quantum yield of the 2Ap in positions AP9 and AP49 was nearly doubled. The redistributions within the different conformations for these two locations were also the more important, clearly showing that the protein was more likely to bind on that portion of cTAR. Slighter changes were also observed for AP17, AP21, AP35, AP53 or AP55, indicating the protein was able to bind all over the cTAR sequence with a low affinity. Since the fluorescence of the 2-AP was differently affected but no dramatic differences were observed between the different locations of the 2AP, we can rule out that despite the fact that 2AP allowed us to evidence two preferential binding sites on G10 and G50, no strong affinity binding sites for (11-55)NCp7 were observed in cTAR.

Discussion

In this work, we investigated the steady-state and time-resolved fluorescence of different 2Ap substituted cTAR in order to monitor the structure and dynamics of the stem-loop sequence and to locate specific positions for (11-55)NCp7 binding. We probed heterogeneous base stacking patterns of the 2Ap. Indeed, in the absence of protein, the 2AP inserted in the cTAR sequence adopts at least five different conformations. The strongly quenched conformations represented nearly 90% of the whole conformations. The fluorescence signals of the 2Ap located in different positions within

cTAR was in line with the mfold structure prediction of cTAR, except for the very last ds-segment of cTAR that we found annealed at 20°C in our conditions. It was also in good agreement with the NMR characterization of the top-half of cTAR [51]. Moreover, cTAR appeared highly dynamic at the pico-nanosecond time scale, especially the lower half part of cTAR that underwent segmental motions, in line with the fraying process that governs the dynamics of the both ends of cTAR stem. The structures and dynamics of the terminal hairpin-loop were highly heterogeneous, while the internal loop appeared unexpectedly constrained. The latter appeared to confirm previous observations showing that the residues of the internal loop undergo slow motions (micro-to-millisecond) at the NMR timescale [51].

Taken together, these observations constitute thus a refined structural model regarding the environment and the space available for rotational motion of the bases within the PBS loop, valuable for validating molecular dynamics simulations.

The binding of NCp7 was monitored through structure and dynamics perturbations of the DNA in the environment of the 2Ap in response to the binding of (11-55)NCp7. The binding of the protein induced local fluorescence changes which were dependent of both the location of the 2Ap within the sequence and of the amount of protein bound to cTAR. The protein was shown to bind all over the cTAR sequence with no evident high-affinity binding sites. Nevertheless, two preferential binding site were evidenced through AP9 and AP49 fluorescence signals. Our observations thus suggest the existence of at least two binding mode for the protein: the first one, which appeared to be sequence-independent, was responsible for the slight increase of the fluorescence and for the very limited redistribution towards the less stacked conformers of the 2AP, which suggest that these 'outside complexes' are very transient and that the protein remain highly mobile. The second, observed only for positions AP9 and AP49, induced stronger conformational redistribution of the 2AP. Since dark-species were associated to conformations where the probe stacks with one of its neighbors, resulting in an "immediate" quenching [39], a shift towards less stacked populations is likely the consequence of the direct interaction of the Trp37 of NCp7 with the 2Ap adjacent guanine residue in the DNA sequence, thus preventing its collision with 2-AP and reducing the quenching [31, 34, 39]. Moreover, both the stacked conformations of AP9 and AP49 (the sum of α_0 and α_1 equal 57% and 55% respectively) and the QY (~0.10-0.11) were very similar to those observed for the AATGApC or AACGApC sequences in the presence of saturating concentrations of (11-55)NCp7 where the quantum yield was ~0.10-0.13 and the quenched and unstacked conformations represented of about 50% and 30-35% respectively. , being the signature of the direct interaction of the Trp37 with the Guanine flanking the 2AP in a model extrapolated from the well-resolved structure of (11-55)NCp7 binding on the ACGCC sequence (Morellet98). Taken together, these data clearly indicated two preferential binding sites of the (11-55)NCp7 located on both sides of the long

central ds-segment of cTAR, and involving a direct interaction of the Trp37 of the (11-55)NCp7 with the G10 and the G50 residues, resulting in a significant decrease of the stacking interactions of the 2AP with its neighbouring bases. This type of interaction leads to longer time-residency and to a stronger local destabilization of the DNA structure.

The Trp37 invasion and the “freezing” of the ODN mobility likely constitute a prerequisite for the NC-induced destabilization of the nucleic-acids. 2-AP offers thus the possibility to site-specifically probe the destabilization within the cTAR stem. The fast binding-unbinding events which make the protein rapidly scanning the DNA sequence has been shown to be modulated by positions on which (11-55)NCp7 makes some “breathings”, defining thus some preferential positions. The major changes in the DNA structure and dynamics were observed in the lower half of the stem. G10 and G50, the two guanines involved in the preferential binding of NC, are flanked by a C-A mismatch. This suggests that the destabilisation of the cTAR could be initiated preferentially where the sequence is less stable. The bases within mismatch base-pairs are more prone than those involved in conventional base pair to extrahelical conformations. They thus undergo larger conformational explorations. Since NCp7 was found to preferentially bind the neighbouring of these positions, it may be concluded that the destabilizing activity of NCp7 could result from a conformational capture of the transiently destabilized conformations of cTAR [59].

The anisotropy decay evidenced that NCp7 strongly constrains the ODN dynamics when bound. The restriction in the flexibility of the oligonucleotide likely prevents a rapid pairing of the captured DNA domain with its complementary sequence on the cTAR opposite strand, thus allowing the formation of a transient locally destabilized domain. When one NCp7 is bound on AP9 in a 1:1 ratio with cTAR, no perturbation in the signal of the 2AP in position AP45 which is only 4bp apart from AP9 is observed. Taken together with the absence of spectral shift in the fluorescence emission spectra, this clearly indicate that the structural variations of the cTAR oligonucleotide remains limited, a prerequisite for fast binding-unbinding events [25] imagined to explain the chaperone properties of the protein. Moreover, changes in the fluorescence signal of AP45 occurred for higher concentrations of NC compared to that of the lower labeling positions within the stem. We can thus hypothesis that binding of the protein in the proximity of AP45 occurs when the ds-segments containing AP45 was already previously destabilized by the binding of (11-55)NCp7 to the G10 or to the G50. Thus, an effective chaperone activity of the NCp7 would require a sufficient local concentration of NCp7 on the oligonucleotide, which is obviously the case in the high molecular weights complexes observed in the viral context. The requirement for a sufficient level of coating to achieve destabilization of the nucleic acid sequence far from its initiation site demonstrates that the destabilizing activity of NCp7 could be cooperative. The initial and limited destabilization can indeed be extended and stabilized by the binding of supplementary NCp7 or even by the fusion of

contiguous destabilized area, resulting in a larger destabilization domain.

Since these domains result from the fast stochastic binding of NCp7 on cTAR, they should undergo complex dynamics on large time-scales. In this context, the action of NCp7 to freeze the cTAR dynamics promote the formation of a vast number of reactive cTAR Y-shape with multiple dynamics and conformations, generating the reacting species needed for the necessary annealing of TAR with its complementary cTAR occurring during the first strand transfer at the early stage of the viral replication.

Materials and methods

The NC(11-55) (Figure 1) was synthesized as previously described [60] and stored lyophilized in its zinc-bound form. An extinction coefficient of $5.700 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nm was used to determine its concentration.

The oligonucleotides labelled selectively in the different positions with 2'-deoxyribosyl-2-aminopurine (2-AP) (Figure 1) were synthesized at 1 μmol scale by IBA GmbH Nucleic Acids Product Supply (Göttingen, Germany) and purified by the manufacturer by double reverse-phase HPLC.

Steady-state fluorescence

Experiments were performed in 25 mM Tris-HCl, pH 7.5, 30 mM NaCl and 0.2 mM MgCl_2 at 20°C. Excitation and emission spectra were registered with a Fluorolog or FluoroMax spectrofluorometers (Jobin Yvon Instruments, S.A. Inc.) equipped with a thermostated cell compartment. All fluorescence intensities were corrected for buffer emission and lamp fluctuations.

Time-resolved fluorescence spectroscopy

Time-resolved fluorescence measurements were performed with a time-correlated, single-photon counting technique. Excitation pulses were generated by a pulse-picked frequency-tripled titanium-sapphire laser (Tsunami, Spectra Physics) pumped by a Millennia X laser (Spectra Physics). The excitation pulses were at 315nm (third harmonic), with a repetition rate of 4MHz. The emission was collected at 570nm through a 16mm band-pass monochromator (Jobin-Yvon H10). The single-photon events were detected with a micro-channel plate Hamamatsu R3809U photomultiplier coupled to a Philips 6954 pulse pre-amplifier and recorded on a multi-channel analyser (Ortec 7100) calibrated at 25.5ps/channel. The instrumental response function (IRF) was recorded using a polished aluminum reflector. The full-width at half-maximum (FWHM) was ~ 40 ps. In fluorescence lifetime measurements, the emission was monitored at the magic angle (54.7°) to eliminate the contribution from the decay of anisotropy. Experimentally measured time-fluorescence fluorescence decays were deconvoluted with the IRF and fitted to retrieve the most probable continuous lifetime distribution defined by the maximum entropy method [61]. As the resulting lifetime distribution four well individualized lifetimes clusters, the decays were summarized assuming a sum of discrete exponentials, $I(t)=\sum\alpha_i\times\exp(-t/\tau_i)$ where α_i are the relative amplitude associated to the fluorescence lifetime τ_i representing a distribution of conformation in which 2Ap experience similar quenching rates. The mean lifetime ($\langle\tau\rangle$) was calculated from individual fluorescence lifetimes τ_i and their relative amplitude α_i according to $\langle\tau\rangle=\sum\alpha_i\tau_i$. The population α_0 of dark species was calculated by $\alpha_0 = 1 - \tau_{free}/(\tau_{sample} \times Rm)$

where τ_{free} is the lifetime of the free 2-AP riboside, τ_{sample} is the measured lifetime of 2-AP within the oligonucleotide (either free or bound to NCp7) and Rm is the ratio of their corresponding

steady-state fluorescence. The remaining amplitudes of each fluorescent population were recalculated according to $\alpha_c = \alpha_i (1 - \alpha_0)$. All fitting procedures were carried out both with the maximum entropy method (MEM) (Pulse5 software) and a self-made software based on an iterative application of the Marquardt's method in nonlinear least-squares analysis (kindly provided by P^r G. Krishnamoorthy).

In time resolved anisotropy measurements, the emission was collected at directions parallel $I_{//}(t)$ and perpendicular $I_{\perp}(t)$, alternatively, to the excitation beam polarization (interposition of a quartz crystal rotating the beam polarization by 90°). Time-resolved anisotropy decays were analysed according to:

$$I_{//} = I(t)[1 + 2r(t)]/3 \quad \text{Equation (1)}$$

$$I_{\perp} = I(t)[1 - r(t)]/3 \quad \text{Equation (2)}$$

$$r(t) = \frac{I_{//} - G \times I_{\perp}}{I_{//} + 2G \times I_{\perp}} = r_0 \sum_i \beta_i \times \exp(-t/\phi_i) \quad \text{Equation (3)}$$

where G is a correcting factor at the wavelength λ of emission (G values were 1-1.02). The shorter component ϕ_i , associated to the internal motion of the probe, could be modelled as due to hindered rotation [62, 63]. The cone semi-angle (θ_0) of the angular range of the hindered rotation is given by

$$\theta_0 = \cos^{-1}(0.5 \times (\sqrt{1 + 8S} - 1))$$

where S is the generalized order parameter;

$$S = \left(\frac{\beta'_2 + \beta'_3}{\beta'_1 + \beta'_2 + \beta'_3} \right)^{0.5}$$

β'_1 , β'_2 and β'_3 are the amplitudes of rotational correlation times, recalculated such that $\sum \beta'_i = r_0$.

Since the shortest lifetime and rotational correlation times are in the range of 0.08 - 0.45 ns, too close to both the width of the IRF and the time per channel used, we measured the rotational correlation time (ϕ) of the free 2-AP riboside in sucrose-water mixtures with increasing viscosity (η) in the range of 1-8 cP to check the reliability of these values. The ϕ values obtained at 20°C were 0.080 ± 0.001 ns, 0.237 ± 0.002 ns, 0.478 ± 0.029 ns and 0.850 ± 0.006 ns for viscosities of respectively 1.0 cP, 1.9 cP, 3.9 cP and 7.8 cP, as expected from the Stokes-Einstein relationship ($\phi = \eta \cdot V/kT$) ($r=0.994$; $p<0.001$).

References

1. Morellet, N., et al., *Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR*. EMBO J, 1992. **11**(8): p. 3059–3065.
2. Summers, M.F., et al., *Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1*. Protein Sci, 1992. **1**(5): p. 563–574.
3. Darlix, J.L., et al., *Nucleocapsid protein of human immunodeficiency virus as a model protein with chaperoning functions and as a target for antiviral drugs*. Adv Pharmacol, 2000. **48**: p. 345–372.
4. Levin, J.G., et al., *Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism*. Prog Nucleic Acid Res Mol Biol, 2005. **80**: p. 217–286.
5. Herschlag, D., *RNA chaperones and the RNA folding problem*. J Biol Chem, 1995. **270**(36): p. 20871–4.
6. Rein, A., L.E. Henderson, and J.G. Levin, *Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication*. Trends Biochem Sci, 1998. **23**(8): p. 297–301.
7. Darlix, J.L., et al., *Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA*. C R Acad Sci III, 1993. **316**(8): p. 763–71.
8. Driscoll, M.D. and S.H. Hughes, *Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences*. J Virol, 2000. **74**(19): p. 8785–92.
9. Guo, J., et al., *Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA*. J Virol, 1997. **71**(7): p. 5178–5188.
10. Lapadat-Tapolsky, M., et al., *Possible roles of HIV-1 nucleocapsid protein in the specificity of proviral DNA synthesis and in its variability*. J Mol Biol, 1997. **268**(2): p. 250–60.
11. You, J.C. and C.S. McHenry, *Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription*. J Biol Chem, 1994. **269**(50): p. 31491–31495.
12. Godet, J., et al., *During the early phase of HIV-1 DNA synthesis, nucleocapsid protein directs hybridization of the TAR complementary sequences via the ends of their double-stranded stem*. J Mol Biol, 2006. **356**(5): p. 1180–1192.
13. Vo, M.-N., et al., *HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper"*. J Mol Biol, 2009. **386**(3): p. 789–801.
14. Lapadat-Tapolsky, M., et al., *Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1*. Nucleic Acids Res, 1995. **23**(13): p. 2434–41.
15. Azoulay, J., et al., *Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations*. J Mol Biol, 2003. **326**(3): p. 691–700.
16. Beltz, H., et al., *Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7*. J Mol Biol, 2003. **328**(1): p. 95–108.
17. Bernacchi, S., et al., *HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence*. J Mol Biol, 2002. **317**(3): p. 385–399.
18. Cosa, G., et al., *Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein*. Biophys J, 2004. **87**(4): p. 2759–2767.
19. Cosa, G., et al., *Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned opening of DNA hairpins*. J Phys Chem B, 2006. **110**(5): p. 2419–2426.
20. Liu, H.-W., et al., *Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription*. Biophys J, 2005. **89**(5): p. 3470–3479.
21. Beltz, H., et al., *Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis*. J Mol Biol, 2005. **348**(5): p. 1113–1126.
22. Liu, H.-W., et al., *Insights on the role of nucleic acid/protein interactions in chaperoned nucleic acid rearrangements of HIV-1 reverse transcription*. Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5261–5267.
23. Zeng, Y., et al., *Probing nucleation, reverse annealing, and chaperone function along the reaction path of HIV-1 single-strand transfer*. Proc Natl Acad Sci U S A, 2007. **104**(31): p. 12651–12656.
24. Williams, M.C., et al., *Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching*. Proc Natl Acad Sci U S A, 2001. **98**(11): p. 6121–6126.
25. Cruceanu, M., et al., *Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function*. J Mol Biol, 2006. **363**(5): p. 867–877.
26. Berkowitz, R., J. Fisher, and S.P. Goff, *RNA packaging*. Curr Top Microbiol Immunol, 1996. **214**: p. 177–218.
27. Fisher, R.J., et al., *Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides*. J Virol, 1998. **72**(3): p. 1902–1909.
28. Mély, Y., et al., *Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3Lys), in vitro, is essentially not specific*. J Biol Chem, 1995. **270**(4): p. 1650–1656.
29. Tsuchihashi, Z. and P.O. Brown, *DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein*. J Virol, 1994. **68**(9): p. 5863–5870.
30. Urbaneja, M.A., et al., *Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA: elucidation of the structural determinants for function*. J Mol Biol, 1999. **287**(1): p. 59–75.
31. Vuilleumier, C., et al., *Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study*. Biochemistry, 1999. **38**(51): p. 16816–16825.
32. Amarasinghe, G.K., et al., *NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition*. J Mol Biol, 2000. **301**(2): p. 491–511.
33. Shubsda, M.F., et al., *Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein*. Biochemistry, 2002. **41**(16): p. 5276–5282.
34. Morellet, N., et al., *Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC)*. J Mol Biol, 1998. **283**(2): p. 419–434.

35. Amarasinghe, G.K., et al., *Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition.* J Mol Biol, 2001. **314**(5): p. 961--970.
36. De Guzman, R.N., et al., *Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element.* Science, 1998. **279**(5349): p. 384-8.
37. Tisne, C., B.P. Roques, and F. Dardel, *Heteronuclear NMR studies of the interaction of tRNA(Lys)₃ with HIV-1 nucleocapsid protein.* J Mol Biol, 2001. **306**(3): p. 443-54.
38. Avilov, S., et al., *Site-specific characterization of HIV-1 nucleocapsid protein binding to oligonucleotides with two binding sites.* Biochemistry, 2009.
39. Avilov, S.V., et al., *Probing dynamics of HIV-1 nucleocapsid protein target hexanucleotide complexes by 2-aminopurine.* Nucleic Acids Res, 2008. **36**(3): p. 885--896.
40. Guest, C.R., et al., *Dynamics of mismatched base pairs in DNA.* Biochemistry, 1991. **30**(13): p. 3271-9.
41. Jean, J.M. and K.B. Hall, *2-Aminopurine fluorescence quenching and lifetimes: role of base stacking.* Proc Natl Acad Sci U S A, 2001. **98**(1): p. 37-41.
42. Nordlund, T.M., et al., *Structure and dynamics of a fluorescent DNA oligomer containing the EcoRI recognition sequence: fluorescence, molecular dynamics, and NMR studies.* Biochemistry, 1989. **28**(23): p. 9095-103.
43. Wan, C., et al., *Ultrafast unequilibrated charge transfer: A new channel in the quenching of fluorescent biological probes.* Chemical Physics Letters, 2005. **412**(1-3): p. 158-163.
44. Ward, D.C., E. Reich, and L. Stryer, *Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives.* J Biol Chem, 1969. **244**(5): p. 1228--1237.
45. Jean, J.M. and K.B. Hall, *Stacking-unstacking dynamics of oligodeoxynucleotide trimers.* Biochemistry, 2004. **43**(31): p. 10277-84.
46. Kenfack, C.A., et al., *Time-resolved fluorescent properties of 8-vinyl-deoxyadenosine and 2-amino-deoxyribosylpurine exhibit different sensitivity to their opposite base in duplexes.* J Phys Chem B, 2008. **112**(32): p. 9736--9745.
47. Sowers, L.C., et al., *Base pairing and mutagenesis: observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR.* Proc Natl Acad Sci U S A, 1986. **83**(15): p. 5434-8.
48. Eritja, R., et al., *Synthesis and properties of defined DNA oligomers containing base mismatches involving 2-aminopurine.* Nucleic Acids Res, 1986. **14**(14): p. 5869-84.
49. Fagan, P.A., et al., *NMR study of the conformation of the 2-aminopurine:cytosine mismatch in DNA.* Biochemistry, 1996. **35**(13): p. 4026-33.
50. Hochstrasser, R.A., et al., *Melting of a DNA helix terminus within the active site of a DNA polymerase.* Biochemistry, 1994. **33**(39): p. 11971-9.
51. Zargarian, L., et al., *Structural and dynamic characterization of the upper part of the HIV-1 cTAR DNA hairpin.* Nucleic Acids Res, 2009.
52. Jiao, Y., S. Stringfellow, and H. Yu, *Distinguishing "looped-out" and "stacked-in" DNA bulge conformation using fluorescent 2-aminopurine replacing a purine base.* J Biomol Struct Dyn, 2002. **19**(5): p. 929-34.
53. Jean, J.M. and K.B. Hall, *2-Aminopurine electronic structure and fluorescence properties in DNA.* Biochemistry, 2002. **41**(44): p. 13152-61.
54. Xu, D., K.O. Evans, and T.M. Nordlund, *Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption.* Biochemistry, 1994. **33**(32): p. 9592-9.
55. Beltz, H., et al., *Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7.* J Mol Biol, 2004. **338**(4): p. 711--723.
56. Wan, C., et al., *Femtosecond direct observation of charge transfer between bases in DNA.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14052-5.
57. Larsen, O.F., et al., *Probing the structure and dynamics of a DNA hairpin by ultrafast quenching and fluorescence depolarization.* Biophys J, 2001. **81**(2): p. 1115--1126.
58. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy.* 2nd ed. 1999: Kluwer Academic/Plenum Publishers.
59. Leulliot, N. and G. Varani, *Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture.* Biochemistry, 2001. **40**(27): p. 7947-56.
60. de Rocquigny, H., et al., *First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form.* Biochem Biophys Res Commun, 1991. **180**(2): p. 1010-8.
61. Brochon, J.C., *Maximum entropy method of data analysis in time-resolved spectroscopy.* Methods Enzymol, 1994. **240**: p. 262-311.
62. Kinoshita, K., Jr., S. Kawato, and A. Ikegami, *A theory of fluorescence polarization decay in membranes.* Biophys J, 1977. **20**(3): p. 289-305.
63. Lipari, G. and A. Szabo, *Effect of librational motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes.* Biophys J, 1980. **30**(3): p. 489--506.
64. Bombarda, E., et al., *Time-resolved fluorescence investigation of the human immunodeficiency virus type 1 nucleocapsid protein: influence of the binding of nucleic acids.* Biophys J, 1999. **76**(3): p. 1561--1570.
65. Egelé, C., et al., *The single-finger nucleocapsid protein of moloney murine leukemia virus binds and destabilizes the TAR sequences of HIV-1 but does not promote efficiently their annealing.* Biochemistry, 2007. **46**(50): p. 14650--14662.

Table 1 A: Time-resolved intensity parameters for the 2-Ap-substituted cTAR.

| | NCp7 (Eq) | Q.Y. | α_0 | τ_1 | α_1 | α_{1c} | τ_2 | α_2 | α_{2c} | τ_3 | α_3 | α_{3c} | τ_4 | α_4 | α_{4c} | $\langle\tau\rangle$ |
|-------------|--------------|--------|------------|----------|------------|---------------|----------|------------|---------------|----------|------------|---------------|----------|------------|---------------|----------------------|
| Free 2AP | | | | | | | | | | | | | 10.20 | 1.00 | 1.00 | 10.20 |
| AP9 | 0 | 0.0140 | 0.71 | 0.08 | 0.72 | 0.21 | 0.58 | 0.14 | 0.04 | 2.60 | 0.10 | 0.03 | 7.30 | 0.04 | 0.01 | 0.73 |
| AP9 | 1 | 0.0237 | 0.67 | 0.08 | 0.60 | 0.20 | 0.66 | 0.18 | 0.06 | 2.80 | 0.14 | 0.05 | 7.10 | 0.08 | 0.03 | 1.08 |
| AP9 | 4 | 0.0454 | 0.67 | 0.17 | 0.35 | 0.12 | 0.76 | 0.26 | 0.09 | 3.00 | 0.24 | 0.08 | 7.50 | 0.15 | 0.05 | 2.04 |
| AP9 | 7 | 0.0665 | 0.51 | 0.11 | 0.40 | 0.19 | 0.77 | 0.25 | 0.12 | 3.40 | 0.21 | 0.10 | 8.20 | 0.14 | 0.07 | 2.05 |
| AP9 | 10 | 0.0834 | 0.48 | 0.11 | 0.30 | 0.16 | 0.71 | 0.28 | 0.15 | 3.00 | 0.22 | 0.11 | 7.70 | 0.20 | 0.10 | 2.40 |
| AP9 | 15 | 0.0990 | 0.48 | 0.11 | 0.20 | 0.10 | 0.65 | 0.30 | 0.16 | 3.32 | 0.26 | 0.13 | 7.80 | 0.23 | 0.12 | 2.87 |
| AP17 | 0 | 0.0056 | 0.76 | 0.05 | 0.81 | 0.19 | 0.49 | 0.12 | 0.03 | 2.30 | 0.05 | 0.01 | 6.91 | 0.02 | 0.00 | 0.35 |
| AP17 | 1 | 0.0086 | 0.77 | 0.07 | 0.76 | 0.18 | 0.48 | 0.14 | 0.03 | 2.65 | 0.07 | 0.02 | 7.35 | 0.03 | 0.01 | 0.55 |
| AP17 | 4 | 0.0170 | 0.73 | 0.09 | 0.68 | 0.18 | 0.61 | 0.15 | 0.04 | 2.89 | 0.10 | 0.03 | 7.82 | 0.06 | 0.02 | 0.94 |
| AP17 | 7 | 0.0252 | 0.70 | 0.09 | 0.59 | 0.18 | 0.555 | 0.19 | 0.06 | 2.84 | 0.13 | 0.04 | 8.19 | 0.09 | 0.03 | 1.24 |
| AP17 | 10 | 0.0299 | 0.69 | 0.11 | 0.58 | 0.18 | 0.71 | 0.19 | 0.06 | 3.06 | 0.13 | 0.04 | 8.29 | 0.10 | 0.03 | 1.43 |
| AP17 | 15 | 0.0337 | 0.65 | 0.09 | 0.58 | 0.21 | 0.70 | 0.18 | 0.06 | 3.05 | 0.13 | 0.05 | 8.26 | 0.10 | 0.04 | 1.43 |
| AP21 | 0 | 0.0049 | 0.92 | 0.06 | 0.72 | 0.06 | 0.46 | 0.12 | 0.01 | 2.682 | 0.09 | 0.01 | 8.48 | 0.07 | 0.01 | 0.90 |
| AP21 | 1 | 0.0069 | 0.92 | 0.061 | 0.67 | 0.05 | 0.519 | 0.13 | 0.01 | 2.852 | 0.10 | 0.01 | 8.810 | 0.10 | 0.01 | 1.27 |
| AP21 | 4 | 0.0167 | 0.85 | 0.041 | 0.61 | 0.09 | 0.438 | 0.14 | 0.02 | 3.049 | 0.11 | 0.02 | 9.157 | 0.14 | 0.02 | 1.71 |
| AP21 | 7 | 0.0259 | 0.77 | 0.045 | 0.59 | 0.13 | 0.418 | 0.14 | 0.03 | 2.934 | 0.13 | 0.03 | 8.980 | 0.14 | 0.03 | 1.72 |
| AP21 | 10 | 0.0295 | 0.70 | 0.037 | 0.61 | 0.18 | 0.427 | 0.16 | 0.05 | 2.930 | 0.12 | 0.04 | 8.747 | 0.12 | 0.04 | 1.46 |
| AP21 | 15 | 0.0326 | 0.69 | 0.054 | 0.60 | 0.18 | 0.489 | 0.15 | 0.05 | 3.114 | 0.13 | 0.04 | 8.755 | 0.12 | 0.04 | 1.60 |
| AP28 | 0 | 0.0325 | 0.69 | 0.11 | 0.68 | 0.21 | 0.76 | 0.10 | 0.03 | 3.60 | 0.07 | 0.02 | 8.33 | 0.14 | 0.04 | 1.57 |
| AP28 | 1 | 0.0300 | 0.63 | 0.11 | 0.69 | 0.26 | 0.68 | 0.13 | 0.05 | 3.00 | 0.08 | 0.03 | 8.00 | 0.10 | 0.04 | 1.20 |
| AP28 | 4 | 0.0218 | 0.73 | 0.11 | 0.62 | 0.17 | 0.67 | 0.18 | 0.05 | 2.90 | 0.10 | 0.03 | 8.10 | 0.09 | 0.02 | 1.21 |
| AP28 | 7 | 0.0199 | 0.72 | 0.08 | 0.68 | 0.19 | 0.76 | 0.12 | 0.03 | 2.60 | 0.12 | 0.03 | 7.60 | 0.08 | 0.02 | 1.07 |
| AP28 | 10 | 0.0203 | 0.81 | 0.11 | 0.55 | 0.10 | 0.76 | 0.19 | 0.04 | 3.80 | 0.15 | 0.03 | 8.20 | 0.10 | 0.02 | 1.59 |
| AP28 | 15 | 0.0237 | 0.84 | 0.08 | 0.42 | 0.07 | 0.67 | 0.21 | 0.03 | 3.72 | 0.19 | 0.03 | 7.80 | 0.18 | 0.03 | 2.29 |
| AP35 | 0 | 0.0860 | 0.20 | 0.18 | 0.43 | 0.34 | 0.96 | 0.20 | 0.16 | 2.90 | 0.32 | 0.26 | 8.21 | 0.05 | 0.04 | 1.61 |
| AP35 | 1 | 0.0973 | 0.12 | 0.19 | 0.45 | 0.40 | 1.10 | 0.20 | 0.18 | 3.10 | 0.30 | 0.26 | 8.45 | 0.05 | 0.04 | 1.66 |
| AP35 | 4 | 0.1163 | 0.09 | 0.19 | 0.48 | 0.43 | 1.23 | 0.19 | 0.17 | 3.80 | 0.24 | 0.22 | 8.60 | 0.08 | 0.07 | 1.92 |
| AP35 | 7 | 0.1376 | 0.11 | 0.16 | 0.48 | 0.43 | 1.24 | 0.16 | 0.14 | 4.20 | 0.23 | 0.20 | 8.28 | 0.13 | 0.12 | 2.32 |
| AP35 | 10 | 0.1548 | 0.15 | 0.18 | 0.42 | 0.36 | 1.00 | 0.13 | 0.11 | 3.50 | 0.19 | 0.16 | 7.43 | 0.25 | 0.21 | 2.72 |
| AP35 | 15 | 0.1617 | 0.20 | 0.17 | 0.39 | 0.31 | 0.85 | 0.13 | 0.10 | 3.70 | 0.19 | 0.15 | 7.52 | 0.29 | 0.23 | 3.05 |

Table 1 B: Time-resolved intensity parameters for the 2-*Ap*-substituted cTAR.

| | NCp7 (Eq) | Q.Y. | α_0 | τ_1 | α_1 | α_1c | τ_2 | α_2 | α_2c | τ_3 | α_3 | α_3c | τ_4 | α_4 | α_4c | $\langle\tau\rangle$ |
|-------|--------------|--------|------------|----------|------------|-------------|----------|------------|-------------|----------|------------|-------------|----------|------------|-------------|----------------------|
| AP45 | 0 | 0.0180 | 0.56 | 0.11 | 0.77 | 0.34 | 0.76 | 0.12 | 0.05 | 2.65 | 0.06 | 0.03 | 6.95 | 0.04 | 0.02 | 0.61 |
| AP45 | 1 | 0.0182 | 0.59 | 0.11 | 0.75 | 0.31 | 0.76 | 0.13 | 0.05 | 2.86 | 0.07 | 0.03 | 7.20 | 0.04 | 0.02 | 0.67 |
| AP45 | 4 | 0.0244 | 0.67 | 0.12 | 0.54 | 0.18 | 0.65 | 0.23 | 0.08 | 2.62 | 0.14 | 0.05 | 7.60 | 0.07 | 0.02 | 1.11 |
| AP45 | 7 | 0.0299 | 0.70 | 0.13 | 0.38 | 0.12 | 0.58 | 0.33 | 0.10 | 2.46 | 0.20 | 0.06 | 8.30 | 0.09 | 0.03 | 1.48 |
| AP45 | 10 | 0.0377 | 0.73 | 0.13 | 0.34 | 0.09 | 0.69 | 0.28 | 0.08 | 3.10 | 0.21 | 0.06 | 7.80 | 0.15 | 0.04 | 2.06 |
| AP45 | 15 | 0.0533 | 0.70 | 0.10 | 0.28 | 0.08 | 0.69 | 0.25 | 0.07 | 3.16 | 0.22 | 0.07 | 7.81 | 0.23 | 0.07 | 2.69 |
| AP49 | 0 | 0.0083 | 0.85 | 0.14 | 0.62 | 0.09 | 0.81 | 0.23 | 0.04 | 3.02 | 0.10 | 0.02 | 6.11 | 0.04 | 0.01 | 0.82 |
| AP49 | 1 | 0.0178 | 0.78 | 0.16 | 0.57 | 0.13 | 0.81 | 0.24 | 0.05 | 3.15 | 0.11 | 0.02 | 8.07 | 0.07 | 0.02 | 1.20 |
| AP49 | 4 | 0.0414 | 0.52 | 0.13 | 0.58 | 0.28 | 0.93 | 0.22 | 0.10 | 3.58 | 0.12 | 0.06 | 8.52 | 0.07 | 0.03 | 1.31 |
| AP49 | 7 | 0.0660 | 0.37 | 0.13 | 0.52 | 0.33 | 0.91 | 0.24 | 0.15 | 3.66 | 0.14 | 0.09 | 8.70 | 0.09 | 0.06 | 1.58 |
| AP49 | 10 | 0.0835 | 0.27 | 0.13 | 0.49 | 0.36 | 0.89 | 0.24 | 0.17 | 3.61 | 0.15 | 0.11 | 8.62 | 0.11 | 0.08 | 1.72 |
| AP49 | 15 | 0.1077 | 0.24 | 0.16 | 0.44 | 0.33 | 1.00 | 0.24 | 0.18 | 3.89 | 0.18 | 0.14 | 8.65 | 0.13 | 0.10 | 2.13 |
| AP53 | 0 | 0.0162 | 0.82 | 0.05 | 0.76 | 0.14 | 0.57 | 0.07 | 0.01 | 3.92 | 0.06 | 0.01 | 9.55 | 0.11 | 0.02 | 1.37 |
| AP53 | 1 | 0.0178 | 0.79 | 0.04 | 0.74 | 0.16 | 0.53 | 0.09 | 0.02 | 3.31 | 0.07 | 0.01 | 9.27 | 0.10 | 0.02 | 1.26 |
| AP53 | 4 | 0.0244 | 0.78 | 0.06 | 0.61 | 0.13 | 0.65 | 0.15 | 0.03 | 3.37 | 0.13 | 0.03 | 8.92 | 0.12 | 0.03 | 1.65 |
| AP53 | 7 | 0.0320 | 0.74 | 0.06 | 0.52 | 0.14 | 0.65 | 0.20 | 0.05 | 3.53 | 0.15 | 0.04 | 8.82 | 0.14 | 0.04 | 1.88 |
| AP53 | 10 | 0.0405 | 0.74 | 0.08 | 0.43 | 0.12 | 0.68 | 0.21 | 0.06 | 3.48 | 0.18 | 0.05 | 8.78 | 0.17 | 0.05 | 2.34 |
| AP53 | 15 | 0.0487 | 0.73 | 0.10 | 0.38 | 0.12 | 0.72 | 0.21 | 0.06 | 3.53 | 0.20 | 0.06 | 8.82 | 0.20 | 0.06 | 2.67 |
| AP55 | 0 | 0.0156 | 0.72 | 0.06 | 0.56 | 0.16 | 0.26 | 0.32 | 0.09 | 2.82 | 0.05 | 0.02 | 8.89 | 0.06 | 0.02 | 0.83 |
| AP55 | 1 | 0.0169 | 0.74 | 0.11 | 0.64 | 0.17 | 0.39 | 0.22 | 0.06 | 2.91 | 0.08 | 0.02 | 8.73 | 0.07 | 0.02 | 0.97 |
| AP55 | 4 | 0.0247 | 0.68 | 0.14 | 0.63 | 0.20 | 0.71 | 0.17 | 0.05 | 3.26 | 0.13 | 0.04 | 8.14 | 0.07 | 0.02 | 1.17 |
| AP55 | 7 | 0.0343 | 0.64 | 0.11 | 0.50 | 0.18 | 0.61 | 0.22 | 0.08 | 2.97 | 0.19 | 0.07 | 7.66 | 0.09 | 0.03 | 1.44 |
| AP55 | 10 | 0.0415 | 0.64 | 0.12 | 0.46 | 0.17 | 0.73 | 0.22 | 0.08 | 3.05 | 0.22 | 0.08 | 7.57 | 0.11 | 0.04 | 1.71 |
| AP55 | 15 | 0.0486 | 0.60 | 0.10 | 0.43 | 0.17 | 0.68 | 0.22 | 0.09 | 3.04 | 0.23 | 0.09 | 7.69 | 0.12 | 0.05 | 1.82 |
| AP28c | 0 | 0.0340 | 0.64 | 0.07 | 0.72 | 0.26 | 0.48 | 0.09 | 0.03 | 3.16 | 0.05 | 0.02 | 8.33 | 0.14 | 0.05 | 1.41 |
| AP28c | 1 | 0.0350 | 0.61 | 0.07 | 0.7 | 0.27 | 0.52 | 0.11 | 0.04 | 3.10 | 0.06 | 0.02 | 8.28 | 0.13 | 0.05 | 1.36 |
| AP28c | 4 | 0.0301 | 0.60 | 0.06 | 0.69 | 0.28 | 0.58 | 0.14 | 0.06 | 3.08 | 0.08 | 0.03 | 8.26 | 0.09 | 0.04 | 1.13 |
| AP28c | 7 | 0.0321 | 0.62 | 0.06 | 0.63 | 0.24 | 0.59 | 0.17 | 0.06 | 3.05 | 0.10 | 0.04 | 8.18 | 0.10 | 0.04 | 1.27 |
| AP28c | 10 | 0.0382 | 0.53 | 0.06 | 0.64 | 0.30 | 0.56 | 0.16 | 0.08 | 2.89 | 0.10 | 0.05 | 8.04 | 0.10 | 0.05 | 1.21 |
| AP35c | 0 | 0.0940 | 0.41 | 0.11 | 0.03 | 0.01 | 0.61 | 0.32 | 0.19 | 2.69 | 0.57 | 0.33 | 7.21 | 0.09 | 0.05 | 2.40 |
| AP35c | 1 | 0.1180 | 0.30 | 0.10 | 0.05 | 0.03 | 0.68 | 0.33 | 0.23 | 2.74 | 0.50 | 0.35 | 7.52 | 0.12 | 0.09 | 2.51 |
| AP35c | 4 | 0.1450 | 0.26 | 0.08 | 0.08 | 0.06 | 0.66 | 0.29 | 0.21 | 2.87 | 0.42 | 0.31 | 7.66 | 0.20 | 0.15 | 2.96 |
| AP35c | 7 | 0.1579 | 0.31 | 0.12 | 0.00 | 0.00 | 0.66 | 0.32 | 0.22 | 2.98 | 0.44 | 0.30 | 7.81 | 0.25 | 0.17 | 3.43 |
| AP35c | 10 | 0.1701 | 0.16 | 0.05 | 0.14 | 0.11 | 0.69 | 0.28 | 0.24 | 3.12 | 0.36 | 0.31 | 7.96 | 0.22 | 0.18 | 3.05 |

Table 2A: Time-resolved anisotropy parameters for the 2-Ap-substituted cTAR.

| | | Φ_1 | β_1 | Φ_2 | β_2 | Φ_3 | β_3 | r_0 | $\langle\Phi\rangle$ | S | θ_0 ($\delta\epsilon\gamma$) |
|-----------------|-----------|----------|-----------|----------|-----------|----------|-----------|-------|----------------------|------|---------------------------------------|
| Free 2AP | | 0.08 | 1.00 | - | - | - | - | 0.27 | 0.1 | 0.00 | 90.0 |
| AP9 | 0 | 0.22 | 0.42 | 1.1 | 0.32 | 4.9 | 0.26 | 0.32 | 1.7 | 0.76 | 34 |
| AP9 | 1 | 0.23 | 0.33 | 1.4 | 0.28 | 7.7 | 0.38 | 0.31 | 3.4 | 0.82 | 29.3 |
| AP9 | 4 | 0.22 | 0.27 | 1.1 | 0.22 | 19.4 | 0.51 | 0.30 | 10.2 | 0.85 | 25.9 |
| AP9 | 7 | 0.22 | 0.23 | 1.1 | 0.17 | 23.0 | 0.60 | 0.32 | 14.0 | 0.88 | 23.7 |
| AP9 | 10 | 0.18 | 0.20 | 1.0 | 0.13 | 32.3 | 0.67 | 0.30 | 21.8 | 0.89 | 21.9 |
| AP9 | 15 | 0.18 | 0.21 | 1.0 | 0.08 | 36.0 | 0.71 | 0.30 | 25.7 | 0.89 | 22.5 |
| AP17 | 0 | 0.21 | 0.22 | 1.2 | 0.37 | 9.2 | 0.41 | 0.35 | 4.2 | 0.88 | 23 |
| AP17 | 1 | 0.19 | 0.33 | 1.8 | 0.21 | 16.8 | 0.46 | 0.35 | 8.3 | 0.82 | 28.8 |
| AP17 | 4 | 0.08 | 0.22 | 1.2 | 0.21 | 21.5 | 0.56 | 0.34 | 12.4 | 0.88 | 23.4 |
| AP17 | 7 | 0.20 | 0.22 | 2.8 | 0.22 | 32.0 | 0.57 | 0.35 | 18.8 | 0.89 | 22.8 |
| AP17 | 10 | 0.20 | 0.23 | 2.2 | 0.19 | 30.4 | 0.58 | 0.34 | 18.1 | 0.88 | 23.7 |
| AP17 | 15 | 0.20 | 0.22 | 2.2 | 0.19 | 33.7 | 0.60 | 0.35 | 20.5 | 0.88 | 23.0 |
| AP21 | 0 | 0.20 | 0.33 | 1.5 | 0.35 | 7.6 | 0.32 | 0.35 | 3.0 | 0.82 | 29 |
| AP21 | 1 | 0.12 | 0.34 | 1.4 | 0.28 | 10.5 | 0.38 | 0.40 | 4.4 | 0.81 | 29.8 |
| AP21 | 4 | 0.15 | 0.33 | 2.0 | 0.28 | 17.6 | 0.39 | 0.39 | 7.5 | 0.82 | 29.0 |
| AP21 | 7 | 0.12 | 0.27 | 1.7 | 0.21 | 24.3 | 0.52 | 0.40 | 13.0 | 0.86 | 25.8 |
| AP21 | 10 | 0.11 | 0.32 | 2.3 | 0.19 | 34.0 | 0.49 | 0.40 | 17.2 | 0.83 | 28.5 |
| AP21 | 15 | 0.11 | 0.32 | 2.0 | 0.19 | 34.7 | 0.49 | 0.40 | 17.5 | 0.83 | 28.5 |
| AP28 | 0 | 0.20 | 0.47 | 1.2 | 0.32 | 7.4 | 0.20 | 0.32 | 2.0 | 0.72 | 37 |
| AP28 | 1 | 0.24 | 0.47 | 1.4 | 0.28 | 13.7 | 0.38 | 0.31 | 5.7 | 0.76 | 33.5 |
| AP28 | 4 | 0.20 | 0.37 | 1.2 | 0.18 | 21.0 | 0.38 | 0.34 | 8.3 | 0.78 | 32.6 |
| AP28 | 7 | 0.20 | 0.43 | 1.2 | 0.10 | 22.0 | 0.48 | 0.34 | 10.8 | 0.76 | 34.0 |
| AP28 | 10 | 0.20 | 0.34 | 1.2 | 0.16 | 35.0 | 0.50 | 0.34 | 17.8 | 0.81 | 29.6 |
| AP28 | 15 | 0.20 | 0.21 | 1.2 | 0.26 | 35.0 | 0.53 | 0.30 | 18.9 | 0.89 | 22.5 |
| AP35 | 0 | 0.17 | 0.20 | 1.4 | 0.28 | 8.3 | 0.51 | 0.30 | 4.7 | 0.89 | 22 |
| AP35 | 1 | 0.17 | 0.23 | 1.2 | 0.18 | 14.0 | 0.57 | 0.33 | 8.2 | 0.87 | 23.9 |
| AP35 | 4 | 0.19 | 0.13 | 1.2 | 0.15 | 31.1 | 0.72 | 0.30 | 22.6 | 0.93 | 17.4 |
| AP35 | 7 | 0.15 | 0.10 | 1.2 | 0.18 | 34.0 | 0.72 | 0.27 | 24.7 | 0.95 | 15.1 |
| AP35 | 10 | 0.21 | 0.08 | 1.2 | 0.12 | 33.2 | 0.80 | 0.23 | 26.7 | 0.96 | 13.5 |
| AP35 | 15 | 0.23 | 0.10 | 1.2 | 0.19 | 33.2 | 0.70 | 0.23 | 23.5 | 0.95 | 15.2 |

Table 2B: Time-resolved anisotropy parameters for the 2-Ap-substituted cTAR.

| | | Φ_1 | β_1 | Φ_2 | β_2 | Φ_3 | β_3 | r_0 | $\langle\Phi\rangle$ | S | θ_0 ($\delta\epsilon\gamma$) |
|-------|----|----------|-----------|----------|-----------|----------|-----------|-------|----------------------|------|---------------------------------------|
| AP45 | 0 | 0.17 | 0.35 | 0.6 | 0.31 | 4.3 | 0.33 | 0.34 | 1.6 | 0.80 | 30 |
| AP45 | 1 | 0.22 | 0.40 | 1.6 | 0.30 | 8.5 | 0.30 | 0.34 | 3.1 | 0.77 | 32.7 |
| AP45 | 4 | 0.29 | 0.36 | 2.1 | 0.33 | 19.5 | 0.30 | 0.36 | 6.6 | 0.80 | 30.9 |
| AP45 | 7 | 0.16 | 0.39 | 1.9 | 0.31 | 23.5 | 0.30 | 0.34 | 7.7 | 0.78 | 32.2 |
| AP45 | 10 | 0.07 | 0.43 | 2.7 | 0.09 | 31.0 | 0.48 | 0.34 | 15.1 | 0.75 | 34.2 |
| AP45 | 15 | 0.15 | 0.30 | 2.6 | 0.35 | 33.0 | 0.35 | 0.36 | 12.5 | 0.84 | 27.5 |
| AP49 | 0 | 0.14 | 0.41 | 1.1 | 0.33 | 5.6 | 0.26 | 0.34 | 1.9 | 0.77 | 33 |
| AP49 | 1 | 0.22 | 0.31 | 1.7 | 0.36 | 14.2 | 0.32 | 0.34 | 5.2 | 0.83 | 28.2 |
| AP49 | 4 | 0.15 | 0.25 | 1.5 | 0.25 | 20.3 | 0.50 | 0.34 | 10.5 | 0.87 | 24.8 |
| AP49 | 7 | 0.21 | 0.16 | 2.2 | 0.34 | 28.5 | 0.50 | 0.34 | 15.0 | 0.92 | 19.4 |
| AP49 | 10 | 0.17 | 0.13 | 2.4 | 0.40 | 31.7 | 0.47 | 0.34 | 15.9 | 0.93 | 17.4 |
| AP49 | 15 | 0.14 | 0.10 | 1.8 | 0.33 | 34.1 | 0.56 | 0.34 | 19.7 | 0.95 | 15.2 |
| AP53 | 0 | 0.13 | 0.44 | 0.8 | 0.30 | 4.7 | 0.26 | 0.34 | 1.5 | 0.75 | 35 |
| AP53 | 1 | 0.18 | 0.46 | 1.2 | 0.27 | 7.3 | 0.27 | 0.34 | 2.4 | 0.73 | 36 |
| AP53 | 4 | 0.23 | 0.35 | 2.1 | 0.27 | 15.2 | 0.38 | 0.34 | 6.4 | 0.81 | 30 |
| AP53 | 7 | 0.18 | 0.28 | 1.9 | 0.25 | 20.8 | 0.47 | 0.34 | 10.3 | 0.85 | 26 |
| AP53 | 10 | 0.20 | 0.27 | 2.3 | 0.27 | 24.7 | 0.45 | 0.34 | 11.8 | 0.85 | 26 |
| AP53 | 15 | 0.20 | 0.25 | 2.4 | 0.28 | 26.2 | 0.47 | 0.34 | 13.0 | 0.87 | 25 |
| AP55 | 0 | 0.45 | 0.70 | - | - | 4.4 | 0.30 | 0.33 | 1.6 | 0.55 | 49 |
| AP55 | 1 | 0.43 | 0.55 | 1.7 | 0.12 | 8.5 | 0.32 | 0.33 | 3.2 | 0.67 | 40.7 |
| AP55 | 4 | 0.33 | 0.37 | 1.6 | 0.26 | 13.7 | 0.37 | 0.33 | 5.6 | 0.79 | 31.2 |
| AP55 | 7 | 0.22 | 0.28 | 1.7 | 0.30 | 23.2 | 0.42 | 0.33 | 10.3 | 0.85 | 26.4 |
| AP55 | 10 | 0.19 | 0.30 | 1.9 | 0.29 | 28.4 | 0.41 | 0.33 | 12.3 | 0.84 | 27.5 |
| AP55 | 15 | 0.20 | 0.26 | 1.9 | 0.32 | 29.5 | 0.42 | 0.33 | 13.1 | 0.86 | 25.3 |
| AP28c | 0 | 0.16 | 0.44 | 0.9 | 0.32 | 4.2 | 0.24 | 0.34 | 1.3 | 0.75 | 35 |
| AP28c | 1 | 0.17 | 0.37 | 1.0 | 0.32 | 6.4 | 0.32 | 0.34 | 2.4 | 0.80 | 31.0 |
| AP28c | 4 | 0.15 | 0.35 | 1.0 | 0.20 | 10.6 | 0.45 | 0.35 | 5.0 | 0.80 | 30.3 |
| AP28c | 7 | 0.13 | 0.33 | 1.5 | 0.21 | 14.4 | 0.47 | 0.35 | 7.1 | 0.82 | 29.0 |
| AP28c | 10 | 0.18 | 0.31 | 2.4 | 0.25 | 16.5 | 0.44 | 0.35 | 7.9 | 0.83 | 28.0 |
| AP35c | 0 | 0.08 | 0.16 | 0.7 | 0.28 | 4.5 | 0.57 | 0.34 | 2.8 | 0.92 | 19 |
| AP35c | 1 | 0.17 | 0.16 | 1.2 | 0.30 | 8.8 | 0.54 | 0.33 | 5.1 | 0.92 | 19.5 |
| AP35c | 4 | 0.11 | 0.18 | 1.0 | 0.21 | 11.8 | 0.61 | 0.33 | 7.4 | 0.90 | 20.9 |
| AP35c | 7 | 0.10 | 0.17 | 1.5 | 0.20 | 15.7 | 0.64 | 0.35 | 10.4 | 0.91 | 19.7 |
| AP35c | 10 | 0.11 | 0.17 | 1.5 | 0.17 | 15.0 | 0.66 | 0.34 | 10.1 | 0.91 | 20.2 |

Figure legends:

Figure 1: 2-Ap substituted cTAR and (11-55)NCp7 sequence.

cTAR sequence is from the Lai strain of HIV-1 and is selectively modified either in position 9, 17, 21, 28, 35, 45, 49, 53 or 55 with 2-Ap. The short cTAR_{sup} AP28c and AP35c are modified in the same position than cTAR AP28 and cTAR AP35 respectively.

Figure 2: Quantum yields of the different 2AP-modified cTAR.

The Q.Y. of the 2Ap labelled cTAR at 20°C (blue) or at 65°C (green) are represented together with the Q.Y. of an equimolar mixture of cTAR and (11-55)NCp7 (dark blue) or in the presence of a saturating concentration of (11-55)NCp7 (red).

Figure 3: Temperature dependence of the 2AP substituted cTARs fluorescence.

2-Ap quantum yield are represented as a function of temperature. Two types of transitions were observed when cTAR melts. (A) The 2-Ap transitioned from strongly to less quenched conformations in AP9 (black squares), AP17 (red circles), AP21 (green triangles), AP45 (blue inverted triangles) and AP49 (cyan pentagons). (B) A down transition was observed for AP28 (black squares), AP53 (red circles) and AP55 (green triangles). (C) No melting transition was observed for the 2-Ap in position AP35.

Figure 4: Fluorescence changes in response to (11-55)NCp7 binding.

Quantum yield of the 2Ap are represented (a) in absence of NCp7, and in presence of (b) one (11-55)NCp7 per cTAR or (c) fifteen (11-55)NCp7 per cTAR. The binding of one (11-55)NCp7 to cTAR induced fluorescence changes of the 2-Ap located mostly in positions 9 and 49, suggesting a preferential binding site at these positions. In presence of saturating concentrations of NCp7, the non-specific binding of (11-55)NCp7 all over the cTAR induced an increase of the quantum yield of the 2-Ap at all the positions.

Figure 5: Proportion of the 2-Ap stacked conformations as a function of (11-55)NCp7.

- (A) A strong unstacking effect of (11-55)NCp7 was observed for positions AP9 (black circles) and AP49 (red squares) with a decrease of a factor of about two in the amplitudes associated to the stacked conformers of the 2AP in cTAR.
- (B) When the 2AP was located in the different ds-segments of cTAR probed by AP17 (black double triangles), AP21 (green triangles) AP45 (purple crosses), AP53 (blue squares) and AP55 (red circle), an overall (11-55)NCp7 concentration-dependent decrease of the stacked conformations

was observed, showing that (11-55)NCp7 is able to bind unspecifically all over the cTAR sequence.

(C) No change in the amplitude of the stacked conformations is observed when the 2AP is inserted in a ss-part of cTAR as seen in AP35 (red squares) or AP28 (red triangles).

Figure 6: Effect of (11-55)NCp7 on the cTAR AP9 fluorescence.

(A) Emission fluorescence spectra of cTAR AP9 in the absence (black line) or in the presence of increasing NCp7/cTAR ratios (one (red), four (green), seven (blue), ten (light-blue) and fifteen (purple)).

(B) Representative anisotropy decay of cTAR AP9 for increasing NCp7/cTAR ratios (color codes are the same as in (A))

(C) Evolution of the cone semi-angle in which the 2AP rotates as a function of (11-55)NCp7 concentrations. The binding of (11-55)NCp7 freezes the local mobility of cTAR, as traduced by the reduction of the cone semi angle.

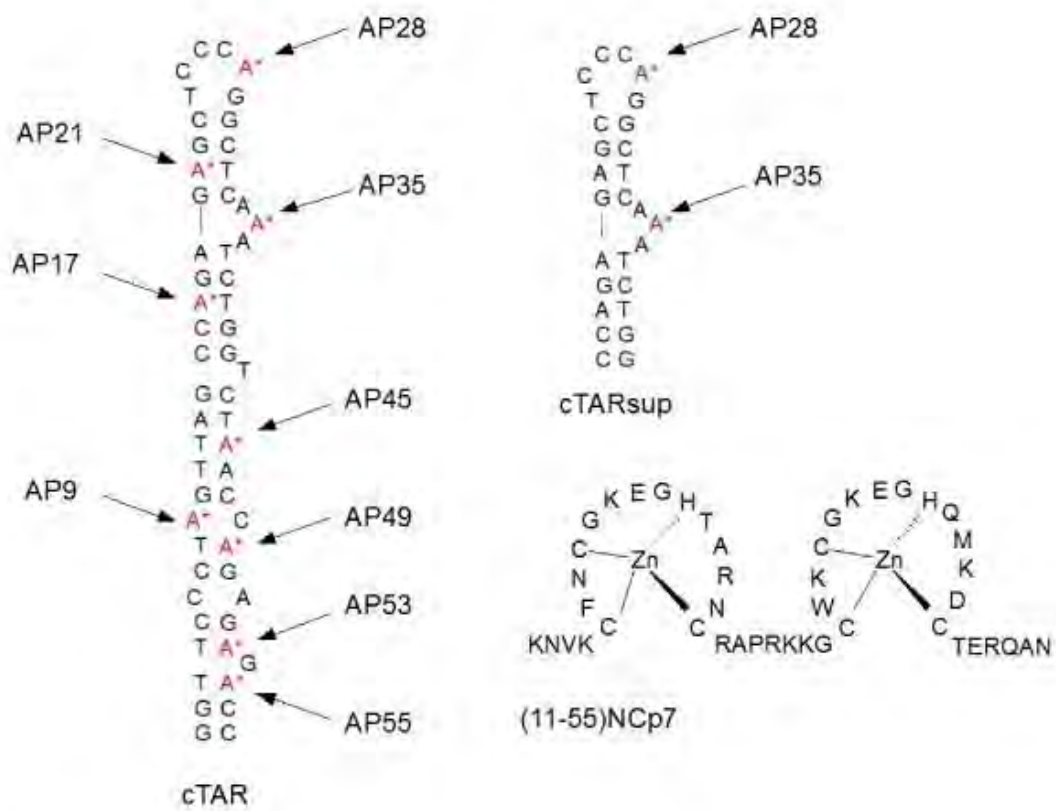


Figure 1.

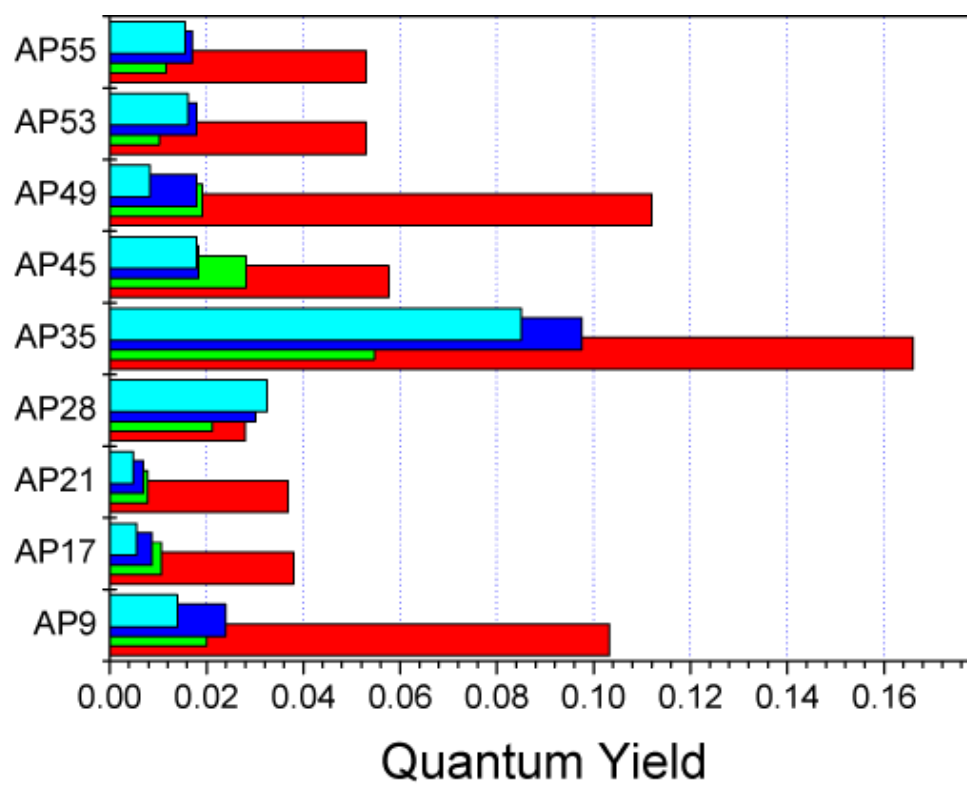


Figure 2.

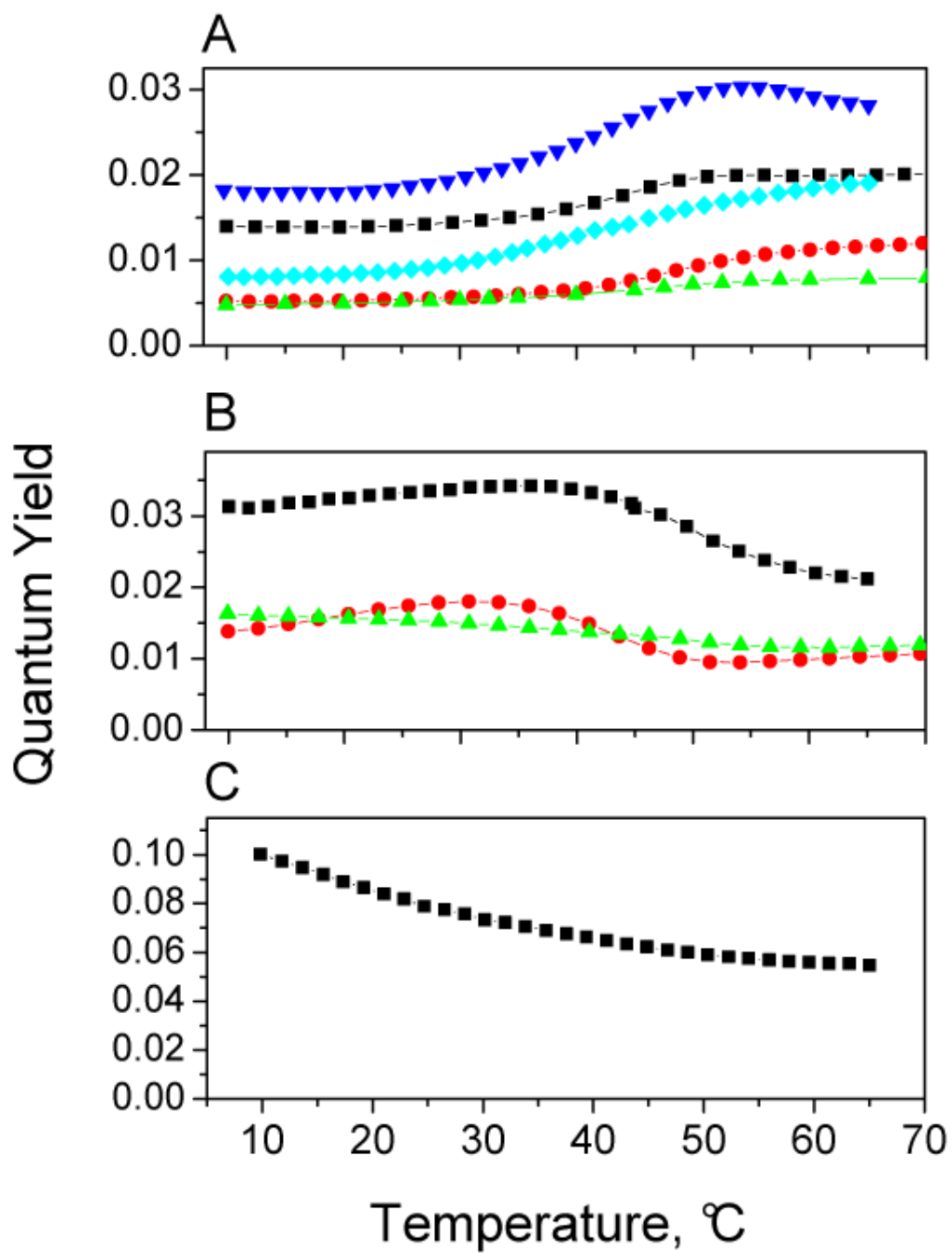


Figure 3.

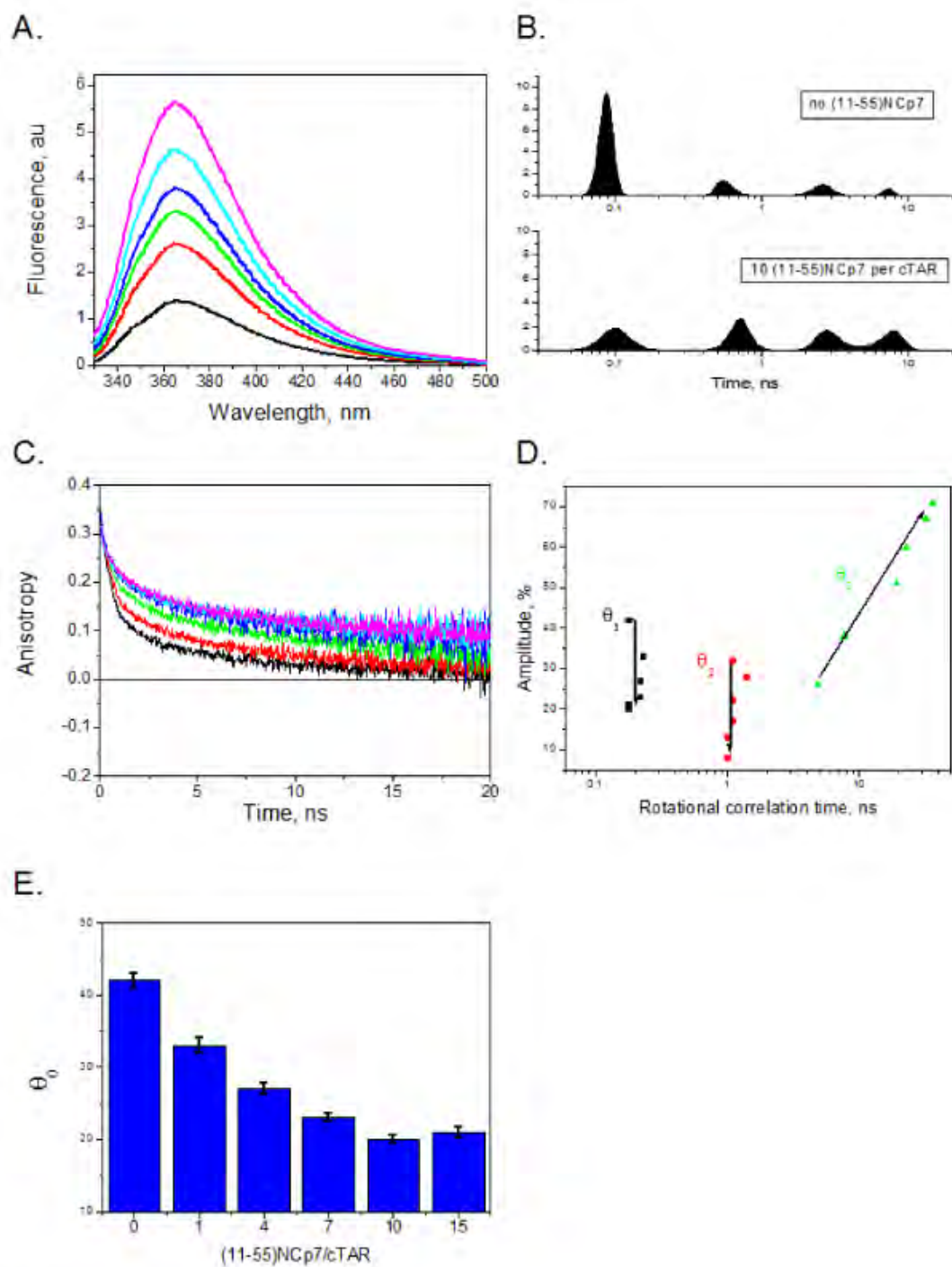


Figure 4.

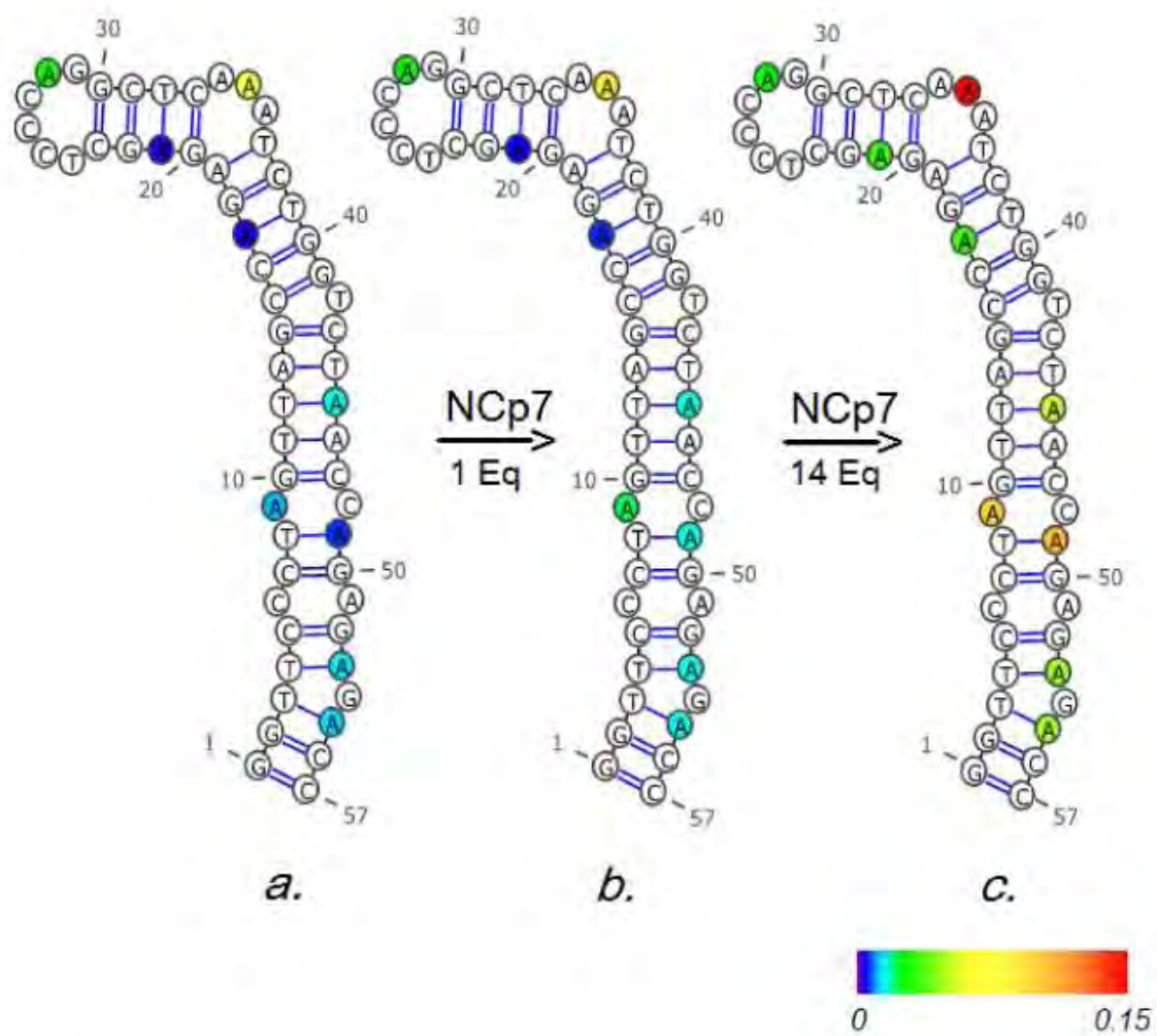


Figure 5.

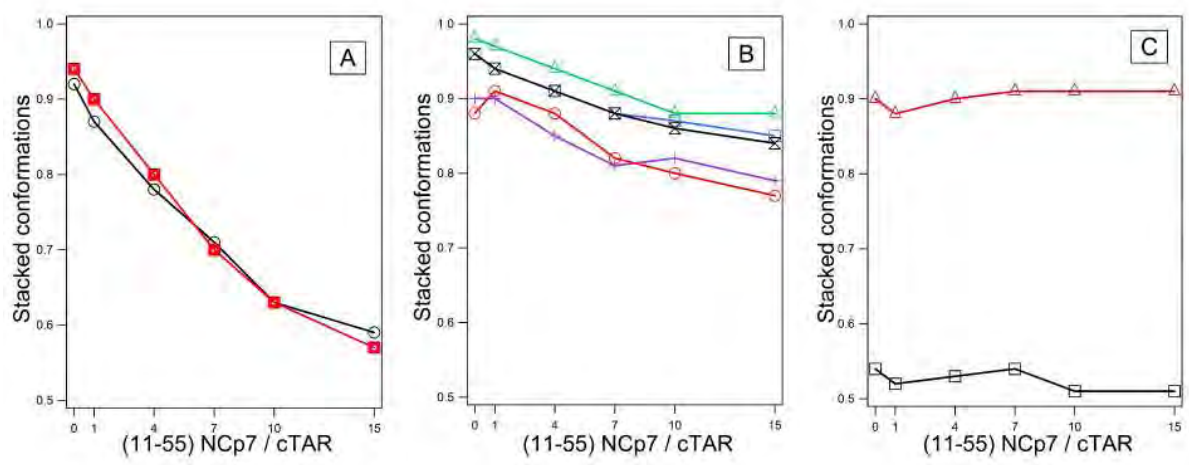


Figure 6.

Supplementary

Titration experiments:

Reverse titrations were performed by adding increasing concentrations of cTAR Lai to a fixed amount of (11-55)NCp7 (Figure 3.). The binding was monitored via the Trp37 fluorescence quenching induced by its stacking within oligonucleotides bases [31, 34, 36, 64]. Binding experiments were performed in a buffer containing 25mM of TRIS, 30mM NaCl, 0.2mM of MgCl₂, pH 7.5. Fluorescence intensities were corrected for dilution, buffer fluorescence and screening effect due to the absorption of nucleic acids. Affinity constants were determined from direct fitting of the experimental data corrected for dilution and screening effect according to the rewritten Scatchard equation

$$I = I_i + \frac{(I_p - I_i)}{n} \times \left(\frac{(1 + K_\mu(1 + nx)[NC]_{tot}) - \sqrt{(1 + K_\mu(1 + nx)[NC]_{tot})^2 - 4nxK_\mu^2[NC]_{tot}^2}}{2K_\mu[NC]_{tot}} \right) \quad \text{Eq. (S1)}$$

where I is the fluorescence intensity at the given cTAR concentration, I_i and I_p are the fluorescence emission intensities at 320nm (where the 2Ap does not fluoresce) in the absence of oligonucleotide and at the plateau value respectively. $[NC]_{tot}$ is the total concentration of NCp7, K_μ is the apparent dissociation constant, x the ratio [cTAR]/[NCp7]. The number of binding sites n is ascertained from the intercept of the initial slope of the titration curve with the asymptote slope of the final plateau. The parameters are recovered from non-linear fits of Eq (S1) to experimental datasets with a home-written Igor Pro 6[®] (WaveMetrics) fitting procedure.

The number of peptides bound per cTAR was 10.7+/-0.3 for (11-55)NCp7 (Figure S1 inset) giving an occluded binding site of about 5 nucleotides per peptide, in agreement with the previously reported binding size of HIV-1 NC on nucleic acids. [27, 30, 31]. No significant differences were induced by the substitution of an adenine by a 2-AP in the cTAR sequence (for any 2Ap insertion), indicating that 2-AP does not affect NCp7 binding (Figure S1). The apparent binding constant K_μ can be retrieved from the experimental data, using a model of n identical and independent binding sites for NCp7, according to Equation (1) [55]. Although it is obvious that the latter assumption is approximate (NCp7 binding depends on the oligonucleotide sequence and strandedness [27, 31, 38], this approach yield a K_μ of $1.4 \times 10^7 \text{M}^{-1}$, which is in good agreement with previous works on cTAR Mal strain [21, 55, 65].

Figure S1 legend: Titration curve of (11-55)NCp7 affinity for cTAR Lai.

Substituting the natural A by a 2-Ap did not affected the binding of (11-55) NCp7, as seen from the

perfect superimposition of the titration curves of the unlabelled or labelled cTAR (only AP45 and AP28 are represented together with the unlabelled cTAR for clarity). The red line correspond to the fit of the data with equation S1.

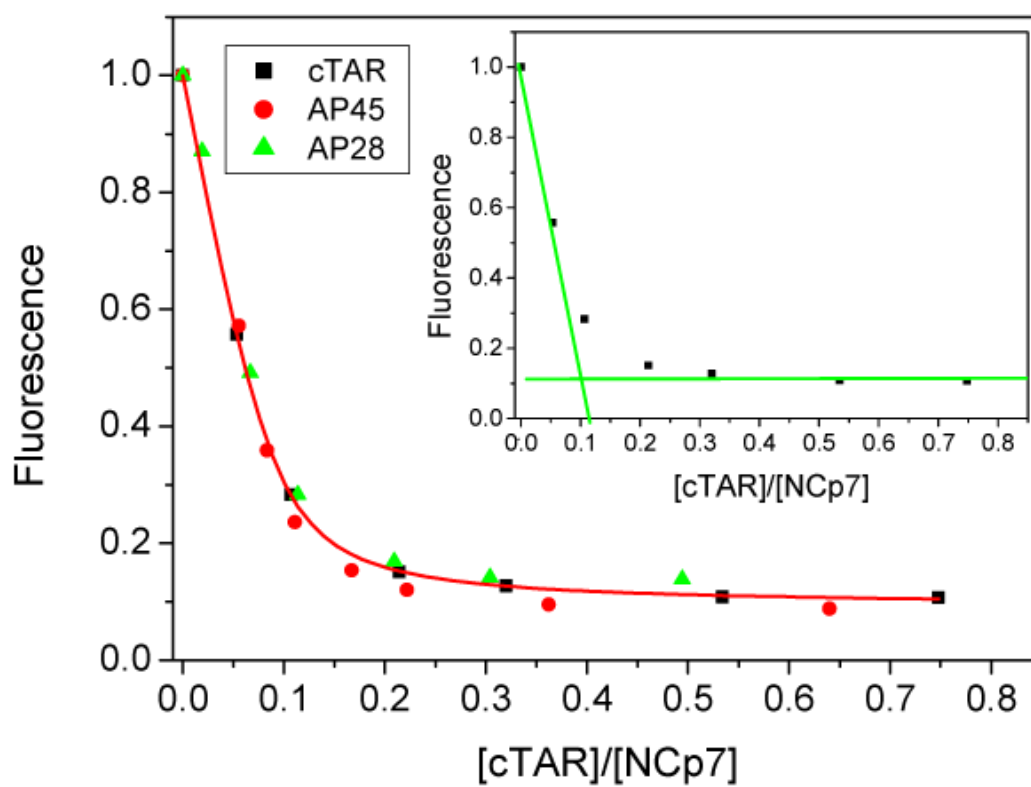


Figure S1.

5.7 Analyse structurale de cTAR

La séquence en tige boucle de cTAR (souche Mal) a été sélectivement modifiée avec une 2-Ap insérée respectivement dans chacun des segments double-brin (Figure 5.6.), afin d'obtenir des informations locales sur la structure et la dynamique de ces domaines.

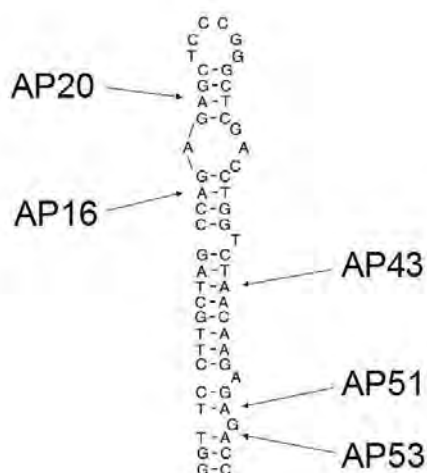


FIGURE 5.6 – Séquence de cTAR Mal et position d'insertion des 2-Ap

La stabilité thermodynamique globale de cette séquence a déjà été déterminée [268]. Afin d'obtenir des informations sur la stabilité des différentes zones double-brin, des expériences de fusion induite par la température ont été réalisées. La cinétique d'ouverture de la tige est suivie localement par le changement de l'intensité de fluorescence de la 2-Ap. Pour chaque température, le signal brut d'émission de fluorescence à 370 nm est divisé par le signal d'une solution de référence de 2Ap-riboside (correction de la perte d'intensité de l'émission de fluorescence en fonction de la température) afin de reconstituer une courbe de fusion.

Pour les 2-Ap en position 16, 20 et 46, l'émission de fluorescence est fortement éteinte (*quen- chée*) en raison de la structure double-brin qui impose un empilement (*stacking*) important de la sonde avec ses bases voisines. Lorsque la température augmente, la fluorescence augmente fortement (Figure 5.7b), ce qui traduit que le *stacking* de la 2-Ap au sein d'une séquence simple brin devient moins important que dans une séquence double-brin (Figure

5.7 Analyse structurale de cTAR

5.7a.)

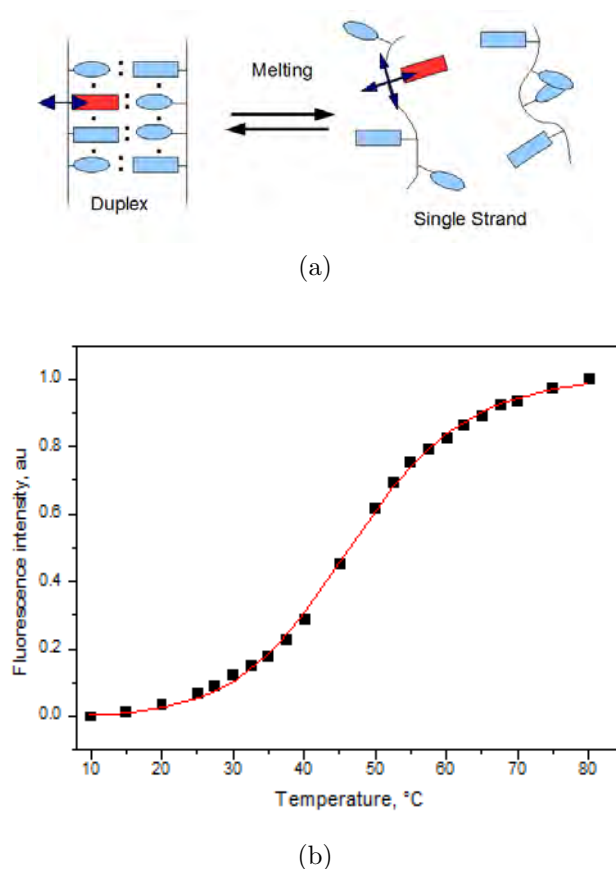


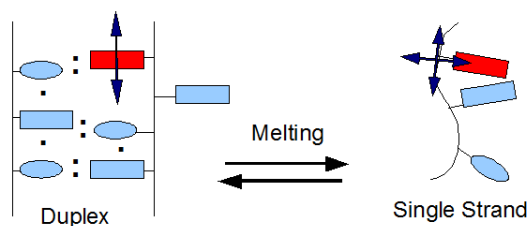
FIGURE 5.7 – (a) Schéma de la fusion d'un ADN double brin
(b) Courbe de fusion de cTAR-AP16 suivie par l'émission de fluorescence de la 2Ap

Pour les 2-Ap en position 51 et 53, proche de la guanine *bulgée* terminale, un phénomène contraire est observé avec une diminution de la fluorescence au fur et à mesure que la température augmente. La 2-Ap passe d'un état plus quenché à faible température vers un état moins quenché lorsque la température augmente. Ceci suggère fortement que dans l'état double-brin, la 2Ap est protégée du *stacking* avec G52, confirmant que la guanine *bulgée* est extra-hélicale (Figure 5.8a) [402].

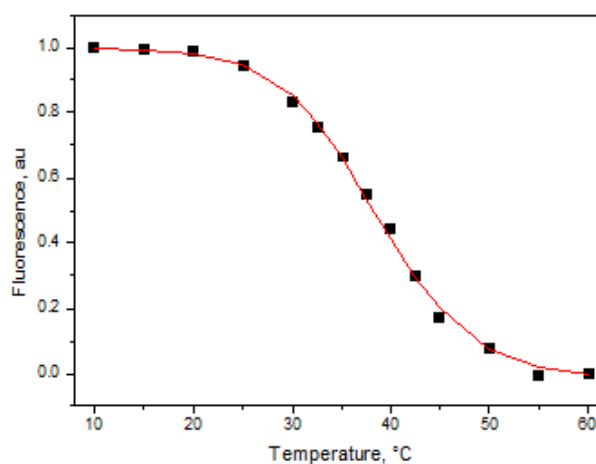
Les données thermodynamiques des courbes de fusion sont extraites à l'aide d'un modèle de Van't Hoff, impliquant une transition ordre-désordre au sein d'un système à deux états [403]. Ce modèle, appliqué à une structure en tige-boucle (système monomoléculaire), prend la forme suivante :

$$\ln K(T) = \frac{\Delta H^0}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right) \quad (5.1)$$

5.7 Analyse structurale de cTAR



(a)



(b)

FIGURE 5.8 – (a) Schéma de la fusion de la partie basse de cTAR.

(b) Courbe de fusion de cTAR-AP51 suivie par l'émission de fluorescence de la 2Ap

permettant de calculer ΔG^0 et ΔS^0 selon

$$\Delta G^0 = \Delta H^0 \left(1 - \frac{T}{T_m} \right)$$

$$\Delta S^0 = \frac{\Delta H^0}{T_m}$$

En utilisant la relation entre $K(T)$, la constante d'équilibre entre les deux formes, et les signaux de fluorescence associés à chacune de ces formes, on obtient alors

$$F(T) = \frac{F_{ini} + F_{fin} \times \exp\left(-\frac{\Delta H^0}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_m}\right)\right)}{1 + \exp\left(-\frac{\Delta H^0}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_m}\right)\right)} \quad (5.2)$$

où R est la constante des gaz parfaits, T_m la température de fusion et F_{ini} et F_{fin} , les

5.7 Analyse structurale de cTAR

signaux de fluorescence correspondant à l'état double brin et simple brin, respectivement. Ce modèle permet de décrire fidèlement les données (Figure 5.7b. et 5.8b.). Les paramètres thermodynamiques correspondants sont présentés dans le tableau 5.1.

| | ΔH^0 ($kCal.mol^{-1}$) | ΔS^0 (eu) | $\Delta G_{37^\circ C}^0$ ($kCal.mol^{-1}$) | T_m ($^\circ C$) |
|-----------------------|-------------------------------------|--------------------------|--|-------------------------|
| cTAR-AP ₁₆ | -25.8 ± 1 | -80 ± 2 | -0.74 ± 0.1 | 46.6 ± 0.4 |
| cTAR-AP ₂₀ | -28.6 ± 1 | -90 ± 2 | -0.55 ± 0.1 | 43.2 ± 0.6 |
| cTAR-AP ₄₃ | -47.7 ± 1 | -151 ± 2 | -0.8 ± 0.05 | 43.4 ± 0.4 |
| cTAR-AP ₅₁ | -45.5 ± 1 | -146 ± 2 | -0.37 ± 0.05 | 39.6 ± 0.5 |
| cTAR-AP ₅₃ | -25.3 ± 1 | -82 ± 2 | 0.08 ± 0.01 | 36.1 ± 0.4 |

TABLE 5.1 – Paramètres thermodynamiques de la stabilité des différentes zones double-brin de cTAR

Les stabilités (ΔG^0) des segments double-brins de cTAR apparaissent hétérogènes (Tableau 5.1 et Figure 5.9). Le segment central, sondé par AP16 et AP46, apparaît comme le plus stable. Le segment supérieur (AP20), est probablement destabilisé par la boucle apicale et la boucle interne. Les segments terminaux du bas de la tige sont les moins stables et peuvent fondre indépendamment de la zone centrale de la tige.

Utilisation du transfert d'énergie entre bases

Le spectre d'excitation de fluorescence de la séquence de cTAR-AP43 présente une bande aux alentours de 260 – 270 nm. À ces longueurs d'onde, l'excitation directe de la 2-Ap est peu probable alors que les bases naturelles absorbent. Cette excitation de la séquence cTAR-AP43 traduit donc l'existence d'un transfert d'énergie entre les bases naturelles et la 2-Ap [404]. Le transfert d'énergie de base à base au sein d'une séquence d'ADN résulte notamment de transfert de charges [405]. Le transfert d'énergie des bases naturelles vers la 2-Ap se limite à une distance de quelques bases. L'Adénine se montre un ordre de grandeur plus efficace que les autres bases [406]. En outre, le transfert d'énergie est généralement plus efficace lorsque la température est inférieure à $\sim 15^\circ C$ et la bande d'excitation correspondant au transfert d'énergie est fortement dépendante du stacking des bases. Pour toutes ces raisons, l'étude du transfert d'énergie de la 2-Ap en position AP46 devrait permettre de sonder la structure du motif 42-TAA-44 au sein de cTAR, et donc de caractériser la conformation de

5.7 Analyse structurale de cTAR

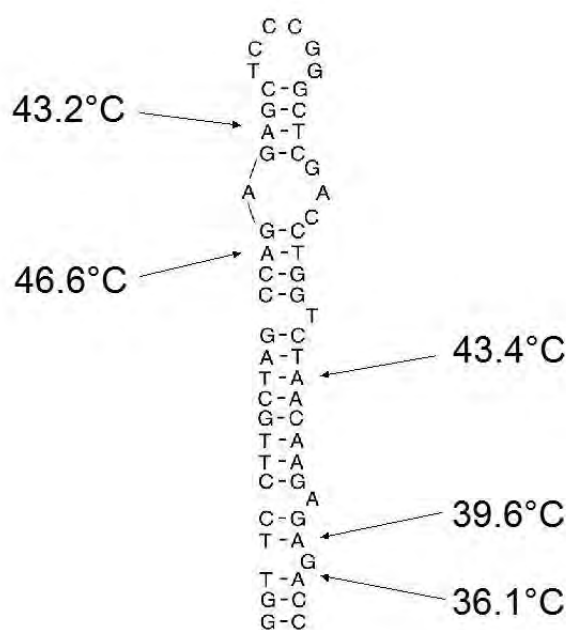


FIGURE 5.9 – Stabilités locales de la séquence de cTAR Mal.

la base A44 qui est hybridée de manière imparfaite avec C10.

Afin de caractériser le stacking de la 2-Ap en position 43, nous avons évalué l'influence de la température sur le transfert d'énergie. L'importance du transfert d'énergie est minorée par l'effet d'écran causé par les bases qui absorbent autour de 260 nm. Pour calculer le transfert d'énergie, les mesures d'intensité de fluorescence sont corrigées par un facteur CF [407] selon :

$$CF = 2.303 \times \frac{A(\lambda_{ex})}{1 - 10^{-A(\lambda_{ex})}}$$

Il apparait clairement sur la figure 5.10a que le transfert d'énergie diminue avec l'augmentation de la température. D'autre part, le maximum du pic d'excitation de la 2-Ap se décale vers le bleu avec l'augmentation de la température. Puisqu'un *red-shift* de la sonde incorporée dans un ADN traduit une faible exposition de l'aminopurine au solvant, les variations du *shift* traduisent la variation de l'exposition de la sonde au solvant. Si l'on représente la position (en nm) du maximum du pic d'excitation de la 2-Ap en fonction de la température, on obtient une relation linéaire biphasique dont l'intersection des deux

5.7 Analyse structurale de cTAR

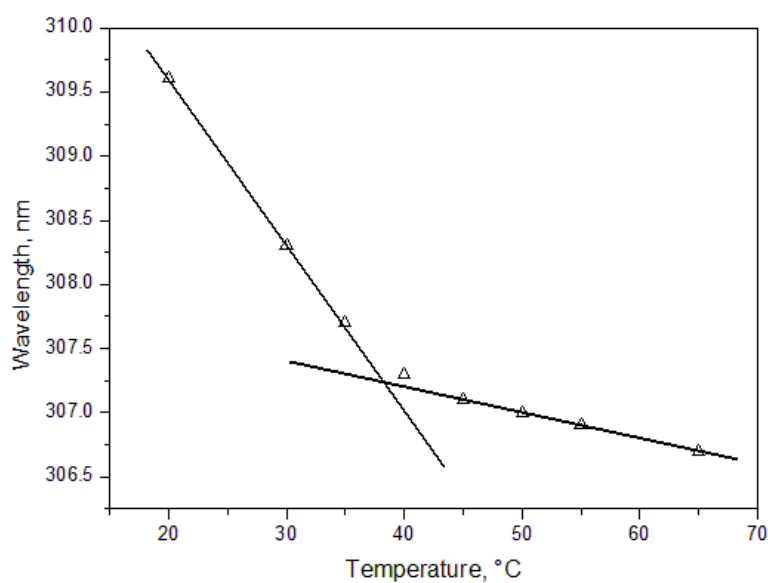
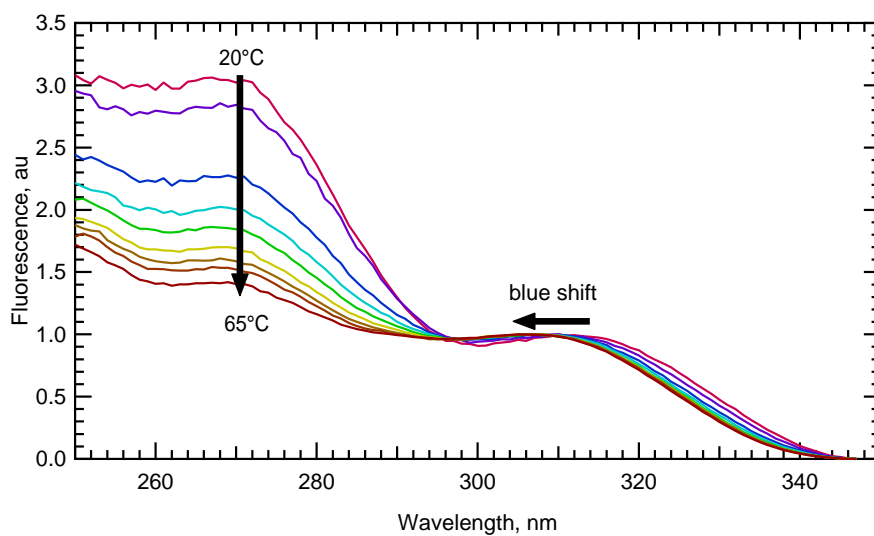


FIGURE 5.10 – (a) Spectre d'excitation de cTAR AP43 en fonction de la température. (b) Longueur d'onde du maximum d'excitation de la 2-Ap de cTAR-AP43 en fonction de la température.

droites correspond à une température proche de la température de fusion de la séquence (Figure 5.10b).

Ces observations suggèrent qu'en deçà de la température de fusion de la tige-boucle de cTAR, la base A47 est principalement orientée vers l'intérieur de la tige, en accord avec le

5.7 Analyse structurale de cTAR

fait que la 2-Ap est capable de s'apparier avec une cytosine [408].

La position AP35 dans la séquence de cTAR Lai est également intéressante à sonder par cette approche. En effet, AP35, positionnée au coeur de la boucle interne, est flanquée de deux adénines. De plus, les mesures d'anisotropie à l'état stationnaire suggèrent que la 2-Ap en cette position est fortement contrainte (Tableau 5.2).

| | Rss | Bornes IC _{95%} | | r_0 | $\langle\tau\rangle$ (ns) | $\langle\phi\rangle$ (ns) | Bornes IC _{95%} | |
|------------|-------|--------------------------|-------|-------|---------------------------|---------------------------|--------------------------|------|
| free 2AP | 0.003 | 0.002 | 0.003 | 0.350 | 10.200 | 0.08 | 0.07 | 0.10 |
| cTAR AP9 | 0.089 | 0.083 | 0.095 | 0.320 | 0.690 | 0.26 | 0.24 | 0.29 |
| cTAR AP28 | 0.061 | 0.059 | 0.063 | 0.320 | 1.560 | 0.37 | 0.36 | 0.38 |
| cTAR AP35 | 0.120 | 0.117 | 0.123 | 0.300 | 1.600 | 1.06 | 1.02 | 1.11 |
| cTAR AP43 | 0.118 | 0.116 | 0.120 | 0.340 | 0.610 | 0.32 | 0.32 | 0.33 |
| cTAR AP49 | 0.110 | 0.107 | 0.113 | 0.340 | 0.800 | 0.38 | 0.37 | 0.40 |
| cTAR AP28c | 0.046 | 0.044 | 0.048 | 0.340 | 2.190 | 0.34 | 0.32 | 0.36 |
| cTAR AP35c | 0.110 | 0.108 | 0.111 | 0.340 | 1.890 | 0.90 | 0.88 | 0.92 |

TABLE 5.2 – Paramètres d'anisotropie à l'état stationnaire de cTAR Lai ou de sa moitié supérieure

Rss : anisotropie à l'état stationnaire, r_0 : anisotropie fondamentale, $\langle\tau\rangle$: temps de vie moyen, $\langle\phi\rangle$: temps moyen de corrélation de rotation, calculé selon l'équation de Perrin $\langle\phi\rangle = \frac{\langle\tau\rangle}{r_0 - 1}$

Comme pour cTAR Mal AP46, on observe au fur et à mesure que la température augmente :

- un *blue shift* pour atteindre une longueur d'onde au maximum du pic aux alentours de 306 nm (très proche des 305 nm de la sonde libre dans le même tampon)
- une diminution de l'intensité de la bande vers 260-270nm correspondant au transfert d'énergie des bases vers la 2-Ap

L'efficacité du transfert d'énergie est calculée selon :

$$E(t) = I(\lambda_D) \frac{1 - 10^{-A(\lambda_A)}}{1 - 10^{-A(\lambda_D)}}$$

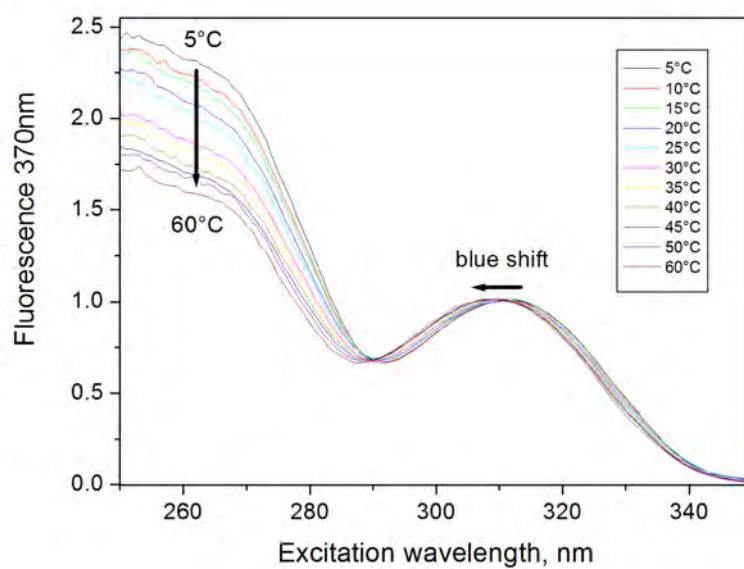
où $I(\lambda_D)$ est normalisée à 1.

En utilisant l'absorbance de l'oligonucléotide aux longueurs d'ondes λ_A et λ_D aux températures T, on obtient des valeurs de transfert relativement faible (maximum 1.55% pour les plus basses températures). Cependant, si l'on considère que seules les 2 bases voisines

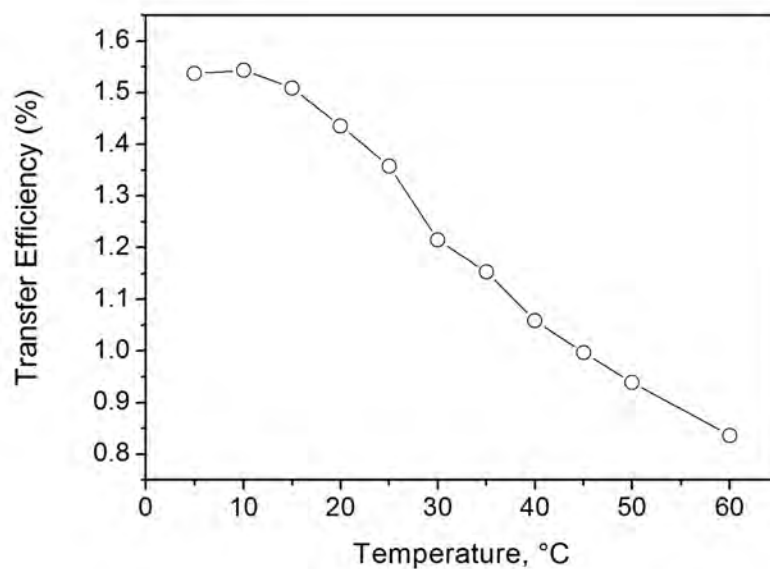
5.7 Analyse structurale de cTAR

de la 2AP sont capables de transférer leur énergie. Le transfert maximum que l'on peut observer est d'environ ($2/57 \sim 3.5\%$). Dans ce cas, l'efficacité de transfert est d'environ $1.55 \times 57/3 \sim 30\%$ ce qui correspond à un taux de transfert de $k_t \sim 3 \times 10^{10} s^{-1}$ (sur la base d'un temps de vie moyen d'une base naturelle de $\sim 10 ps$). Cet ordre de grandeur de vitesse devrait être accessible par spectroscopie ultra-rapide (femtoseconde).

5.7 Analyse structurale de cTAR



(a)



(b)

FIGURE 5.11 – (a) Spectre d'excitation de cTAR AP35 en fonction de la température
(b) Efficacité du transfert d'énergie en fonction de la température

5.8 Action de dTAR sur la stabilité de cTAR

Dans le mécanisme d'hybridation de cTAR avec dTAR en absence de NCp7, cTAR s'hybriderait dans un premier temps avec dTAR dans un complexe boucle-boucle avant de former un duplex étendu cTAR/dTAR [398, 282]. Afin de vérifier cette hypothèse, nous avons utilisé une séquence en tige-boucle mutée, le cTAR_{1,2}wAP₃₃, un mutant de cTAR dont la tige est stabilisée (les bases complémentaires des deux *bulges* terminaux ont été ajoutées à la séquence du brin complémentaire [268]) et portant une boucle de dTAR. Ce mutant peut ainsi former un complexe boucle-boucle avec cTAR mais ne peut pas donner de duplex étendu. Ce mutant peut d'hybrider de manière imparfaite à dTAR (Figure 5.12).

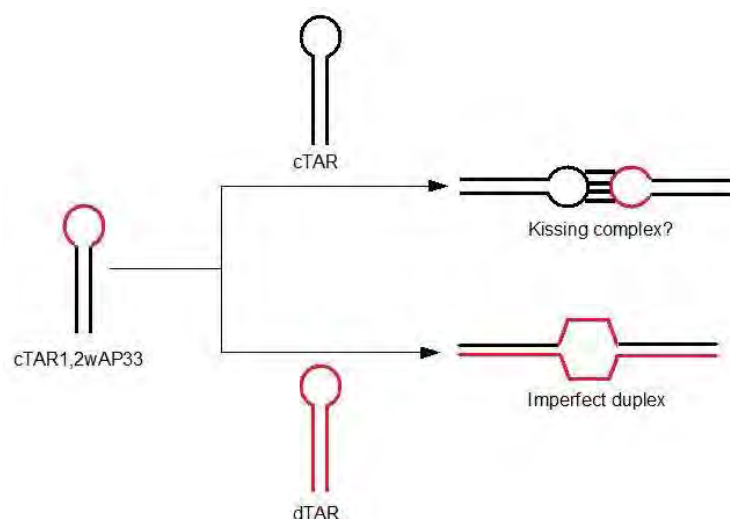


FIGURE 5.12

La formation d'un complexe boucle-boucle stable devrait se traduire par une transition supplémentaire dans les courbes de fusion induites par la température.

Les courbes de fusion de cTAR_{1,2}wAP₃₃ suivies par UV ou par la fluorescence de la 2Ap sélectivement insérée dans la boucle donnent les mêmes paramètres thermodynamiques, très similaires de ceux publiés par Beltz et al. [268], montrant la validité de l'approche (tableau 5.3). L'ajout de cTAR n'a pas d'influence sur la fusion de cTAR_{1,2}wAP₃₃, laissant supposer que si l'interaction boucle-boucle existe elle doit être peu stable. Au contraire de cTAR,

5.8 Action de dTAR sur la stabilité de cTAR

| | ΔH^0 (<i>kCal.mol</i> ⁻¹) | ΔS^0 (<i>eu</i>) | $\Delta G_{37^\circ C}^0$ (<i>kCal.mol</i> ⁻¹) | T_m (<i>°C</i>) |
|---|---|-------------------------------|--|------------------------|
| cTAR _{1,2} ^a | -88 ± 2 | -263 ± 7 | -6.5 ± 0.2 | 63 ± 1 |
| cTAR _{1,2} wAP ₃₃ par UV | -96 ± 2 | -285 ± 8 | -8.3 ± 1 | 66 ± 1 |
| cTAR _{1,2} wAP ₃₃ | -90 ± 8 | -264 ± 10 | -7.9 ± 0.8 | 67 ± 2 |
| cTAR _{1,2} wAP ₃₃ + cTAR | -91 ± 9 | -267 ± 12 | -8 ± 1 | 67 ± 2 |
| cTAR _{1,2} wAP ₃₃ + dTAR | -48 ± 5 | -147 ± 8 | -2.65 ± 0.5 | 55 ± 1 |
| cTAR _{1,2} wAP ₃₃ + dTAR par UV | -103 ± 7 | -300 ± 13 | -10.3 ± 1 | 71 ± 1 |
| cTAR + dTAR par UV | -95 ± 7 | -270 ± 17 | -11.5 ± 1 | 80 ± 1 |
| dTAR-AP ₃₃ | -54 ± 6 | -162 ± 10 | -3.2 ± 0.5 | 57 ± 2 |

TABLE 5.3 – Paramètres thermodynamiques de la stabilité de cTAR_{1,2}, cTAR_{1,2}wAP₃₃ et dTAR par UV ou par fluorescence de la 2-Ap

^a de Beltz & al. 2003 [268]

l'ajout de dTAR induit une diminution de près de 12°C de la température de fusion de cTAR_{1,2}wAP₃₃ (tableau 5.3 et figure 5.13), ce qui suggère que la présence de la séquence complémentaire induit une forte destabilisation de la séquence cTAR_{1,2}wAP₃₃. Cette diminution de la température de fusion n'est pas observée lorsque l'on suit la fusion de la séquence cTAR_{1,2}wAP₃₃ en présence de dTAR par UV puisque dans ce cas, on observe uniquement la fusion du duplex étendu (tableau 5.3 et figure 5.14), dont la formation est rendue possible par l'élévation de température. Ces observations permettent donc de mettre en évidence que la 2-Ap de suivre les fusions locales des acides nucléiques. D'autre part, la présence de dTAR induit une diminution de la stabilité de cTAR. Cette destabilisation induite par la présence d'une séquence complémentaire pourrait expliquer le saut de fluorescence observé au cours de la phase initiale de l'hybridation de 5'FAM-cTAR-TMR-3' avec dTAR [280].

5.8 Action de dTAR sur la stabilité de cTAR

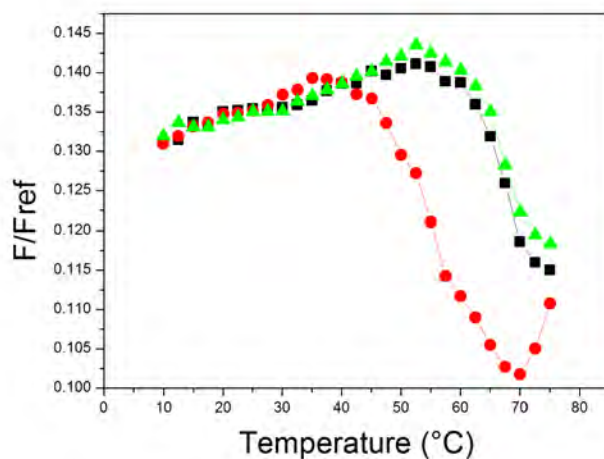


FIGURE 5.13 – Courbe de fusion obtenue en suivant la fluorescence de la 2-Ap en fonction de la température
cTAR_{1,2}wAP₃₃ seul (carrés noirs), cTAR_{1,2}wAP₃₃ + dTAR (ronds rouges) et cTAR_{1,2}wAP₃₃ + cTAR (triangles verts)

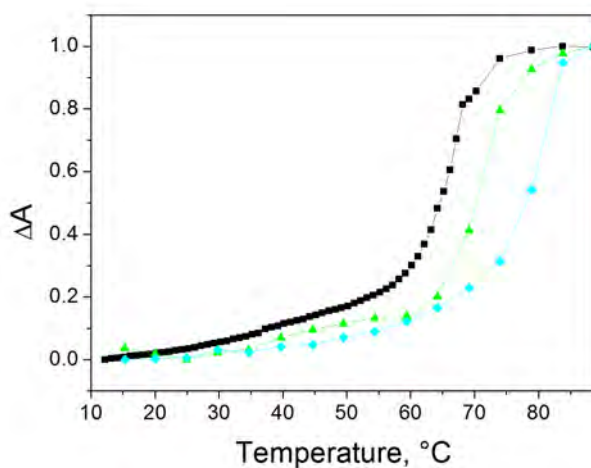


FIGURE 5.14 – Courbe de fusion obtenue en suivant l'absorbance à 260 nm en fonction de la température
dTAR seul (carrés noirs), cTAR_{1,2}wAP₃₃ seul (triangles verts) et cTAR_{1,2}wAP₃₃ + dTAR (ronds cyans)

Spectroscopie de fluorescence à l'échelle de la molécule unique

6.1 Matériels et Méthodes

6.1.1 Plateformes d'imagerie

Microscopie de fluorescence à deux photons (TPE-SFM)

Le TPE-SFM (pour *Two Photon Excitation Scanning Fluorescence Microscope*) a déjà été décrit [267, 382]. Le faisceau d'excitation à deux photons est généré par un laser Titane-Saphir fonctionnant en mode *locké* (Pulse $\sim 100 fs$; Repetition rate 80MHz), pompé par un laser continu Nd:YO₄ (Millenia X Spectra-Physics) émettant à 532 nm. La longueur d'onde d'excitation est ajustée en réglant le monochromateur à 415 nm. La puissance laser est maintenue à $\sim 20 mW$. Le faisceau d'excitation est élargi à l'aide d'un *beam expander* afin de couvrir la totalité de la pupille arrière de l'objectif à immersion ($\times 60$, N.A.=1.2 *Olympus*). Deux miroirs galvanométriques sont ajoutés sur la voie d'excitation afin de pouvoir balayer l'échantillon. La vitesse de balayage est de $4 \mu s$ par pixel, suffisant pour récolter un nombre suffisant de photons et limiter le photoblanchiment. Le volume d'excitation résultant de cette source d'excitation a un volume d'environ 0.3fL et est grossièrement gaussien ($0.3 \mu m$ de rayon; $\sim 1.1 \mu m$ sur l'axe des z). L'émission de fluorescence provenant de l'échantillon est collectée par ce même objectif, filtrée (passe-bande 585/40nm) et redirigé vers une APD selon un mode descanné. Les images sont reconstituées à l'aide d'une procédure *Labview* maison.

Microscopie en champ large/TIRF

Cette plateforme a été construite dans le but de réaliser des expériences de FRET à l'échelle de la molécule unique. Deux sources laser sont disponibles : un laser solide continu 532 nm (100 mW, Cobolt) qui est utilisé pour exciter sélectivement le donneur de fluorescence, et un laser solide continu 635 nm (60 mW, Spectra-Physics) permettant l'excitation de l'accepteur. La lumière d'excitation est sélectionnée par des obturateurs électromécaniques contrôlés par ordinateur. Un télescope ($f'_1 = 3cm$ et $f'_2 = 15cm$) est utilisé pour élargir le faisceau d'excitation. Ce faisceau est ensuite dirigé vers un objectif à immersion à huile ($\times 60$ N.A.=1.4 ou $\times 100$ N.A.=1.4) *Olympus*) à l'aide d'un miroir ajustable

6.1 Matériels et Méthodes

permettant de passer facilement du mode d'illumination en champ large au mode TIRF. L'émission de fluorescence est collectée par le même objectif, puis est restreinte spatialement à l'aide d'une fente, et séparée, à l'aide d'un miroir dichroïque (650 nm), en deux voies de détection sur lesquelles elle est filtrée (passe-bande). Les deux faisceaux de fluorescence sont finalement imagés sur la même caméra CCD amplifiée (EM-CCD Hamamatsu).

6.1.2 Immobilisation des molécules

L'observation de molécules à l'échelle de la molécule unique à l'aide d'un microscope fonctionnant en mode TIRF nécessite d'immobiliser les molécules que l'on cherche à observer. Parmi les méthodes d'immobilisations, l'utilisation de nano-conteneurs immobilisés enfermant les molécules d'intérêt permet de s'affranchir des problèmes d'interaction avec la surface. Nous avons fait le choix de vésicules lipidiques. Un couple de molécules donneur-accepteur est donc enfermé au sein de vésicules lipidiques qui sont fixées sur la surface de la lamelle *via* un couplage biotine/streptavidine (Figure 6.1).

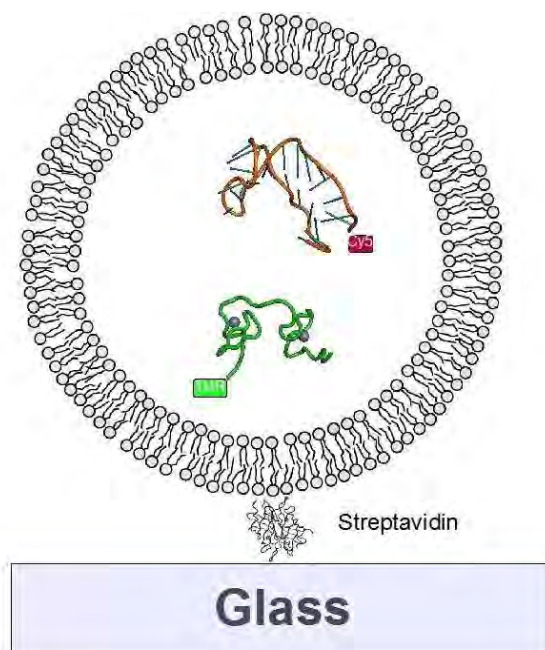


FIGURE 6.1 – Immobilisation des molécules à l'aide de vésicules lipidiques

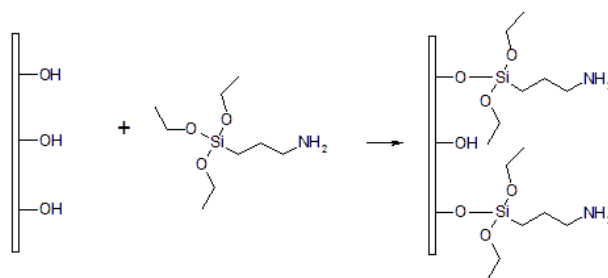
6.1 Matériels et Méthodes

Préparation des surfaces d'ancrage pour les vésicules

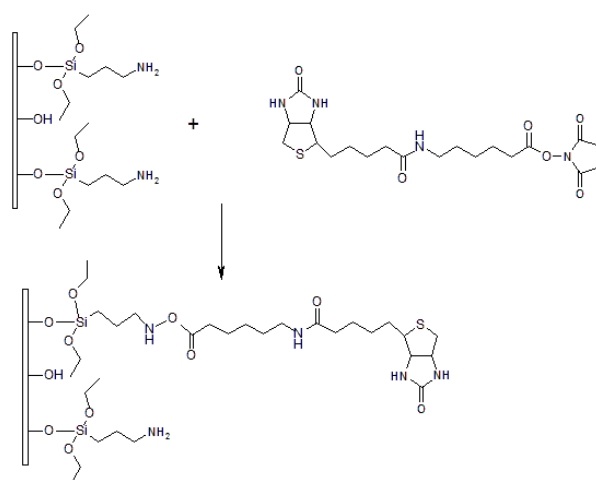
Les lamelles pour microscopie sont d'abord nettoyées à l'aide d'une solution dite *piranha like*, composée d'acide peroxodisulfurique fraîchement préparée, avant d'être abondamment rincées à l'eau Milli-Q et séchées. Toute particule fluorescente résiduelle est ensuite photoblanchie par un traitement d'une heure dans un *UV-Cleaner*.

La surface d'ancrage est ensuite construite couche par couche, à partir du verre nu de la lamelle, selon les trois étapes suivantes :

1. aminosilanisation des hydroxyles de la surface du verre par le 3 Aminopropyl-triéthoxysilane



2. biotinylation des groupements $-NH_2$ de surface par le (+)-Biotin N-hydroxysuccinimide ester (ou biotin amidocaproic acid 3-sulfo-N-HydroxySuccinimide ester)



3. couplage de la streptavidine aux résidus biotine

6.1 Matériels et Méthodes

Cette préparation du support permet l'immobilisation de vésicules lipidiques intactes, immobiles à l'échelle du centième de micromètre, et pouvant être maintenues sans que les vésicules ne fusionnent pendant plusieurs heures [409].

À noter que pour les expériences où les molécules sont directement immobilisées sur la surface, la phase 2 est substituée par une phase de pegylation à l'aide d'un mélange NHS-PEG et NHS-PEG-Biotine.

Préparation des LUV (Large Unilamellar Vesicles)

De la 18 :1 Biotinyl-cap- PE est mélangée dans du CHCl_3 à de l'*Egg-Yolk-PC* selon un rapport molaire de 1 :100 dans un ballon en verre. Le CHCl_3 est ensuite évaporé sous vide (RotaVap) pour obtenir un film lipidique, qui est maintenu sous vide à 50°C pendant au moins 30 min afin d'éliminer le CHCl_3 résiduel. Ce film est ensuite hydraté avec un tampon 10mM TRIS, pH 7.5, 30mM NaCl (sauf mention contraire), 0.2mM MgCl_2 contenant les molécules fluorescentes, ce qui induit le gonflement du film lipidique. Cette préparation est vortexée vigoureusement pour obtenir une solution de 2 mg/mL (en lipide) de MLV (*multilamellar vesicles*). Finalement, une population de LUV (*large unilamellar vesicles*) de taille homogène est obtenue par extrusion de la solution de MLV au travers d'une membrane de polycarbonate contenant des pores de tailles définies de 100 nm ou 200 nm (Avanti Mini Extruder).

Gel filtration

Les LUV sont ensuite purifiées par chromatographie d'exclusion stérique (SEC pour *size exclusion chromatography*), encore appelée gel-filtration, afin de séparer les LUV des fluorophores libres en solution. Une colonne en verre est préparée avec 6 mL de billes de Sepharose 4B (Sigma Aldrich). Les billes sont mises à reposer pour la nuit, permettant la formation d'environ 4 à 5 mL de gel. La colonne est équilibrée avec au moins 10 volumes de tampon d'élution. La solution de vésicules ($\sim 200\mu\text{L}$) est déposée et éluee par gravité. L'éluat est collecté par fraction de 5 gouttes ($\sim 200\mu\text{L}$). Le profil d'élution présente deux pics bien séparés correspondant pour le pic précoce aux liposomes, et pour le pic plus tardif, aux molécules fluorescentes non encapsulées dans les liposomes. Après utilisation, la colonne

6.1 Matériels et Méthodes

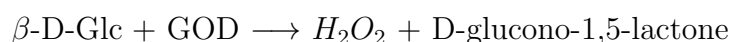
est lavée à l'eau puis stockée dans du TRIS-HCl 10 mM, pH8.0 à 4°C jusqu'à la prochaine utilisation. Lorsque la concentration en fluorophore est inférieure au micromolaire, ce qui est le cas pour la préparation des vésicules de 200 nm de diamètre, une purification à l'aide de spin-colonnes (*illustra MicroSpin S-300 HR, GE Healthcare*), utilisées selon le protocole du fabricant, s'avère suffisante.

Couplage des liposomes sur la lamelle

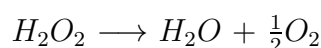
La solution de liposomes est diluée pour obtenir la concentration appropriée, puis est déposée pendant quelques minutes sur la couche de streptavidine. Après un rinçage permettant d'éliminer les vésicules inefficacement attachées, la lamelle est montée sur une lame en utilisant des espaceurs en silicone de 0.5mm d'épaisseur (Press-to-Seal, Invitrogen), définissant un volume de tampon d'environ 300 μ L baignant les liposomes.

Système de piégeage d'oxygène

Un système de piégeage d'oxygène est ajouté au tampon de mesure pour dépléter le milieu de l' O_2 responsable de réactions d'oxydation des états excités des molécules fluorescentes (plus réactifs que l'état fondamental) et dont la conséquence est le photoblanchiment des sondes. Ce système est composé de glucose et de deux enzymes, une Glucose Oxydase (*GOD*) associée à une Catalase (*CAT*), qui agissent en deux étapes. Le *Glc* est oxydé par la *GOD* selon la réaction :



ensuite, le peroxyde d'hydrogène généré est transformé par la *CAT* comme suit :



Ainsi, pour chaque cycle de cette réaction, $\frac{1}{2}O_2$ est consommé.

La déplétion de l' O_2 est importante pour augmenter le temps avant blanchiment des sondes de type cyanine (Cy3, Cy5), alors que peu d'effets sont observés pour le TMR.

L'hydrolyse partielle progressive et spontanée de la D-glucono-1,5-lactone en acide glucuronique induit une modification du pH du tampon de mesure. Il convient donc de renouveler

6.1 Matériels et Méthodes

le tampon régulièrement pour maintenir les propriétés du tampon (toutes les heures environ à 20°C).

6.1.3 Validation de l'approche

Afin de s'assurer que l'approche utilisée pour caractériser l'interaction de la NCp7 avec des ODN à l'échelle de la molécule unique est valide, différents contrôles ont été effectués. Ils concernent d'une part la construction servant à immobiliser les molécules et, d'autre part, la qualification de l'appareillage de mesure.

Caractérisation des nano-conteneurs

Les molécules fluorescentes sont immobilisées par l'intermédiaire de vésicules lipidiques d'un volume de quelques attolitres pour lesquels il convient de vérifier que leur processus de fabrication est fiable et reproductible et que ces objets sont suffisamment stables pour permettre leur utilisation à cette fin. J'ai donc réalisé différentes expériences permettant de caractériser ces nano-objets.

Analyse par FCS de vésicules encapsulant du TMR

Pour vérifier que l'assignement des deux pics du profil d'élution est correct, nous avons utilisé la FCS afin d'explorer les propriétés de diffusion des espèces contenues dans les fractions correspondantes. La FCS permet de caractériser la diffusion d'espèces marquées. La diffusion des espèces contenues dans chacun des deux pics est très différente (Figure 6.2). Le pic d'élution tardif contient des espèces diffusant rapidement ($\tau = 46 \mu\text{s}$) avec un temps de diffusion très similaire au TMR libre en solution. Au contraire, le pic d'élution précoce contient une seule population de particules diffusant lentement. En utilisant le coefficient de diffusion absolu de la rhodamine libre dans l'eau ($D = 421 \mu\text{m}^2\text{s}^{-1}$) [410], nous avons estimé que le coefficient de diffusion de ces espèces vaut $D = 4.2 \pm 0.2 \mu\text{m}^2\text{s}^{-1}$. Un coefficient très similaire de $D = 3.8 \pm 0.2 \mu\text{m}^2\text{s}^{-1}$ a été observé pour des particules de même diamètre, préparées avec un lipide marqué de manière covalente à la Rh6G, et mesurées dans les mêmes conditions.

6.1 Matériels et Méthodes

Le coefficient de diffusion d'une sphère peut être estimé à l'aide de l'équation de Stokes-Einstein

$$D = \frac{k_B \cdot T}{6\pi \cdot \nu \cdot R} \quad (6.1)$$

dans laquelle k_B est la constante de Boltzmann (13.8×10^{-24} J/K), T la température en Kelvin (293 K), ν la viscosité et R , le rayon de l'objet (50×10^{-9} m). En utilisant $\nu = 0.89 \times 10^{-3}$ kg/m/s pour la viscosité de l'eau à température ambiante, et le coefficient de diffusion observé pour ces vésicules, le rayon de ces particules est estimé à 57 ± 3 nm et 63 ± 3 nm, pour les vésicules encapsulant du TMR et celles marquées sur leur membrane par de la rhodamine, respectivement. Ces valeurs sont en parfait accord avec des mesures de DLS (*diffusion light scattering*) montrant que les liposomes extrudés forment une population unique et homogène avec un rayon de 55 ± 5 nm. De plus, la répétition des mesures de FCS au cours du temps ne montre pas de changement de la diffusion des liposomes, montrant que ceux-ci ne fusionnent pas spontanément entre eux et que leur structure reste intacte pendant plusieurs heures.

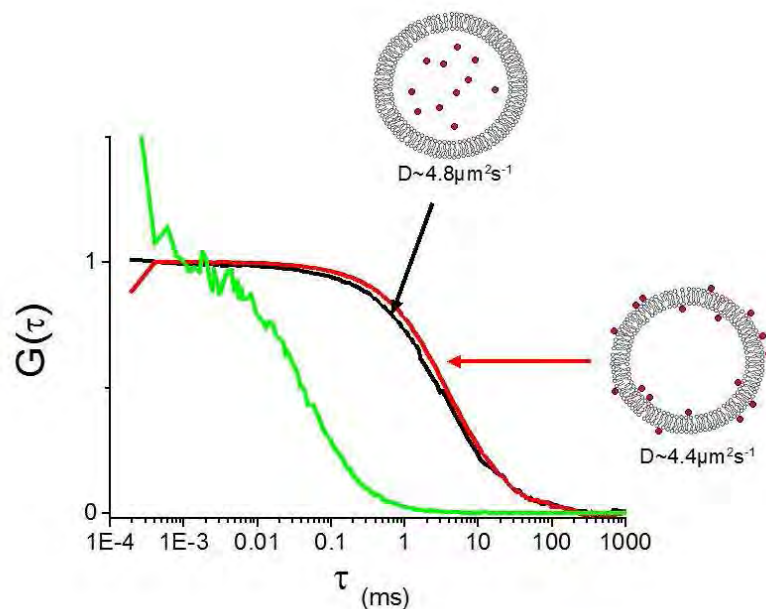


FIGURE 6.2 – Courbe d'autocorrélation du TMR libre (vert), et de liposomes encapsulant du TMR (noir) ou marqués par de la Rhodamine (rouge)

6.1 Matériels et Méthodes

Diffusion des liposomes attachés à la surface

Afin de vérifier le processus d'immobilisation, des vésicules encapsulants du TMR sont attachées sur des lames de verre et la préparation est imagée en utilisant le montage TPE-SFM décrit en 6.1.1. Les nanovésicules encapsulant le TMR apparaissent comme des points fluorescents, immobiles, photoblanchissables, juste au dessus de l'interface verre/eau (Figure 6.3). La taille des vésicules est bien inférieure à la résolution du TPE-SFM ($\lambda/(2 \cdot NA) \sim$

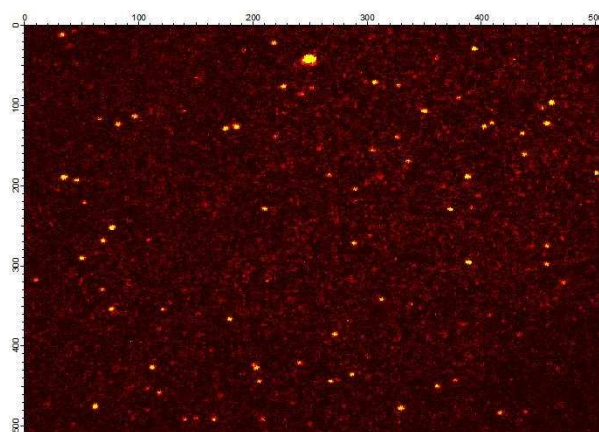


FIGURE 6.3 – Signaux de fluorescence de liposomes immobilisés encapsulant du TMR

400nm). Par conséquent, aucune information sur la taille ou sur la forme de ces vésicules ne peut être obtenue. Par contre, l'homogénéité de la préparation en terme de fluorescence peut être estimée. Une analyse des images montre en effet que les intensités des *spots* sont très homogènes. De plus, la préparation est stable durant plusieurs heures puisque l'on retrouve un profil d'image similaire 24 heures plus tard (stockage à 4°C à l'abri de la lumière). Enfin, la diffusion de ces *spots* est quasi-nulle, en tout cas bien inférieure à la dérive de l'appareil de mesure.

Mesure en microscopie de force atomique de liposomes immobilisés

Dans le but d'obtenir des informations topographiques sur la forme et la stabilité de ces liposomes attachés, des expériences d'AFM ont été réalisées en solution. Les liposomes sont immobilisés sur une surface de verre. L'acquisition des images s'est faite sous phase liquide (tampon TRIS HCl, pH 7.5, 30mM NaCl et 0.2mM MgCl₂). L'instrument utilisé est un

6.1 Matériels et Méthodes

microscope à force atomique commercial (Solver Pro NT-MDT, Moscou). Les *cantilevers* mis en oeuvre sont de type *tapping mode* avec une fréquence de résonance dans le milieu liquide de 32 kHz . Les images ont été acquises avec une résolution de 512×512 , une vitesse de balayage de 2 Hz . Le logiciel de *post-processing* des images est Nova (NT-MDT, Moscou). Les images montrent des liposomes homogènes, immobiles. Leur hauteur moyenne est voisine de 85 nm .

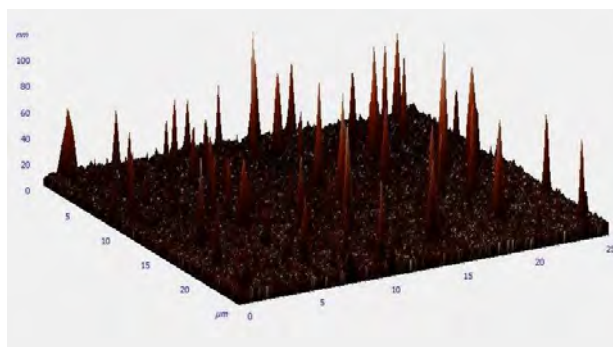
Qualification du montage expérimental

Vérification de l'alignement du montage

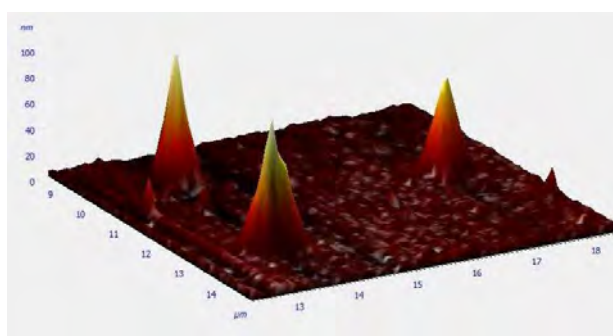
L'alignement de la voie de détection a été réalisé au cours du montage de la plateforme en utilisant une échelle micrométrique gravée sur une lame de microscope. Cette lame, éclairée par transmission de lumière blanche (filtres enlevés) est imagée afin d'aligner au mieux les deux voies de détection.

L'alignement final est vérifié en utilisant des billes fluorescentes émettant dans les deux voies de détection simultanément (crimson fluorescent (625/645) $1\mu\text{m}$ carboxylate-modified FluoSpheres[®]). Une superposition exacte des images provenant des deux voies n'est malheureusement pas possible en raison d'aberrations chromatiques, ce qui empêche de réaliser directement des images de FRET. Cette limitation technique peut être levée au moment du *processing* des images en utilisant une procédure de réaligement utilisant par exemple *UnwarpJ* [411], un *plugin* développé pour *ImageJ* qui permet de générer des déformations élastiques, sur la base de *splines* cubiques, d'une image source de manière à ce qu'elle ressemble au maximum à une image cible. Il est ensuite possible d'appliquer ce 'masque' de déformation à tout un ensemble d'images. Pour autant, les tests de traitement d'image incluant cette procédure de réaligement effectués sur des séquences d'images de billes immobilisées ou de liposomes, ont montré que l'efficacité de celle-ci n'est pas optimale lorsque la correspondance entre les deux images est limitée, ce qui est très souvent le cas dans nos expériences où le nombre de *spots* dans la voie du donneur est très supérieur à ceux de la voie de l'accepteur de fluorescence.

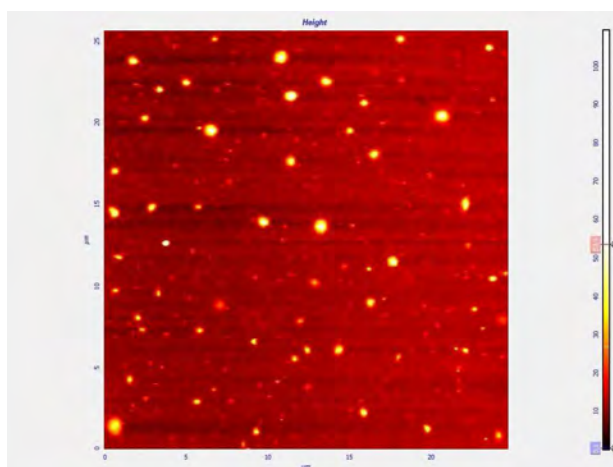
6.1 Matériels et Méthodes



(a)



(b)



(c)

FIGURE 6.4 – Images topographiques des liposomes immobilisés par microscopie à force atomique

Fluorescence à l'échelle de la molécule unique de Quantum dots

L'utilisation de boîtes quantiques, ou quantum dots (QDot) immobilisés en tant que contrôle est trivial puisque le QDot constitue un des nano-objets les plus luminescents actuellement disponibles. Nous avons suivi la fluorescence de QDots couverts par de la

6.1 Matériels et Méthodes

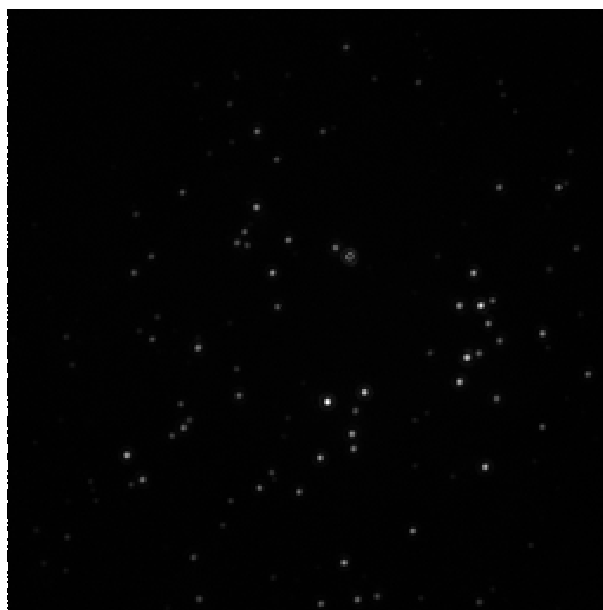


FIGURE 6.5 – Image de QDot605 immobilisés sur une surface de verre (SD image of 150 frames)

streptavidine (*Qdot[®] 605 streptavidin conjugate Invitrogen*) immobilisés sur une surface de verre silanisée et biotinilée (voir Figure 6.5 pour une image représentative). Vu l'étroitesse de la bande d'émission du QDot, centrée sur $605nm$, aucun signal n'est observé dans la voie de l'accepteur et seule la voie du donneur est représentée.

Une trace représentative de la fluorescence d'un QDot est présentée figure 6.6. Les fortes fluctuations de fluorescence observées sont caractéristiques du *blinking* des QDots, et assurent qu'un seul Qdot est observé. Il est également intéressant de mentionner la distribution selon une loi de puissance des temps de résidences correspondant aux états *on* (brillant) et *off* (noir).

6.1 Matériels et Méthodes

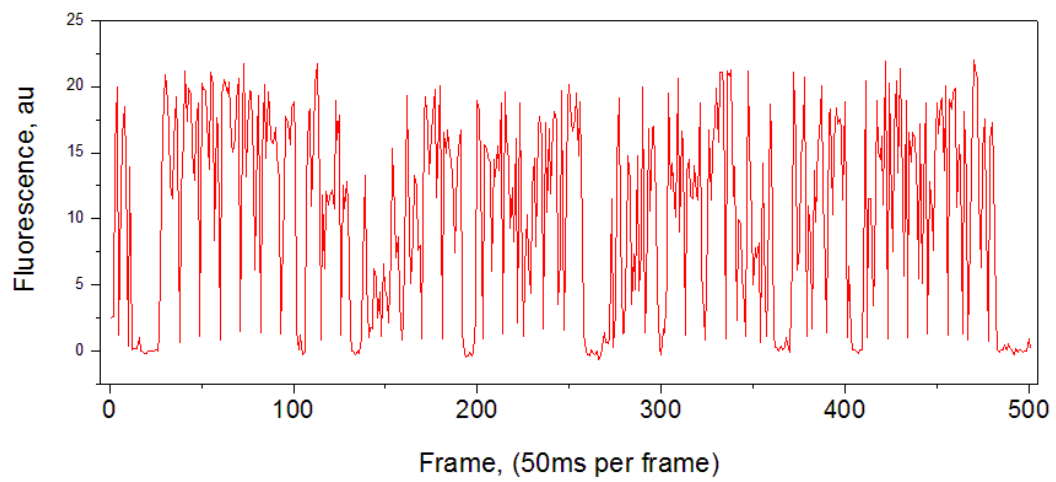


FIGURE 6.6 – Trace de fluorescence d'un QDot605

6.2 Traitement et analyse des signaux de fluorescence

La programmation d'outils permettant le traitement et l'analyse des données a représenté une part significative du travail effectué au cours de ma thèse.

6.2.1 Traitement des images

Les acquisitions des séquences vidéos sont réalisées à l'aide du logiciel *MicroManager* qui contrôle la caméra ainsi que les obturateurs de la plateforme. *Micromanager* génère des séries temporelles d'images (*.tif). Ces images sont ensuite empilées en utilisant une macro maison sous *ImageJ* (code source en annexe) de manière à générer une image 3D ($x, y, z = \text{frame}(t)$). À partir de ce genre d'objet, il est assez facile de générer le profil temporel (Z-profile) d'une aire sélectionnée afin de reconstruire l'évolution temporelle du signal. Les coordonnées des spots fluorescents sont obtenues à partir d'une image de déviation standard (ou SD-image) où la valeur de chaque pixel correspond à la variance du-dit pixel sur les $\sqrt{n+1}$ premières images de la séquence. Le choix d'une SD-image permet de sélectionner les spots où le signal de fluorescence est le plus fluctuant. Un seuillage est appliqué sur l'image, les spots sont identifiés et leurs coordonnées en sont extraites. Ces coordonnées sont ensuite utilisées pour générer le profil temporel de chacun des spots (et de la mesure du bruit environnant associé). Ces profils sont enregistrés dans des fichiers (*.txt). Cette procédure de traitement (code en annexe) peut fonctionner en mode *batch*, permettant de traiter les acquisitions par lot.

6.2.2 Traitement des signaux de fluorescence

Cette étape de traitement est réalisée sous *R*, un logiciel libre dédié au traitement de données et à l'analyse statistique [412]. Une procédure de traitement maison a été écrite afin de permettre le traitement des données générées par nos expériences (code des fonctions en annexe). La figure 6.7 récapitule les traitements qu'il est possible de réaliser. Rapidement, les données issues de *ImageJ* sont importées, débruitées et filtrées. Dans le cadre de données de FRET, les coordonnées des donneurs et des accepteurs de fluorescence sont appariées afin de calculer le FRET. Pour les données d'intensité de fluorescence, une procédure de

6.2 Traitement du signal

classification de courbes permet de générer, si nécessaire, des clusters de courbes aux profils similaires. Les signaux corrigés sont colligés au sein d'une *dataframe* (matrice de données) à partir de laquelle il est possible :

- de représenter graphiquement les traces de fluorescence ou de FRET
- de représenter graphiquement une matrice des transitions inter-états
- de déterminer quand la molécule photoblanchit
- de représenter la localisation spatiale des spots correspondant à ces traces

Suite à une procédure de fit multi-dimensionnel selon un modèle de Markov caché, il est également possible :

- de représenter graphiquement chacune des traces individuelles avec leur fit correspondant
- d'estimer les paramètres de l'état stationnaire du système (proportion relative de chacun des états)
- d'obtenir la matrice de transition ainsi que le positionnement et la dispersion des différents états
- de trouver/quantifier l'existence de transitions particulières ou de séquences de transitions particulières
- de représenter et de fitter les histogrammes des temps de résidence pour chacune des transitions inter-états

Afin de bien comprendre comment est extraite l'information à partir des traces de fluorescence ou de FRET, il convient de détailler *a minima* les modèles de Markov à états cachés.

6.2 Traitement du signal

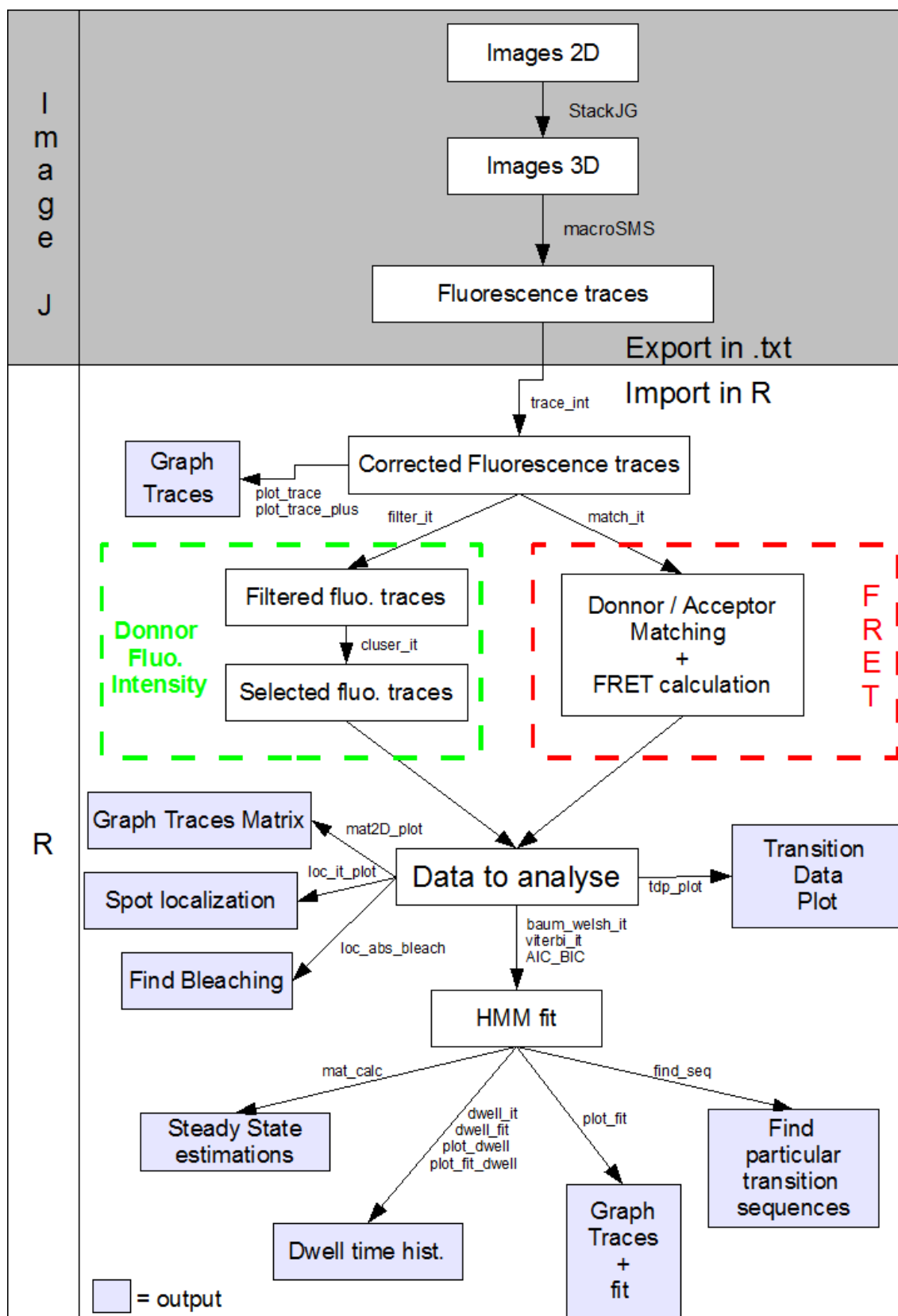


FIGURE 6.7 – Schéma synthétique de la procédure de traitement des données

6.2.3 Modélisation des trajectoires de fluorescence ou de sm-FRET par les méthodes de Markov à états cachés

L'objectif de la spectroscopie de fluorescence à l'échelle de la molécule unique est généralement d'obtenir des informations cinématiques ou mécanistiques d'un système à partir de traces de fluorescence ou de combinaisons de traces de fluorescence (comme le FRET). Afin d'obtenir ces informations, il faut être capable de discriminer les différents états du système étudié. Ceci peut être fait en analysant *à l'oeil* les différentes traces de fluorescence par des méthodes de seuillages plus ou moins complexes ou encore à partir de données condensées sous forme d'histogramme par exemple. Or, puisque les signaux résultants de la détection des photons émis par une seule molécule sont par nature bruités, ces méthodes sont sujettes à biais et les conclusions obtenues par différents expérimentateurs peuvent diverger, en particulier si la connaissance du bruit expérimental, de la photophysique des sondes utilisées ou encore du système étudié lui-même diffère entre les expérimentateurs. De plus, puisque le principe de la méthode consiste à observer de manière individuelle les molécules d'un système, la description de celui-ci requiert l'analyse d'un grand nombre de molécules (typiquement plusieurs centaines voir quelques milliers) afin de reconstituer et de décrire fidèlement le comportement global du système étudié. Dès lors, une analyse basée sur une inspection visuelle des traces peut s'avérer laborieuse. Enfin, il faut être capable de différencier les changements inhérents au bruit de ceux qui correspondent à des transitions rapides mais véritables, là encore, une analyse subjective peut se montrer limitée.

Afin de proposer une solution à ce problème de traitement de signal, plusieurs méthodes ont été proposées dans la littérature (voir [413] pour une revue) incluant le test de Page (qui permet de détecter des changements de la fonction de distribution d'un signal bruité) ou encore des algorithmes basés sur l'utilisation de test de χ^2 avec des données simulées [414]. L'utilisation des méthodes de Markov à états cachés (Hidden Markov Models ou HMM) apporte une réponse intéressante à ce type de problème. L'utilisation des HMM en biophysique a tout d'abord été appliquée à l'étude des canaux-ioniques [415, 416], puis récemment, leur recours s'est avéré intéressant pour le traitement des données de fluorescence en mo-

6.2 Traitement du signal

lécule unique [417, 418] Plus récemment encore, Mc Kinney & al. (2006) [419] ont proposé un HMM qui repose sur des acquisitions de photons binnées temporellement. Les méthodes d'analyses décrites ici sont largement inspirées de ces travaux.

Théorie des HMM

Les méthodes de Markov à états cachés sont des réseaux dynamiques bayésiens dans lesquels le système modélisé est supposé suivre des processus de Markov, c'est à dire qu'au cours du temps, l'état $X^{(n)}$ du système dépend de l'état précédent $X^{(n-1)}$ et uniquement de celui-ci selon :

$$X^{(n)} = X^{(n-1)}.Tr$$

où Tr est la matrice de transition, et qui peut encore s'écrire en fonction des conditions initiales

$$X^{(n)} = X^{(0)}.Tr^n$$

Ainsi, un processus de Markov consiste en une multitude de combinaisons de transitions inter-états, chacune gouvernée par un déclin à une exponentielle.

Lorsque l'on observe les traces de fluorescence, on a accès à des signaux qui, bien que dépendants des états du système observé, peuvent, en raison du bruit, fluctuer fortement par rapport à la valeur de l'état considéré. C'est la raison pour laquelle on dit que les états sont cachés : l'observateur n'est pas capable d'observer directement les différents états du système.

L'objectif de l'HMM est donc d'inférer les états à partir des observations, et ce à l'aide de différentes méthodes d'estimation et d'optimisation. Le concept de l'HMM est de déterminer un modèle de paramètres λ qui maximise la probabilité des observations conditionnellement à λ ($Prob(Obs|\lambda)$). Les paramètres λ sont constitués des probabilités de transition et des probabilités d'émission ainsi que des *priors* du modèle. Les probabilités de transition déterminent les fréquences des transitions qui existent entre les différents états cachés (Fig 6.8.). Les probabilités d'émission représentent la vraisemblance relative d'observer une valeur donnée, conditionnellement à un état de FRET donné (Fig 6.9.). Dans nos analyses et bien que les distributions des états de FRET suivent en réalité des lois *beta*, nous avons

6.2 Traitement du signal

utilisé une approximation par une fonction Gaussienne d'écart-type δ (les erreurs générées par ce type d'approximation sont minimales) [420] tel que :

$$ep_{\phi}(FRET_{Obs}) \propto exp \left[-2 \times \left(\frac{FRET_{Obs} - FRET_{\phi}}{\delta} \right)^2 \right]$$

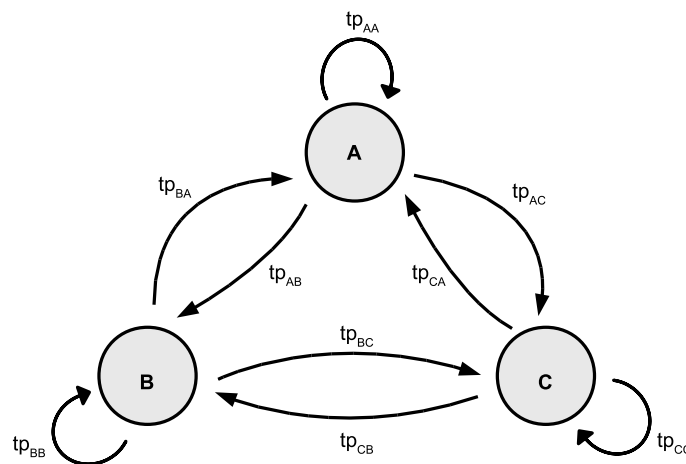


FIGURE 6.8 – Illustration des probabilités de transition entre 3 états A-B-C.

Si le système est dans l'état A au temps t , il a la probabilité tp_{AA} de rester dans cet état, et les probabilités tp_{AB} et tp_{AC} de passer à l'état B ou C, respectivement, au temps $t + 1$.

L'HMM se résume donc à trois étapes :

- déterminer $Prob(Obs|\lambda)$
- trouver la trajectoire la plus probable entre les différents états cachés en fonction des données
- enfin, optimiser les paramètres qui maximisent $Prob(Obs|\lambda)$

La probabilité d'être dans un état donné au moment t correspond au produit de la probabilité d'émission et de la probabilité de transition.

$$Prob(state(t)) = ep_{\phi}(FRET_{data}(t) \times tp(state(t-1), state(t)))$$

6.2 Traitement du signal

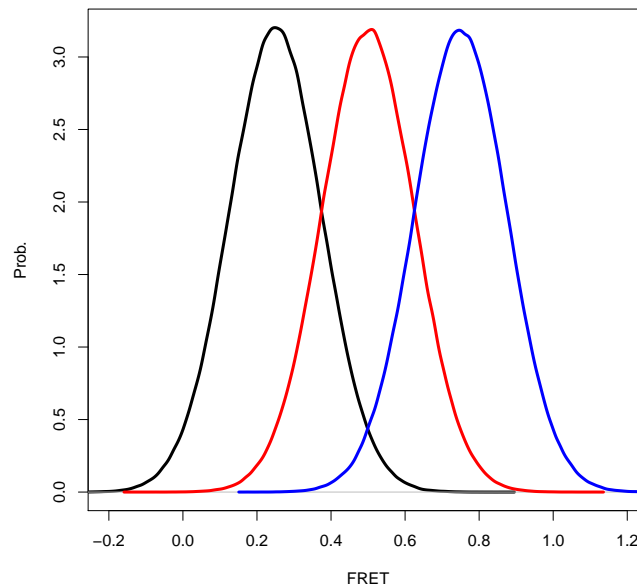


FIGURE 6.9 – Illustration des probabilités d’émission d’un système à trois états. Pour un état donné, la probabilité d’émission est distribuée selon une loi normale avec une moyenne qui correspond à l’état caché (respectivement 0.25, 0.50 et 0.75 pour A, B et C)

La probabilité de la trajectoire entière correspond au produit de l’ensemble des probabilités individuelles

$$Prob(trajetoire) = \prod_{t=1}^N Prob(state(t))$$

Algorithme de Viterbi

Pour le calcul de $Prob(Obs|\lambda)$, on pourrait sommer l’ensemble des probabilités de l’ensemble des trajectoires possibles. Or pour un système à N états de longueur T , il existe N^T trajectoires, ce qui rend rapidement ce calcul impossible. L’algorithme de Viterbi [421] permet de calculer le chemin des états cachés le plus probable ayant produit les données observées. Le but de l’algorithme de Viterbi est de décoder l’état sous-jacent du système pour tout t (cf section annexe pour le code).

Algorithme de Baum-Welch

L’optimisation des paramètres qui maximisent la probabilité $Prob(Obs|\lambda)$ est un problème d’optimisation multi-dimensionnel classique qui peut être résolu par des méthodes

6.2 Traitement du signal

telles que l'algorithme de Brent, des techniques basées sur l'utilisation de gradient [422] ou encore l'algorithme de Baum-Welch [423]. McKinney & al. [419] ont fait le choix de l'algorithme de Brent, qui est plus simple mais plus lent. Nous avons choisi d'utiliser l'algorithme de Baum-Welch qui est un cas particulier d'algorithme d'espérance-maximisation généralisé (generalized expectation-maximisation or GEM) proposé par Dempster & al. (1977) [424], permettant de trouver le maximum de vraisemblance des paramètres de modèles probabilistes lorsque le modèle dépend de variables latentes non observables. L'algorithme permet donc de trouver les probabilités de transition et d'émission d'un HMM à partir des émissions observées par réestimations itératives (cf section annexe pour le code).

Simulations de traces de FRET

Des séries de données de différentes tailles, résultant de processus de Markov de 2 à 5 états, plus ou moins bruitées, ont été générées afin de vérifier que le traitement de ces traces permettait de retrouver les matrices de transitions ayant générées ces signaux.

À titre d'exemple, nous nous intéresserons à la trace présentée figure 6.10, qui représente typiquement le genre de signal expérimental que l'on est capable de mesurer. La détermination des états de FRET sous-jacents n'est pas évidente et on se trouve bien dans une situation où les états du Markov ne nous sont pas accessibles. On imagine également qu'il serait difficile *à l'oeil* de déterminer quels sont les états de FRET de ce système. De même, il est difficile de conclure à partir de l'histogramme présenté figure 6.11b sur les différents états de FRET du système (dont les positions sont représentées par les 3 lignes verticales).

La trace de la figure 6.10 est obtenue à partir de la trace théorique représentée figure 6.11a, résultant d'un processus de Markov à 3 états de FRET (respectivement 0.25, 0.5 et 0.75), générée par la matrice de transition Tr_0

$$Tr_0 = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$$

Pour obtenir la trace de la figure 6.10, la trace théorique a ensuite été bruitée par ajout

6.2 Traitement du signal

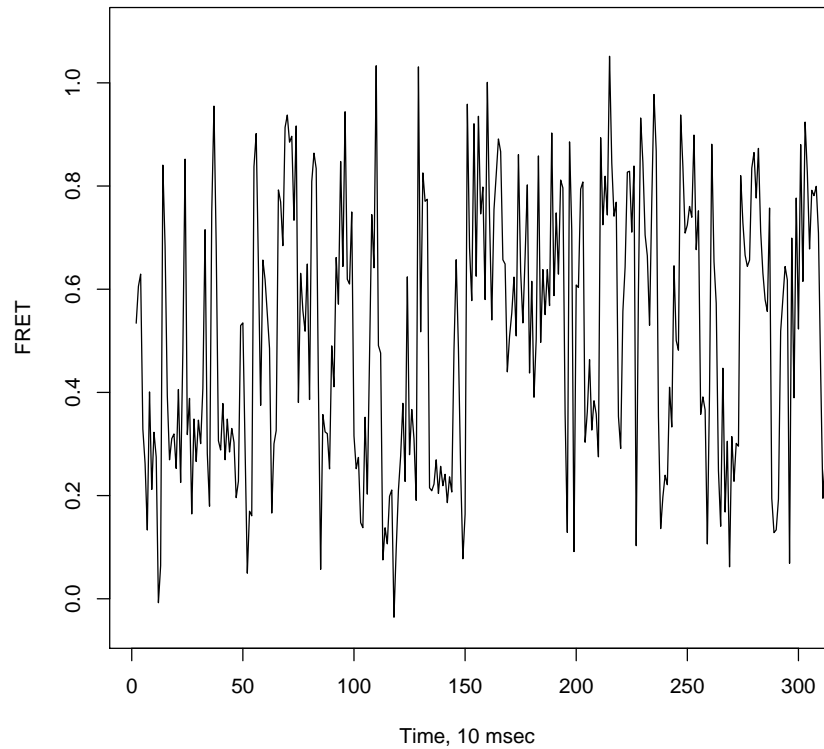


FIGURE 6.10 – Trace de FRET simulée (300 premiers points).
Il est difficile de dire, à la seule vue de ce signal, quels peuvent être les états cachés du HMM.

d'un signal généré de manière aléatoire selon une loi normale $\mathcal{N}(\mu = 0, \sigma = 0.17)$, tel que

$$ep = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

Il convient de remarquer ici que la valeur de $\sigma = 0.17$ est un peu surévaluée par rapport aux observations expérimentales (typiquement $\sigma = 0.1$), mais cette valeur permet de montrer l'efficacité du traitement des données.

L'HMM est ensuite appliquée sur ce signal bruité de manière à retrouver la séquence cachée. L'utilisation de l'algorithme de Baum-Welch permet de trouver, après initialisation, la matrice de transition et les états du système les plus probables. En initialisant grossièrement le système avec 3 états valant 0.2, 0.6 et 0.8 respectivement, on obtient après maximisation de la vraisemblance après 211 itérations, la matrice de transition estimée Tr_{est}

6.2 Traitement du signal

$$Tr_{est} = \begin{pmatrix} 0.80 & 0.10 & 0.10 \\ 0.25 & 0.45 & 0.29 \\ 0.08 & 0.10 & 0.82 \end{pmatrix}$$

et les états de FRET ainsi que leurs écart-types

| | état 1 | état 2 | état 3 |
|-------|--------|--------|--------|
| μ | 0.247 | 0.497 | 0.749 |
| sd | 0.128 | 0.144 | 0.128 |

À l'aide de ces paramètres, et en utilisant l'algorithme de Viterbi, il nous est possible de déterminer le chemin le plus probable entre les différents états de FRET. En superposant la trace de valeurs observée et celle obtenue par l'algorithme de Viterbi, on se rend bien compte que la courbe obtenue (en rouge en haut de la figure 6.12) permet de fitter parfaitement les observations. En comparant directement ce fit (en rouge en bas de la figure 6.12) avec la courbe théorique qui a servi initialement à générer ces données (en bleu sur la même figure), on voit également que les différences entre les deux courbes sont minimales, se résumant à des transitions relativement rapides pour lesquelles un bruit important a rendu la confusion entre deux états contigus possible.

Il est possible de vérifier que les données obtenues décrivent bien le système pour lequel on connaît les différents paramètres, comme la matrice de transition ou la matrice de l'état stationnaire. Il suffit pour cela par exemple de compter les différentes transitions au sein de cette trace de fit et de les rapporter à la longueur de la trace observée pour connaître la matrice de transition observée Tr_{obs} . On trouve ici

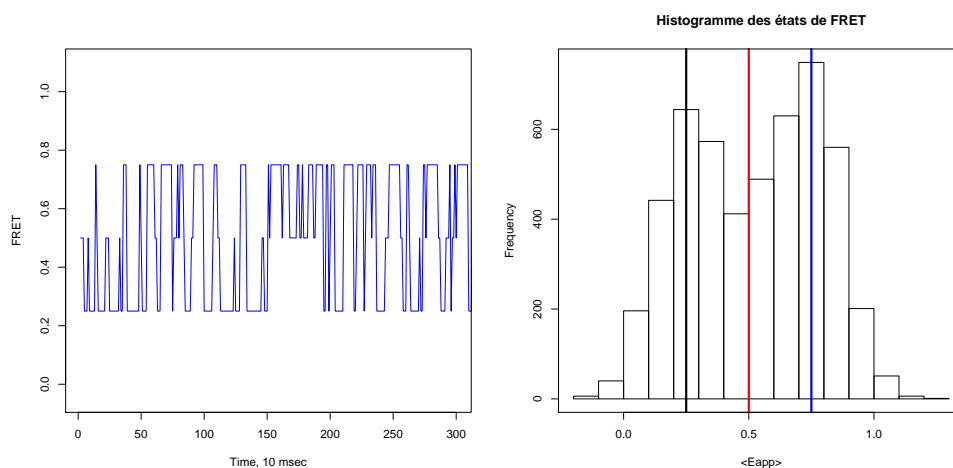
$$Tr_{obs} = \begin{pmatrix} 0.84 & 0.06 & 0.10 \\ 0.19 & 0.55 & 0.25 \\ 0.09 & 0.06 & 0.85 \end{pmatrix}$$

ce qui reste, malgré la longueur conséquente de la trace (5000 points), une moins bonne approximation que la matrice estimée par Baum-Welch.

De même, on peut facilement calculer les proportions de chacun des états de FRET, ce qui correspond à la matrice stationnaire, et pour lesquelles on trouve

6.2 Traitement du signal

| | état 1 | état 2 | état 3 |
|-----------|--------|--------|--------|
| théorique | 0.40 | 0.14 | 0.46 |
| observée | 0.40 | 0.11 | 0.49 |



(a) Trace théorique

(b) Histogramme de la trace de la figure 6.10

FIGURE 6.11

Ce qui est intéressant dans le code développé ici est qu'il s'adapte facilement aux contraintes expérimentales des expériences en molécules uniques. En effet, alors qu'il faut avoir un nombre relativement significatif de points de mesure pour obtenir des résultats robustes (notamment pour pouvoir détecter des événements rares), le temps d'observation des molécules avant que celles-ci ne photoblanchissent est relativement court (typiquement quelques secondes dans le cas de la tétraméthylrhodamine (TMR)). En considérant un temps d'intégration de quelques dizaines de millisecondes, cela génère des traces de fluorescence qui sont de l'ordre de quelques dizaines de points. Il est donc intéressant, et nécessaire, de pouvoir estimer les paramètres du HMM en multivarié, sur plusieurs traces simultanément (ce qui n'est pas possible avec l'application proposée par Mc Kinney & al. (2006) [419] (Hammy) ou encore par Bronson & al. sous Matlab (vbFRET) [425])¹. Ceci peut être

1. Après la rédaction de ce chapitre, Liu & al [426] ont proposé une méthode de fit basée sur des HMM multivariés permettant l'analyse de traces multiples simultanées.

6.2 Traitement du signal

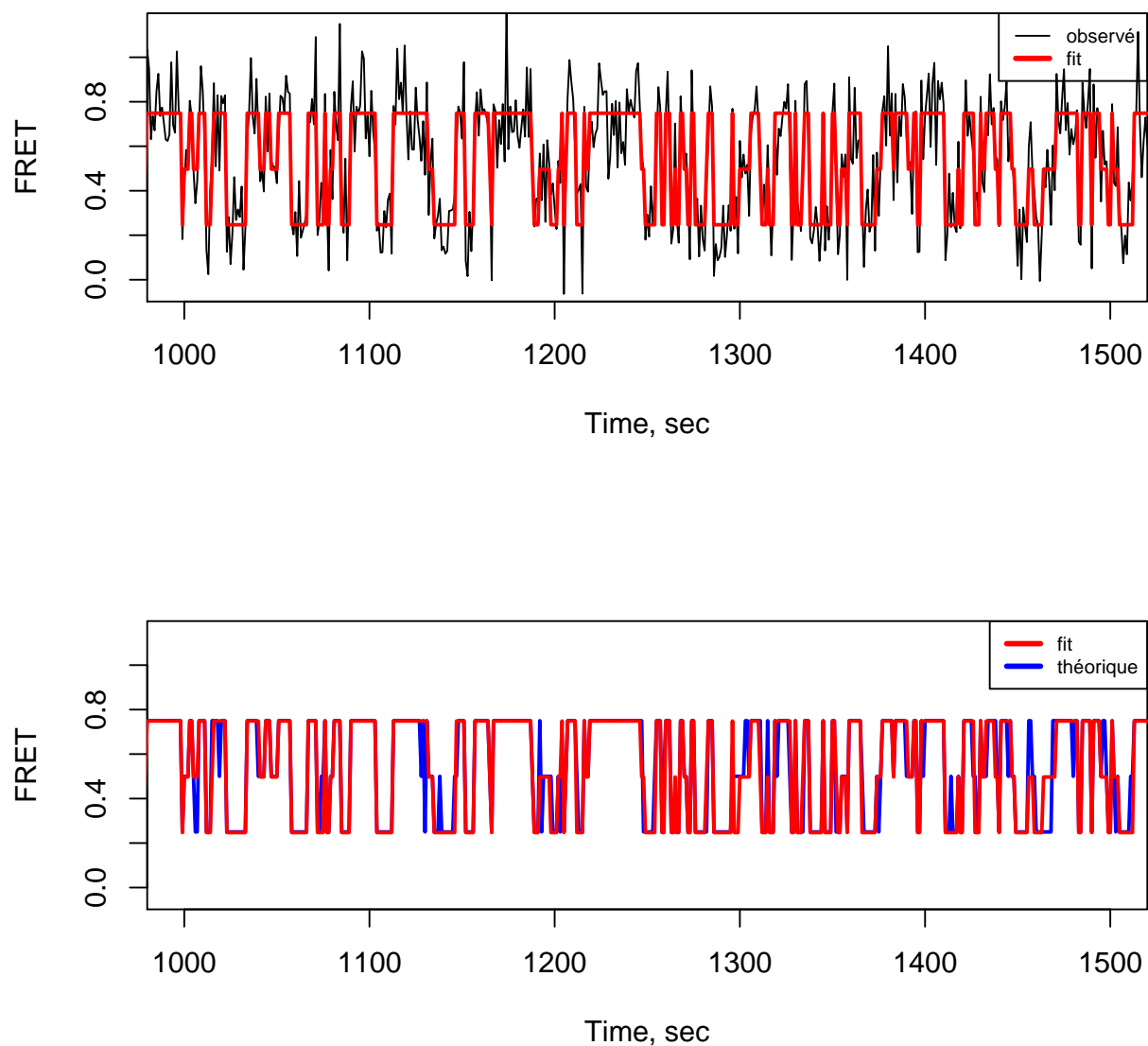


FIGURE 6.12 – Comparaison de la courbe de fit avec les données observées (en haut), et avec les données théoriques (en bas)

vérifié en générant une matrice de données qui contient n colonnes de m points expérimentaux. On peut fitter simultanément les n colonnes et générer des estimations uniques pour l'ensemble de ces données. La figure 6.13 montre par exemple ce que donne ce genre de fit sur les 16 premières fractions de 100 points de la trace utilisée précédemment. Pour tester cette procédure, le fit est réalisé de manière globale sur l'ensemble de ces 50 mesures de

6.2 Traitement du signal

100 points chacune et les mêmes paramètres que ceux trouvés lors du fit de la séquence complète de 5000 points ont été obtenus.

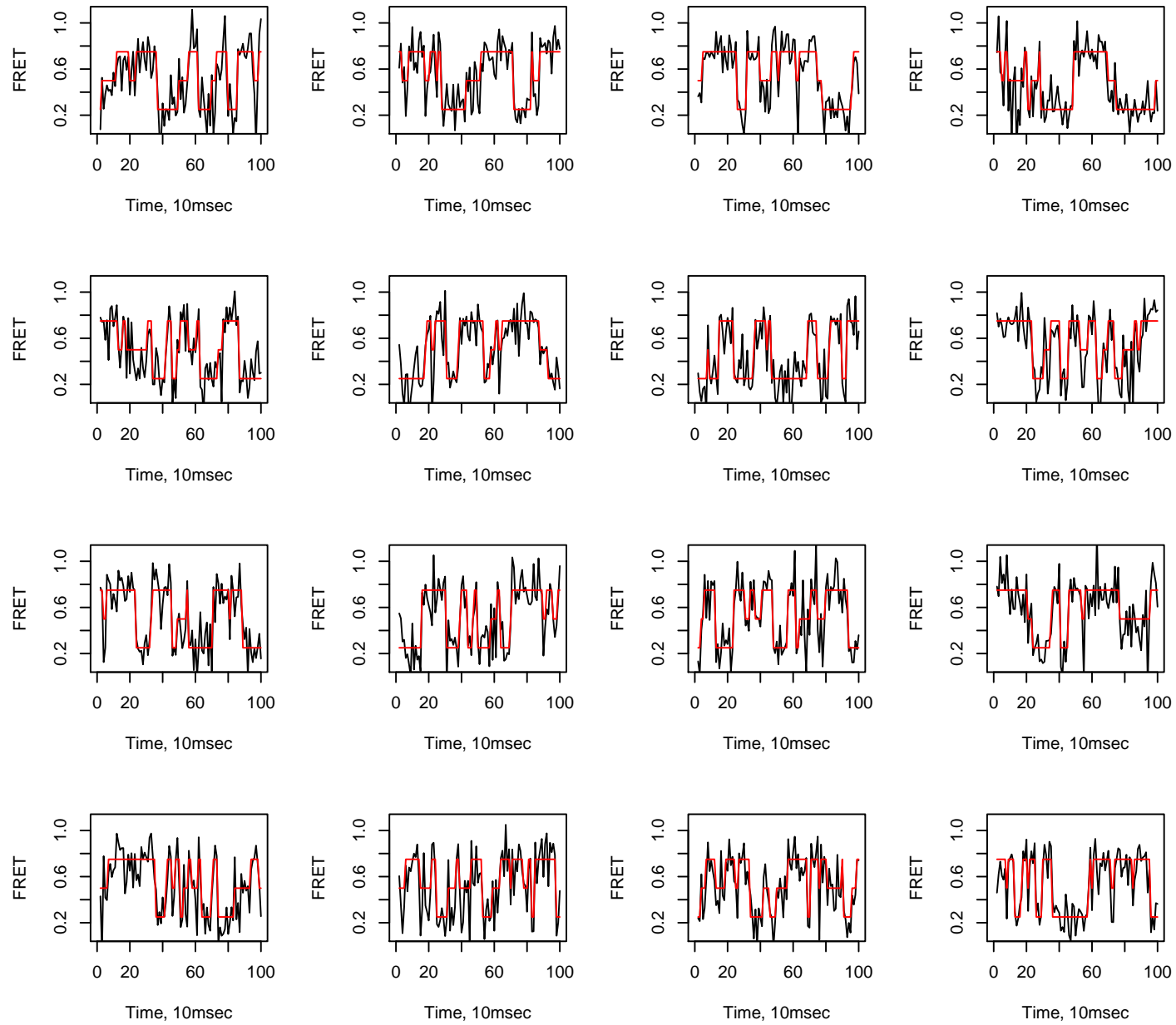


FIGURE 6.13 – Représentation de 16 traces de FRET et de leur fit

6.2 Traitement du signal

Choix du nombre d'états de FRET

Le dernier problème dont nous n'avons pour l'instant pas encore parlé est celui de la détermination du nombre des états de FRET du système. En effet, les processus d'optimisation que nous venons de présenter se calculent pour un nombre fixé d'états du système. Une détermination du nombre de ces états par l'expérimentateur, de même que sa connaissance *a priori*, n'est pas toujours possible. Une fois encore, nous allons faire appel aux méthodes classiques de statistique pour définir quel est le nombre d'états le plus probable. Pour cela, nous aurons recours à l'AIC (Akaike information criterion) [427] et au BIC (Bayesian information criterion) [428]. Tous deux reposent sur l'évaluation de l'adéquation du modèle aux données par l'évaluation du maximum de vraisemblance, mais plus qu'une simple comparaison de la log vraisemblance ($\ln(L)$), ces modèles prennent également en compte le nombre de paramètres qui sont utilisés par le modèle. En effet, on sait qu'augmenter le nombre des paramètres d'un modèle permet toujours d'améliorer l'adéquation entre les données et le modèle et de fait, si l'on ne prend en compte que la vraisemblance, on s'expose à un risque d'*over-fitting*. Ainsi, l'AIC et le BIC permettent de choisir le modèle optimal en pénalisant la $\ln(L)$ par un coefficient dépendant du nombre de paramètres des différents modèles. Par définition,

$$AIC = 2k - 2\ln(L)$$

$$BIC = k \cdot \ln(n) - 2\ln(L)$$

où k correspond au nombre de paramètres, n au nombre d'observation et $\ln(L)$ est la log vraisemblance. Ainsi, le modèle optimal donne la plus petite valeur d'AIC et de BIC.

Le calcul de l'AIC et du BIC pour des modèles à 2, 3, 4 ou encore 5 états donne les résultats suivants

| Nb d'états | 2 | 3 | 4 | 5 |
|------------|-----------|-----------|-----------|-----------|
| $\ln(L)$ | 970.609 | 1019.151 | 1020.810 | 1022.705 |
| AIC | -1933.217 | -2020.302 | -2009.619 | -1995.411 |
| BIC | -1930.958 | -2015.217 | -2000.58 | -1981.287 |

Le nombre d'états optimal (trois) correspond bien à celui qui a permis de générer ces données.

6.3 Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique

Dans un premier temps, nous avons observé la fluorescence de molécules de Rh-NCp7 encapsulées individuellement dans des liposomes de DOPC. Les liposomes ont été formés et immobilisés sur la surface d'observation selon les méthodes présentées dans la section 6.1. En illuminant la surface sur laquelle sont fixés ces liposomes avec un laser à 532 nm , on observe des *spots* de fluorescence sur la voie de détection du donneur ($580 \pm 30\text{ nm}$) qui s'éteignent rapidement après le début de l'illumination (Figure 6.14).

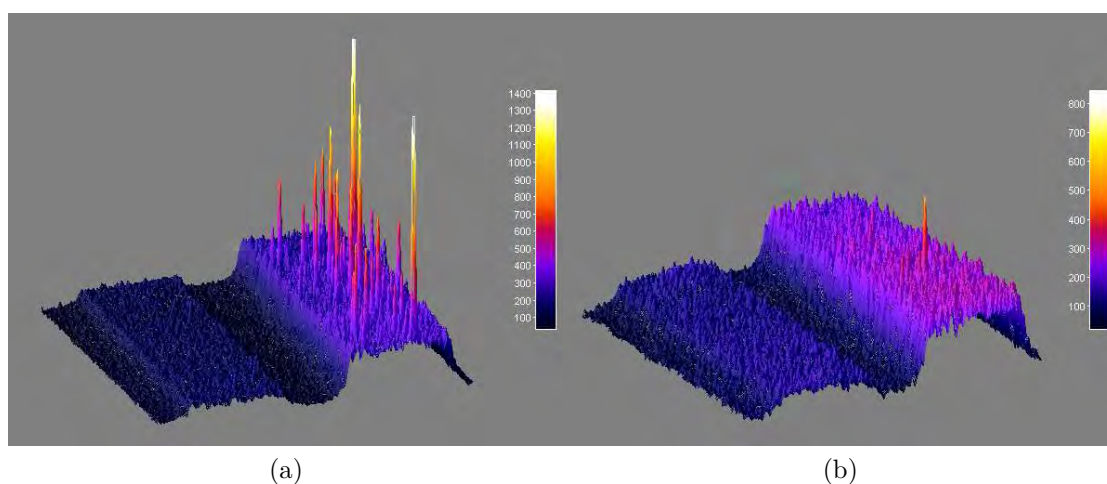


FIGURE 6.14 – Image 3D de la fluorescence de liposomes encapsulant de la Rh-NCp7 attachés sur la surface.

- (a) Les *spots* de fluorescence ne sont observés que dans le canal du donneur.
- (b) Les *spots* disparaissent après quelques secondes d'illumination

Ces *spots* correspondent à des liposomes encapsulant de la Rh-NCp7. Les traces de fluorescence correspondant à ces *spots* ont la forme d'une (ou de plusieurs) marche(s). La fluorescence émise par la protéine marquée est constante au cours du temps, jusqu'à ce que la sonde photoblanchisse (ou photobleache), moment à partir duquel plus aucun signal de fluorescence n'est enregistré. Certaines observations montrent deux, voir trois événements de photobleaching, montrant que les liposomes correspondant avaient encapsulés deux ou trois molécules marquées, respectivement (Figure 6.15).

Puisque chacune des mesures individuelles peuvent correspondre à des comportements non-représentatifs, il convient d'analyser un grand nombre de traces individuelles afin de

6.3 Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique

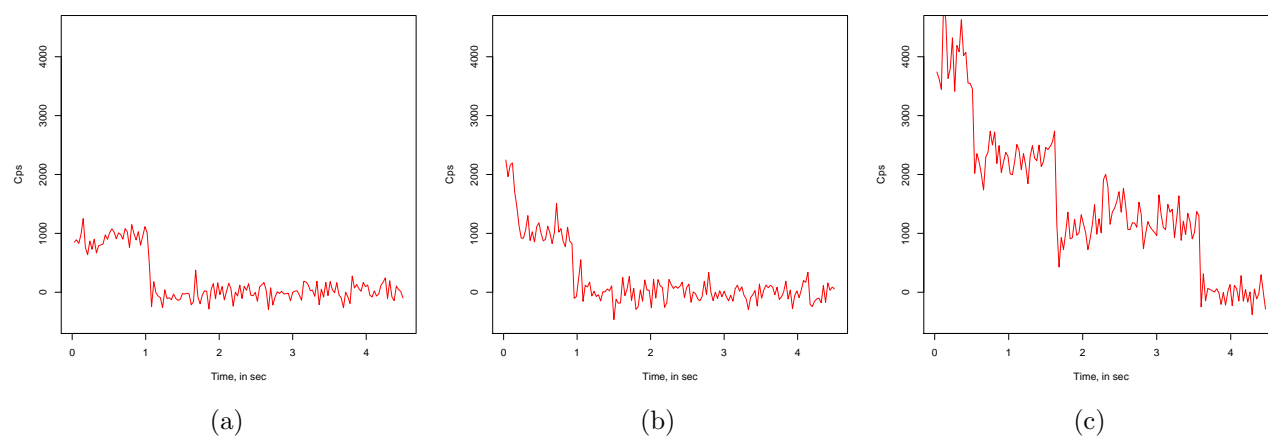


FIGURE 6.15 – Traces de fluorescence de Rh-NCp7 encapsulées dans des liposomes de 200 nm de diamètre. Les traces présentées correspondent à la fluorescence associée à des liposomes encapsulant respectivement (a) une, (b) deux ou (c) trois molécules de Rh-NCp7.

faire ressortir les comportements majoritaires. L'analyse d'un grand nombre de traces de fluorescence (688 pour la Figure 6.17 par exemple) permet de vérifier que la majorité des liposomes encapsulent bien une seule NCp7 marquée. Pour cela, nous avons quantifié le nombre de liposomes encapsulant une, deux ou trois molécules (Figure 6.16). Le nombre de liposomes n'encapsulant aucune Rh-NCp7 n'est bien évidemment pas accessible puisque ces objets ne sont pas fluorescents. Il est possible de déduire le nombre moyen m de molécules par liposome en ajustant la distribution des comptages individuels selon une loi de distribution de Poisson (losanges rouges sur la Figure 6.16) selon :

$$Prob(m) = A \times \frac{\lambda^m}{m!} e^{-\lambda}$$

On trouve $A = 2.6$ et une moyenne $\lambda \sim 0.50$.

Le nombre de molécules fluorescentes par liposome peut-être grossièrement évalué en tenant compte du rayon du liposome et de la concentration en espèces fluorescentes utilisées lors de la formulation des liposomes. Puisque le volume interne d'un liposome d'environ 200 nm de diamètre est de l'ordre de $4 \times 10^{-18} L$ (ou 4 atto litres), on s'attendrait à observer, pour une concentration de 400 nM, en moyenne 1 molécule par liposome. Puisque l'on en observe seulement la moitié ($\lambda \sim 0.50$), il faut en conclure que l'approximation faite ici est

6.3 Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique

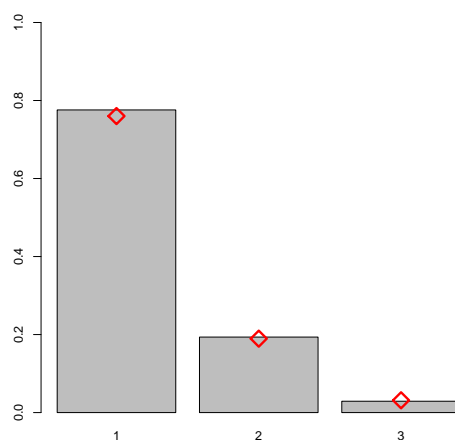


FIGURE 6.16 – Répartition de 688 vésicules en fonction du nombre de Rh-NCp7 encapsulé (barres) ajustée selon une loi de Poisson (losanges rouges).

grossière et que le processus d'encapsulation des Rh-NCp7 doit être affecté par d'autres paramètres.

L'analyse d'un très grand nombre de traces de fluorescence permet également de caractériser les paramètres de fluorescence de la population (ou des populations) de Rh-NCp7. Une représentation conjointe de l'ensemble de ces traces est disponible figure 6.17. On observe aux temps courts de l'acquisition, une distribution de la fluorescence dont le mode vaut 700-800cps/30ms. Cette population se dépeuple assez rapidement (en quelques secondes dans ces conditions d'acquisition) pour peupler, après photoblanchiment, une population pour laquelle le signal enregistré correspond au bruit résiduel. Le temps pendant lequel la molécule fluoresce avant de blanchir a été calculé pour chacune des traces sur la base d'un fit de la trace par HMM en localisant la dernière transition vers l'état non fluorescent. La distribution de ces temps est présentée figure 6.18a.

Il est donc possible à partir de ces événements de calculer pour chaque temps la proportion de molécules qui restent fluorescentes (courbe de type Kaplan-Meyer de la figure 6.18b). Enfin, et puisque le niveau de fluorescence du plateau fluctue légèrement, pour s'assurer que les différences entre les temps de *bleaching* ne soient pas imputables à des différences d'illumination, nous avons calculé le nombre de photons total détectés entre le début de l'acquisition et le moment où la molécule blanchisse. Il est intéressant de noter que

6.3 Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique

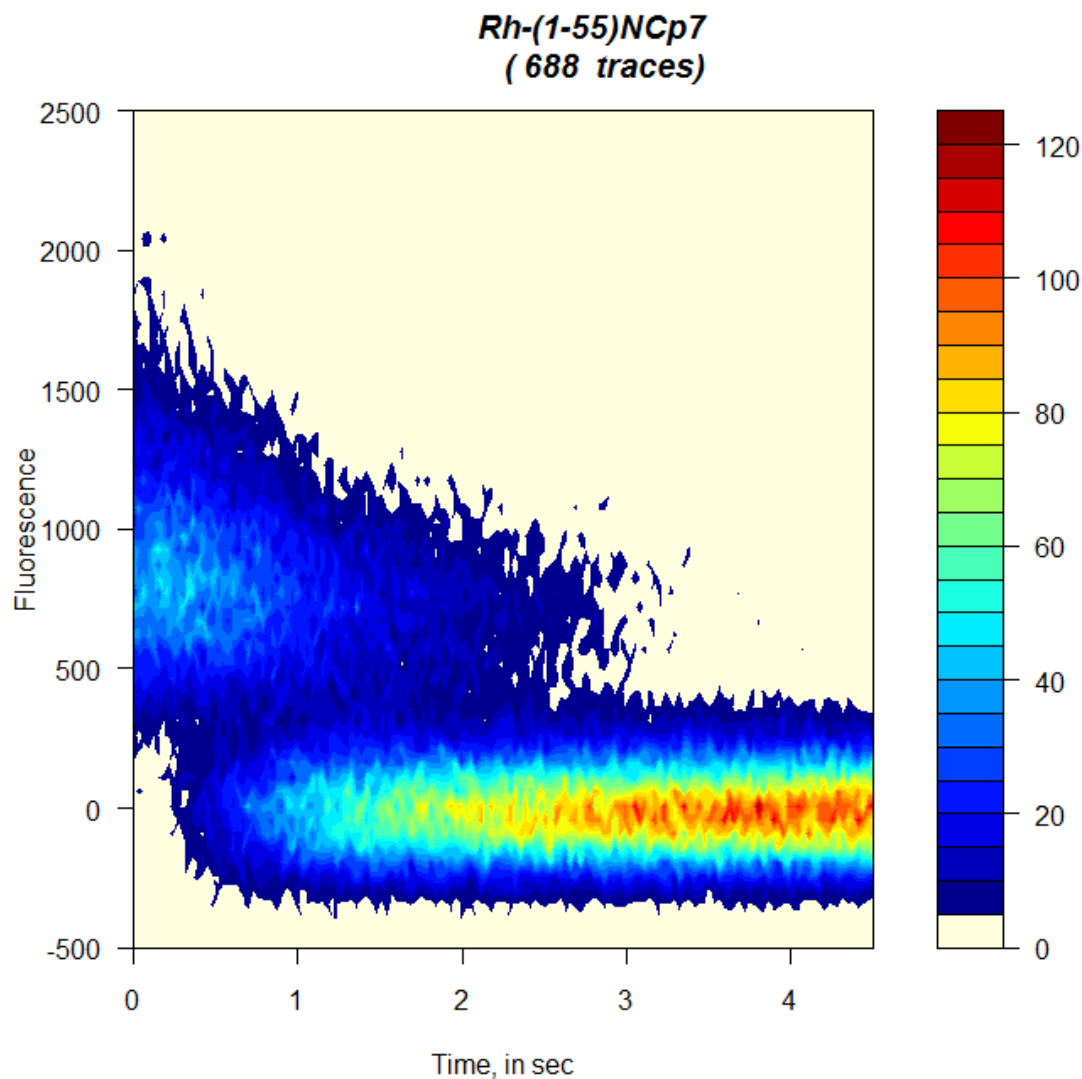


FIGURE 6.17 – Représentation conjointe de la fluorescence en fonction du temps de 688 vésicules encapsulant de la Rh-NCp7.

Chaque pixel représente un nombre de molécules selon le code couleur associé (le nombre de molécules augmente avec la chaleur de la couleur)

la distribution obtenue est largement asymétrique (figure 6.18c).

Toutes ces expériences montrent que :

- l'encapsulation de molécules individualisées au sein de liposomes immobilisés est possible.
- la NCp7 n'interagit pas avec la surface du liposome.
- le TMR qui marque la NCp7 est suffisamment photostable et intense pour être utilisé

6.3 Mesure de la fluorescence de la Rh-NCp7 à l'échelle de la molécule unique

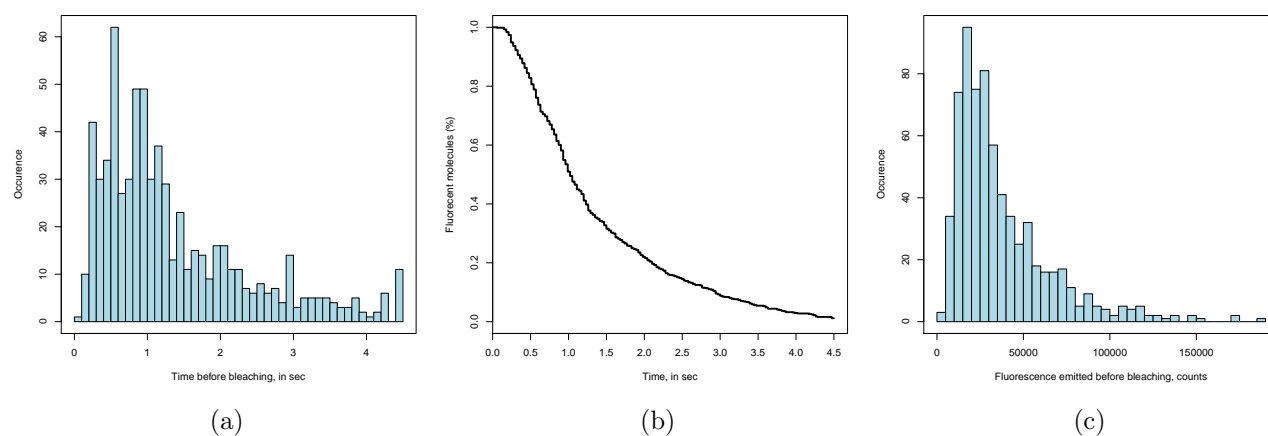


FIGURE 6.18 – Caractéristiques de la fluorescence de 688 molécules de Rh-NCp7 : (a) Distribution des temps individuels avant *bleaching*, (b) fraction de molécules *bleachées* en fonction du temps, (c) distribution des nombres de photons totaux émis avant *bleaching*

comme donneur de fluorescence dans un couple de FRET à l'échelle de la molécule unique.

- le signal de fluorescence du TMR qui marque la NCp7 se comporte de manière biphasique, émettant dans un premier temps une intensité de fluorescence stable pendant quelques secondes avant de photoblanchir et donc de ne plus émettre de fluorescence.

6.4 Mesure de FRET à l'échelle de la molécule unique

Dans un second temps, nous avons réalisé des mesures de FRET à l'échelle de la molécule unique à l'aide d'oligonucléotides doublement marqués.

6.4.1 Duplex d'ADN simplement marqués

Nous avons hybridé les séquences Cy5-AATGCC et Cy3-AAGGCATT de manière à obtenir un duplex doublement marqué. Afin de s'assurer que les deux séquences sont bien hybridées, nous avons réalisé ces expériences dans un tampon à 1M *NaCl* et 10mM *MgCl₂*. En effet, même dans ces conditions, la stabilité du duplex reste limitée et largement dépendante de la température (Figure 6.19).

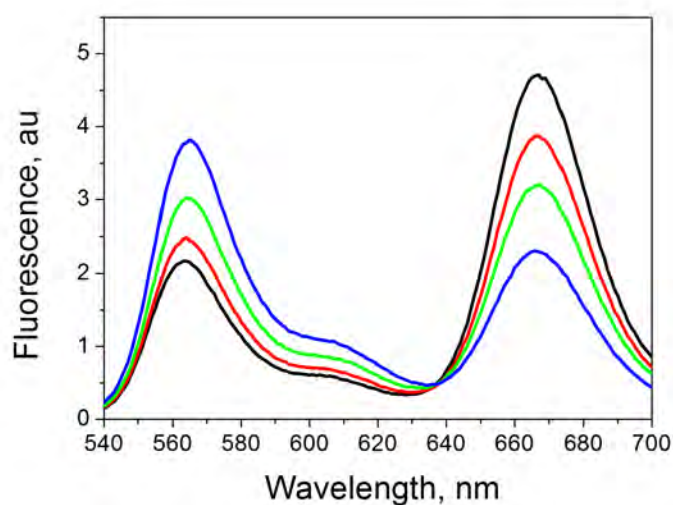


FIGURE 6.19 – Spectre d'émission du duplex Cy5-AATGCC / Cy3-GGCATT en fonction de la température
Noir : 5°C, Rouge : 10°C, Vert : 15°C et Bleu : 20°C

L'excitation par un laser continu à 532 nm permet d'observer simultanément un signal sur la voie du donneur et de l'accepteur. À partir de ces traces de fluorescence, le transfert d'énergie apparent est calculé selon :

$$\langle E_{app}(t) \rangle = \frac{I_A(t)}{I_A(t) + \gamma \times I_D(t)}$$

6.4 Mesure de FRET à l'échelle de la molécule unique

où γ est un facteur qui tient compte des différences de rendements quantiques et des efficacités de détection entre le donneur et l'accepteur [429, 430] ($\gamma = 0.16$)

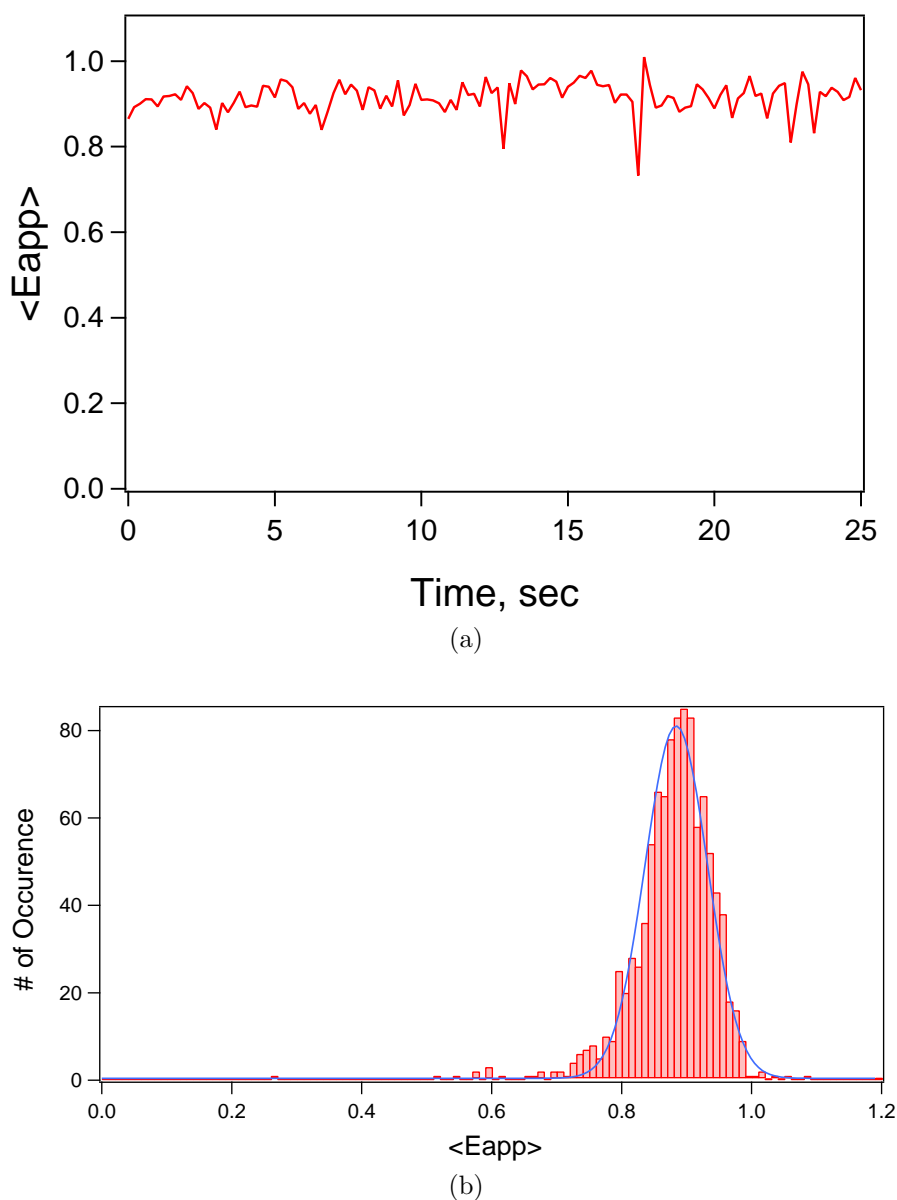


FIGURE 6.20 – Mesures de FRET à l'échelle de la molécule unique.
(a) Trace représentative du FRET au sein du duplex Cy5-AATGCC / Cy3-GGCATT
(b) Distribution des états de FRET de 25 traces

En accord avec la distance attendue entre les sondes au sein de ce duplex, une seule population de FRET apparent est observée avec une valeur moyenne de 0.88. À noter que l'ajout d'un système de piégeage d'oxygène permet d'augmenter significativement la durée de vie des sondes Cy3 et Cy5, autorisant des mesures de l'ordre de plusieurs dizaines de

6.4 Mesure de FRET à l'échelle de la molécule unique

secondes.

6.4.2 PBS doublement marqué

Parallèlement aux expériences réalisées avec un duplex d'ADN doublement marqué, nous avons réalisé des mesures de FRET au sein d'une séquence PBS doublement marquée par du Cy3 et du Cy5. Là encore, il est relativement aisé d'observer simultanément les signaux du donneur et de l'accepteur. Les étapes de photodestruction de l'accepteur sont intéressantes à observer puisque l'extinction de l'accepteur se traduit par une augmentation franche de l'intensité du donneur par disparition du transfert d'énergie (Figure 6.21). D'autre part, le *blinking* de l'accepteur se traduit par une disparition transitoire de la fluorescence de l'accepteur qui n'est pas anti-corrélée avec une exaltation de la fluorescence du donneur (Figure 6.21c). Ces transitions sont donc assignées comme données manquantes lors du calcul du FRET apparent, induisant une trace discontinue de FRET (Figure 6.21d). D'autre part, la valeur de FRET apparent obtenue est en très bon accord avec celle mesurée à l'état stationnaire (Figure 6.21e).

6.4 Mesure de FRET à l'échelle de la molécule unique

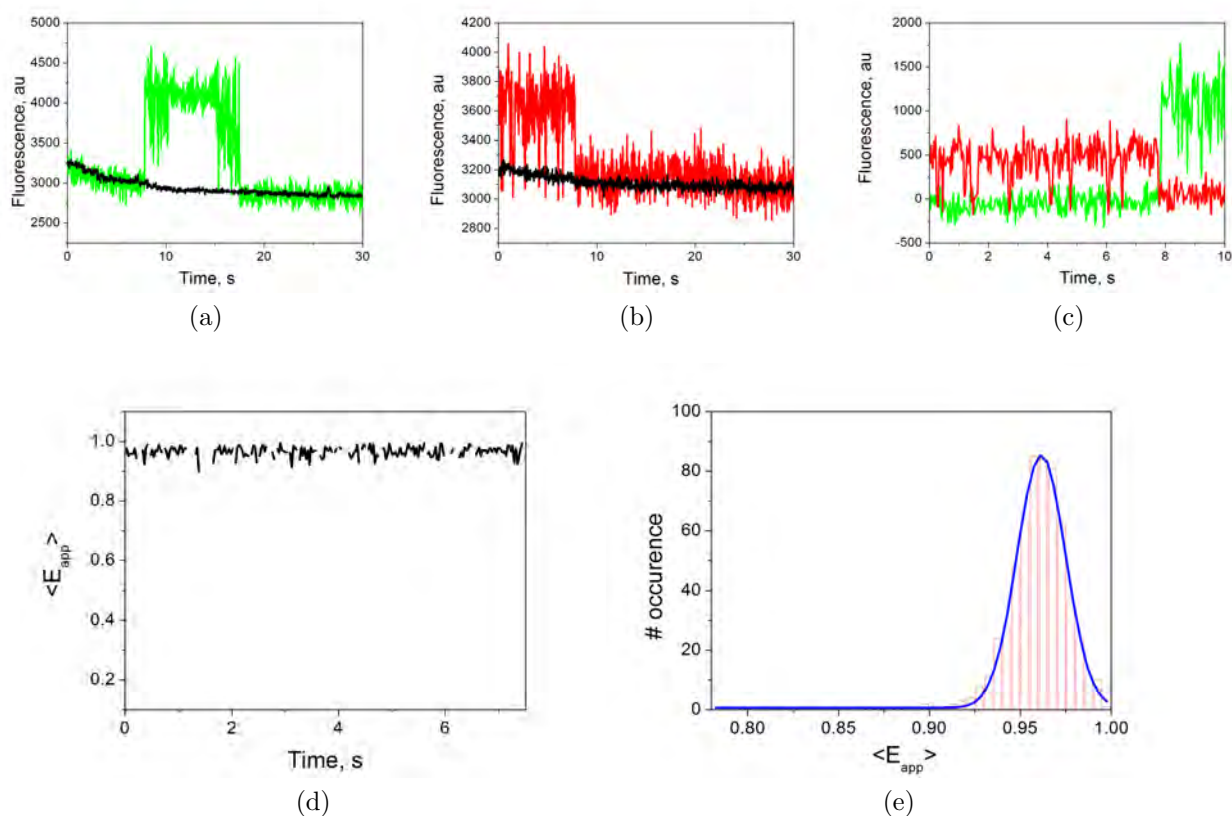


FIGURE 6.21 – Fluorescence de Cy5-PBS-Cy3 à l'échelle de la molécule unique.

(a) Trace brute de fluorescence du Cy3 (vert) et le bruit environnant (noir)

(b) Trace brute de fluorescence du Cy5 (rouge) et le bruit environnant (noir)

(c) Traces corrigées de la fluorescence de Cy5-PBS-Cy3

(d) Évolution temporelle du FRET apparent et (e) sa distribution ajustée selon une distribution gaussienne

Mesures réalisées en présence d'un système de piégeage d'oxygène dans des liposomes de 100 nm de diamètre. Temps d'acquisition : 30 ms

6.5 Interaction de la Rh-NCp7 avec 5'Cy5-AATGCC

Aux échelles de temps accessibles par notre montage expérimental pour les mesures en molécule unique, le FRET au sein de Cy5-PBS-Cy3 apparaît comme stationnaire. Les fluctuations ne sont pas accessibles car elles se déroulent sur des échelles de temps trop rapides et les variations de distances entre les sondes sont probablement limitées. Afin de vérifier qu'il est possible d'observer des transitions de FRET au sein d'un système dynamique, nous nous sommes intéressés à l'étude de la liaison de la Rh-(11-55)NCp7 sur un petit oligonucléotide, le 5'Cy5-AATGCC, constituant un site de liaison affiné pour la (11-55)NCp7 [229].

6.5.1 Mesures d'ensemble

L'interaction de la Rh-(11-55)NCp7 avec le 5'Cy5-AATGCC a d'abord été vérifiée en solution. Les valeurs de FRET ont été utilisées pour mesurer la liaison de concentrations croissantes de 5'Cy5-AATGCC à une concentration fixe de Rh-(11-55)NCp7.

La présence de FRET est clairement démontrée par la diminution progressive de l'émission

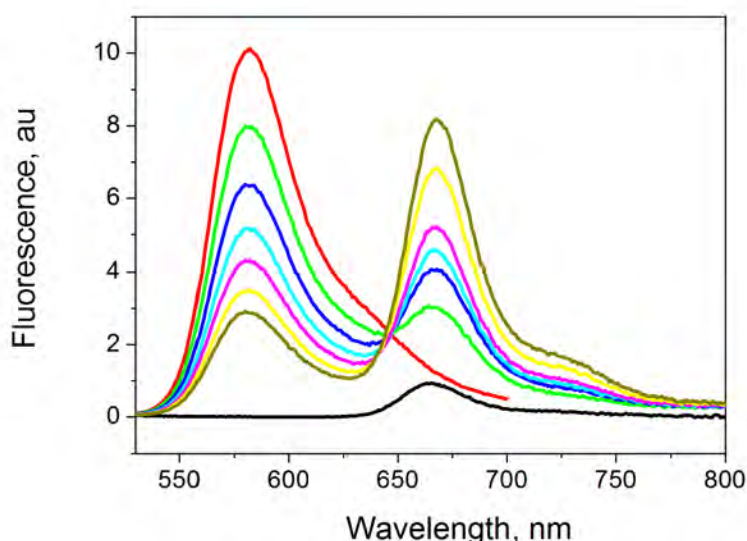


FIGURE 6.22 – Spectre d'émission de fluorescence de la Rh-NCp7 (rouge) en présence de quantités croissantes de 5'Cy5-AATGCC

6.5 Interaction de la Rh-NCp7 avec 5'Cy5-AATGCC

du donneur (TMR) parallèlement à l'augmentation de l'émission de l'accepteur Cy5 (Figure 6.22).

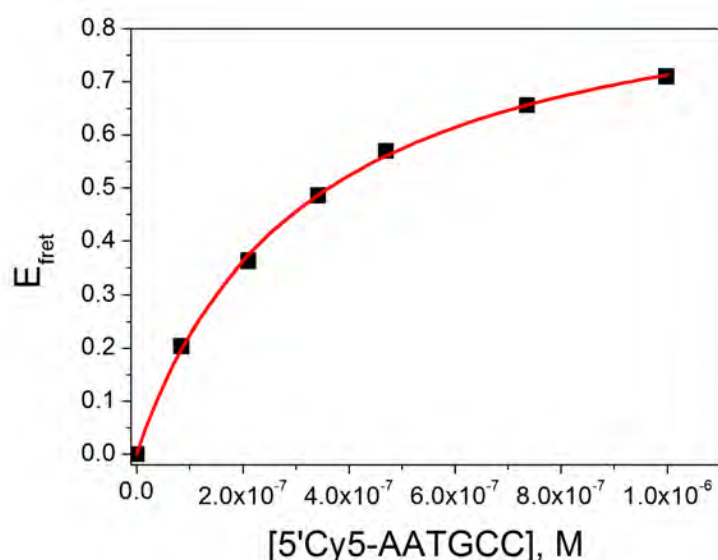


FIGURE 6.23 – Spectre d'émission de fluorescence de la Rh-NCp7 (noir) en présence de quantités croissantes de 5'Cy5-AATGCC

L'efficacité de transfert, calculée par

$$E_{FRET} = 1 - \frac{I_{DA}}{I_D}$$

où I_{DA} et I_D sont les intensités du donneur en absence et en présence de l'accepteur, respectivement, a ensuite été tracée en fonction de la concentration totale de 5'Cy5-AATGCC (Figure 6.23).

L'ajustement par l'équation :

$$E_{FRET} = E_F \frac{(K_d + n [V_{tot}] + [P_{tot}]) - \sqrt{(K_d + n [V_{tot}] + [P_{tot}])^2 - 4n [P_{tot}] [V_{tot}]}}{2 [P_{tot}]} \quad (6.2)$$

donne un $K_{app} = 3.1 \times 10^6 M^{-1}$ très proche de celle mesurée à partir d'expériences de titration par la fluorescence intrinsèque du W37. À concentrations saturantes, la valeur de l'efficacité de FRET final est de 0,94, correspondant à une distance interchromophore

6.5 Interaction de la Rh-NCp7 avec 5'Cy5-AATGCC

d'environ 30 Å, en bon accord avec la distance estimée à partir de la structure par RMN du complexe (12-53)NC/ACGCC [229].

6.5.2 Mesures en molécule unique

L'observation du complexe AATGCC/(11-55)NCp7 encapsulé au sein de liposomes de 100 nm en molécule unique montre des spots fluorescents émettant à la fois dans le canal du donneur et dans celui de l'accepteur. Les traces de fluorescences du donneur et de l'accepteur sont clairement anti-corrélées, comme le montre la figure 6.24. L'évolution temporelle du

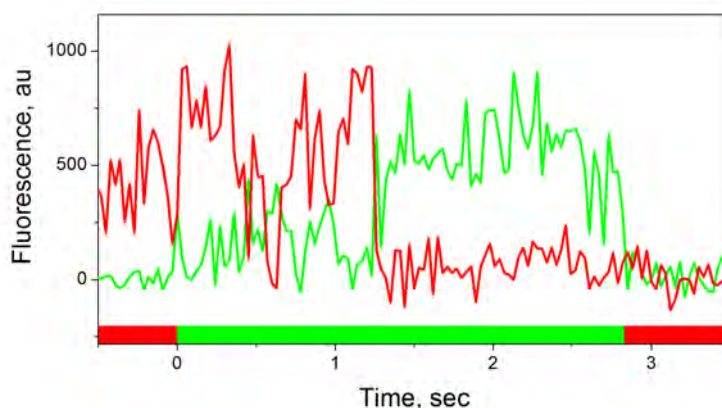


FIGURE 6.24 – Anti-corrélation de la fluorescence de la Rh-(11-55)NCp7 et de 5'Cy5-AATGCC.

L'excitation directe par le laser rouge (635 nm) en début de séquence ($t \leq 0$) montre la présence d'une molécule de 5'Cy5-AATGCC dans le liposome. L'excitation est ensuite basculée sur le laser vert (532 nm), permettant d'observer le transfert d'énergie entre les sondes, jusqu'au photoblanchiment de l'accepteur qui se traduit par une augmentation franche de la fluorescence du donneur. L'excitation directe par le laser rouge en fin de séquence confirme la photodestruction du Cy5.

FRET apparent présente des transitions entre un état où le transfert est fort (proche de 1) et un état où le FRET est pratiquement nul. Ces deux états sont associés à une Rh-NCp7 qui est liée ou libre, respectivement. Alors que l'on s'attend à un système à deux états, l'histogramme des valeurs de FRET apparent (Figure 6.25) montre une distribution continue sur l'ensemble des valeurs possibles de FRET. Ceci montre que le temps d'intégration utilisé pour déterminer la trace de FRET (30 – 60 ms) est supérieur (ou du même ordre de grandeur) à celle des événements générant les fluctuations de FRET. Ainsi, le signal FRET se moyenne au cours du temps d'intégration de la fluorescence. La liaison de la NCp7 sur

6.5 Interaction de la Rh-NCp7 avec 5'Cy5-AATGCC

la séquence AATGCC est donc rapide et la NCp7 se lie et se délie très rapidement.

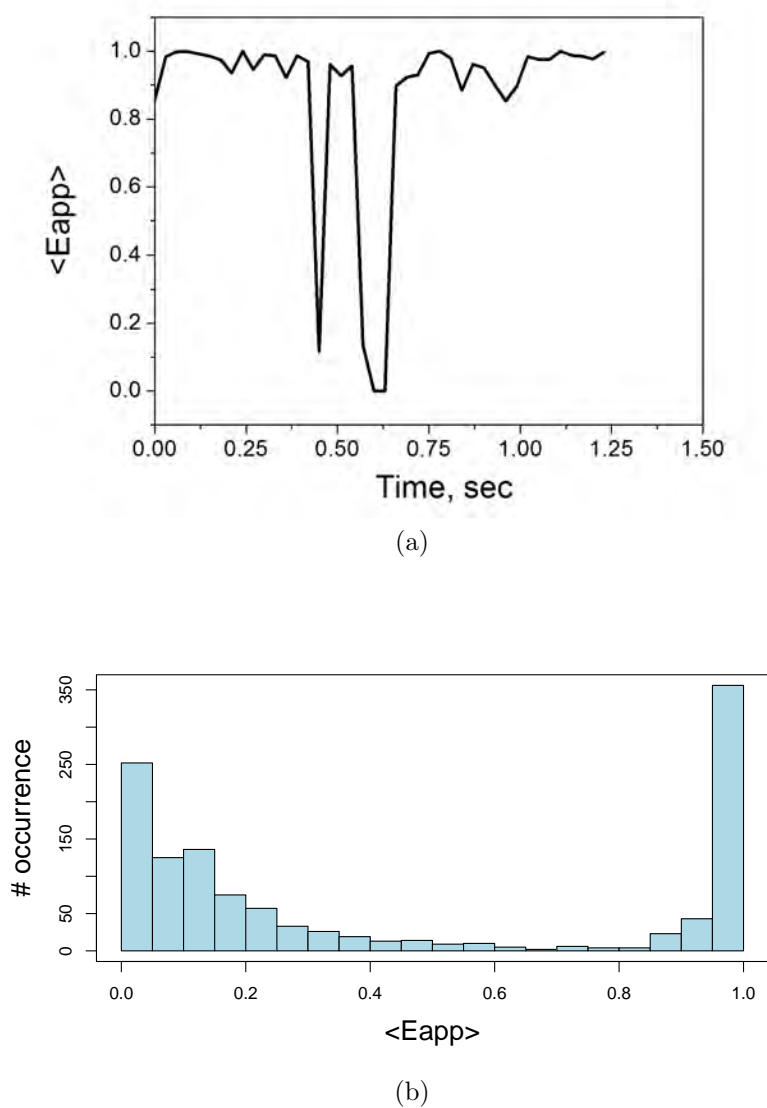


FIGURE 6.25 – (a) Evolution temporelle du FRET apparent correspondant à la trace de la figure 6.24.

(b) Distribution des valeurs de FRET entre Rh-AATGCC et 5'Cy5-AATGCC.

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

Dans le but de caractériser la dynamique de la liaison de la NCp7 avec le (14-39)cTAR, nous avons utilisé de la NCp7 marquée en N_{term} par du TMR (Rh-NCp7) et du (14-39)cTAR marqué à son extrémité 5' terminale par du Cy5.

6.6.1 Mesures d'ensemble

L'ajout progressif de 5'Cy5-(14-39)cTAR à une solution de concentration fixe de NCp7 se traduit par une diminution nette de l'émission de fluorescence du TMR (Figure 6.26). Cette diminution est accompagnée par une légère exaltation de la fluorescence du Cy5.

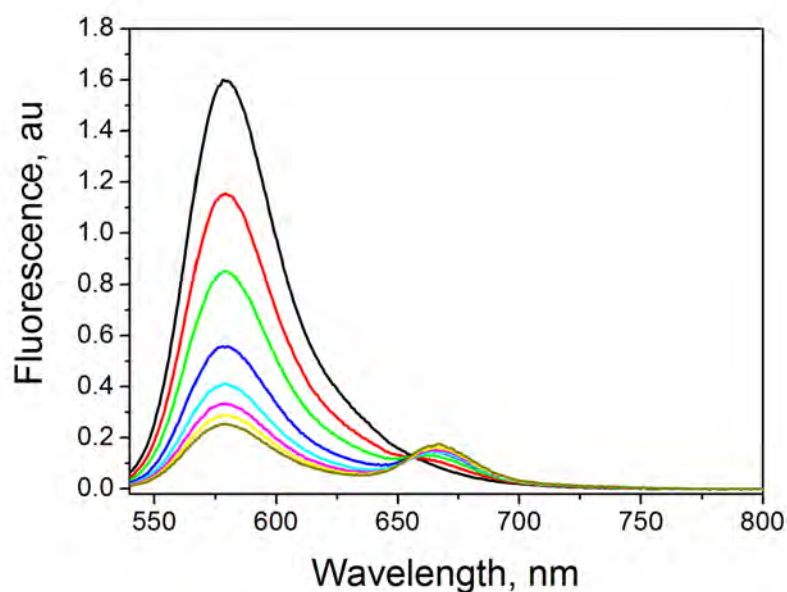


FIGURE 6.26 – Spectre d'émission de fluorescence de la Rh-NCp7 (noir) en présence de quantités croissantes de 5'Cy5-(14-39)cTAR

L'absence d'anti-corrélation linéaire entre les variations de fluorescence du TMR et du Cy5 démontre la présence d'un phénomène, autre que le FRET, qui induit un *quenching* important du TMR marquant la NCp7. L'étude attentive des spectres d'absorption correspondants montre que le pic d'absorption du Cy5 est déplacé vers le bleu à mesure que la concentration de 5'Cy5-(14-39)cTAR augmente (Figure 6.27). L'absorption du Cy5 est également systématiquement plus faible que celle obtenue pour une concentration similaire de

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

5'Cy5-(14-39)cTAR en absence de Rh-NCp7. Un *shift* est également observé au niveau du pic correspondant au TMR. Ces observations montrent sans ambiguïté l'existence d'une interaction sonde-sonde au sein du complexe, suggérant que la NCp7 se lie préférentiellement sur la partie basse du (14-39)cTAR. Ceci est conforté par une diminution très significative de l'extinction de fluorescence du TMR lorsque l'expérience est répétée avec un (14-39)cTAR non-marqué (Figure 6.28).

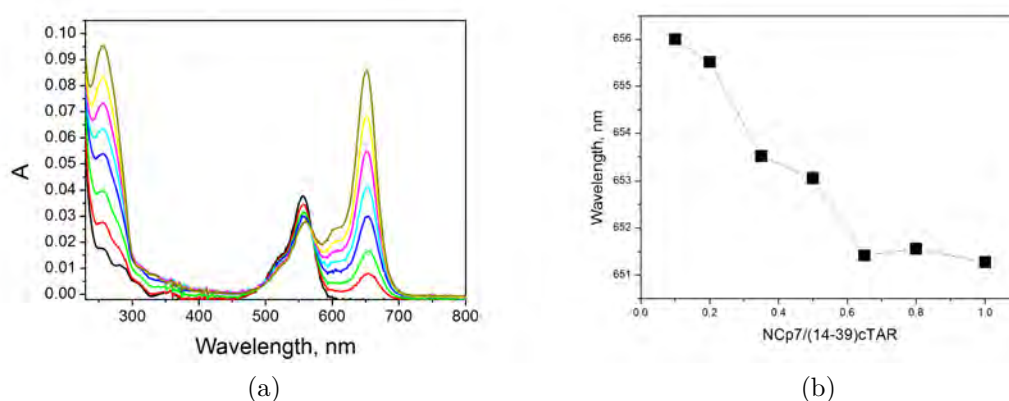


FIGURE 6.27 – (a) Spectre d'absorption UV-visible de la Rh-NCp7 en présence de quantités croissantes de 5'Cy5(14-39)cTAR.

(b) La position du maximum d'absorption du Cy5 se déplace vers le bleu à mesure que le rapport 5'Cy5-(14-39)cTAR/Rh-NCp7 augmente

Les mesures de fluorescence résolue en temps montrent que le déclin de la fluorescence de la Rh-NCp7 est largement modifié en lorsque la protéine est liée au (14-39)cTAR (Figure 6.1). L'existence d'un *quenching* du TMR se traduit par l'apparition d'espèces noires qui représentent près de 85% des conformations dans le cas de la liaison avec le 5'Cy5-(14-39)cTAR. Ces conformations ne représentent plus que 30% lorsque le Cy5 est absent, ce qui confirme bien l'existence d'une interaction sonde-sonde. Des conclusions similaires sont obtenues avec la Rh-NC(11-55), ce qui laisse à penser que la reconnaissance d'un site de liaison plus afin sur le bas de la tige du (14-39)cTAR dépend des doigts de zinc. De plus, la répartition entre les espèces noires et les espèces associées au temps de vie le plus court diffère. Le TMR apparaît plus *quenché* lorsqu'il est lié en N_{ter} de la protéine tronquée. Ceci pourrait s'expliquer par la perte de degrés de liberté de la sonde vis-à-vis de l'ODN consécutive à la délétion de la partie N_{ter} flexible.

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

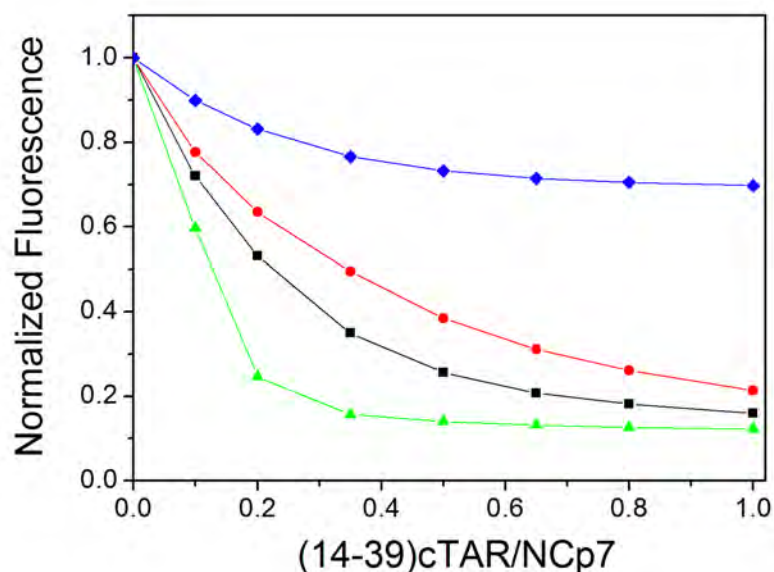


FIGURE 6.28 – Intensité de la fluorescence de la Rh-NCp7 en présence de quantités croissantes de 5'Cy5-(14-39)cTAR.

Les mesures sont effectuées dans un tampon 25mM TRIS HCl, pH 7,5, 0,2 mM de $MgCl_2$ et 30 mM (triangles vert) ou 100 mM de $NaCl$ (en absence (carrés noir) ou en présence (ronds rouge) de 1 mM d'EDTA). Pour comparaison, la mesure dans le même tampon à 100 mM de $NaCl$ avec un (14-39)cTAR non marqué (losanges bleu)

6.6.2 Mesures en molécule unique

Contrairement aux liposomes enfermant uniquement de la Rh-NCp7, la présence d'un ODN se traduit par un signal de la TMR fluctuant fortement jusqu'à ce que la sonde photoblanchisse. Le signal de la TMR varie entre un état non fluorescent et un état dont le niveau de fluorescence correspond à la valeur du plateau observé avec la Rh-NCp7 seule. Conséquence de l'importance du *quenching* du TMR lors de l'interaction de la Rh-NCp7 avec 5'Cy5-cTAR, aucun signal n'est détecté sur la voie de l'accepteur lorsque les liposomes sont excités par le laser à 532 nm. L'excitation directe du Cy5 avec le laser à 635 nm confirme pourtant la présence de l'oligonucléotide dans les liposomes. Dans ces conditions, il n'est donc pas possible de mesurer du FRET entre la Rh-NCp7 et le 5'Cy5-(14-39)cTAR.

Nous avons donc fait l'hypothèse que les états fluorescents de la Rh-NCp7 correspondent

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

| | | α_0 | τ_1 | α_1 | τ_2 | α_2 | τ_3 | α_3 | τ_4 | α_4 | τ | Rm |
|------------|-------------------|------------|----------|------------|----------|------------|----------|------------|----------|------------|--------|------|
| 1-55 NCp7 | - | | | | 0.97 | 0.14 | 2.72 | 0.86 | | | 2.47 | |
| 1-55 NCp7 | 5'Cy5 (14-39)cTAR | | 0.10 | (0.47) | 0.62 | (0.22) | 2.21 | (0.25) | 3.96 | (0.06) | 0.96 | 16 |
| | | 0.84 | | 0.08 | | 0.04 | | 0.04 | | 0.01 | | |
| 1-55 NCp7 | (14-39)cTAR | | 0.17 | (0.35) | 0.67 | (0.25) | 2.33 | (0.28) | 4.65 | (0.12) | 1.43 | 2.5 |
| | | 0.31 | | 0.24 | | 0.17 | | 0.19 | | 0.08 | | |
| 11-55 NCp7 | - | | | | 0.99 | 0.16 | 2.66 | 0.84 | | | 2.38 | |
| 11-55 NCp7 | 5'Cy5 (14-39)cTAR | | 0.13 | (0.27) | 0.61 | (0.27) | 2.44 | (0.39) | 4.32 | (0.07) | 1.44 | 14.8 |
| | | 0.89 | | 0.03 | | 0.03 | | 0.04 | | 0.01 | | |
| 11-55 NCp7 | (14-39)cTAR | | 0.14 | 0.14 | 0.58 | 0.32 | 2.15 | 0.36 | 4.41 | 0.18 | 1.76 | 2.5 |
| | | 0.46 | | 0.08 | | 0.17 | | 0.20 | | 0.09 | | |

TABLE 6.1 – Paramètres des déclin de fluorescence résolus en temps de Rh-NC libres ou complexées avec 5'Cy5-(14-39)cTAR.

Les temps de vie τ_i sont exprimé en *ns*. Les espèces noires, caractérisées par une amplitude α_0 correspondent à l'ensemble des conformations pour lesquelles l'émission de fluorescence est sous la résolution du montage expérimental (~ 40 ps). Rm correspond au ratio des intensités d'émission de fluorescence

aux états pour lesquels la NCp7 est libre dans le liposome, par opposition aux états non-fluorescents associés à la liaison de la NCp7 sur l'oligonucléotide. Les traces de fluorescence des liposomes présentant un signal de fluorescence de Cy5 avec le laser à 635 nm et un signal de TMR sont sélectionnés. Ces traces montrent de multiples transitions entre les états fluorescents et non-fluorescents (Figure 6.29).

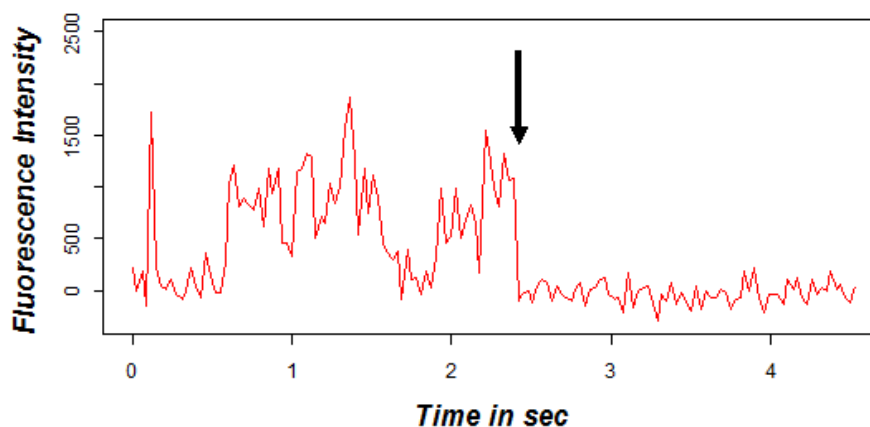


FIGURE 6.29 – Évolution temporelle de l'intensité de fluorescence d'une Rh-(11-55)NCp7 en présence de 5'Cy5-(14-39)cTAR.

Le signal de fluorescence évolue entre un état non-fluorescent pour lequel la protéine est supposée être liée et un état fluorescent maximal correspondant au signal de la Rh-NCp7 seule enregistré dans les mêmes conditions. La flèche correspond au moment où la TMR est photoblanchie

En analysant un grand nombre de trace de fluorescence, il est ainsi possible de re-

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

constituer la distribution des états de fluorescence (Figure 6.30). L'ajustement de cette distribution à l'aide d'un modèle à deux populations distribuées normalement permet de quantifier les proportions des espèces liées (37%) et libres et d'estimer la constante d'affinité. On trouve $K_{app} = 1.5 \pm 0.2 \times 10^6 M^{-1}.s^{-1}$, une valeur presque 10 fois plus faible que celle déjà observée [269].

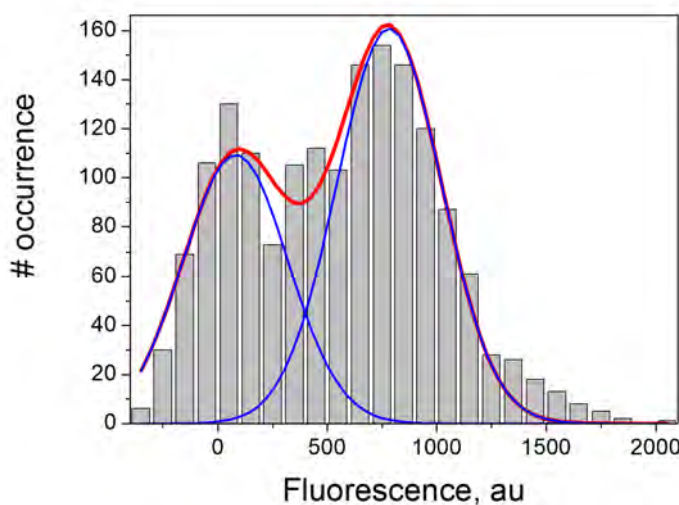


FIGURE 6.30 – Distribution des intensités de fluorescence d'une Rh-(11-55)NCp7 en présence de 5'Cy5-(14-39)cTAR.

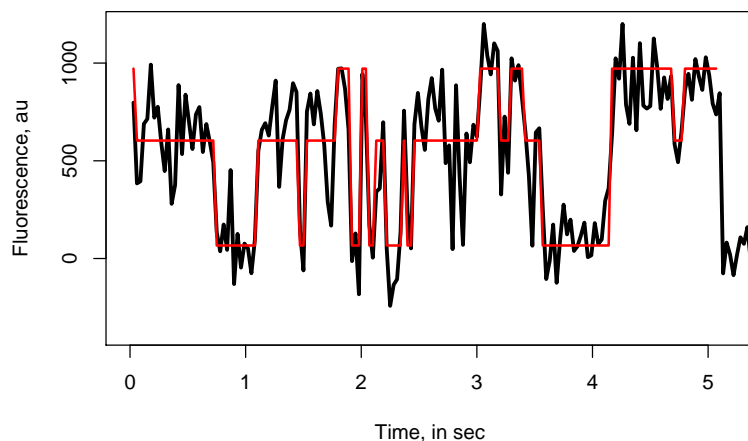
La distribution est ajustée spontanément à l'aide d'un modèle à deux pics gaussiens.

Une analyse plus poussée des traces individuelles de fluorescence montre que l'ajustement par un modèle de HMM est optimal pour un système à trois états (Figure 6.31a). Comme le montre le nouvel ajustement de la distribution des intensités de fluorescence, les états fluorescents correspondent à deux populations distinctes (Figure 6.31b).

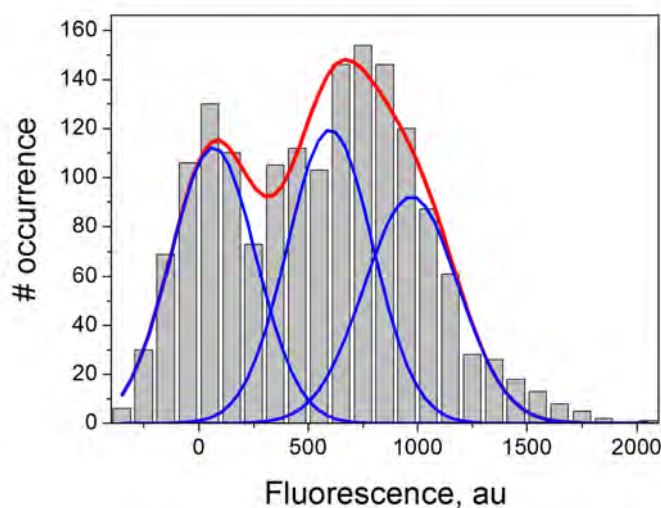
Calcul des temps de résidences

L'ajustement des données par un modèle HMM permet de calculer les temps de résidences moyens du système dans chacun des états (Figure 6.32). Il apparaît évident que les temps de résidences ne traduisent pas directement l'état de liaison de la protéine mais qu'ils incluent des transitions additionnelles d'origine photophysique.

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR



(a)



(b)

FIGURE 6.31 – (a) Évolution temporelle de la fluorescence de Rh-(11-55)NCp7 en présence de 5'Cy5-(14-39)cTAR (noir) ajustée selon un modèle à trois états par HMM (rouge). (b) Distribution des intensités de fluorescence ajusté selon un modèle à trois états

Afin de voir s'il nous était possible de caractériser l'aspect photophysique du système, nous avons effectué des mesures de FCS de la protéine seule et de la protéine complexée à différents oligonucléotides marqués au Cy5 (Tableau 6.2). Les mesures de FCS montrent que le nombre apparent de molécules diffusant au sein du volume focal est largement diminué en présence des ODN. Une partie significative des complexes se trouvent donc sous

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

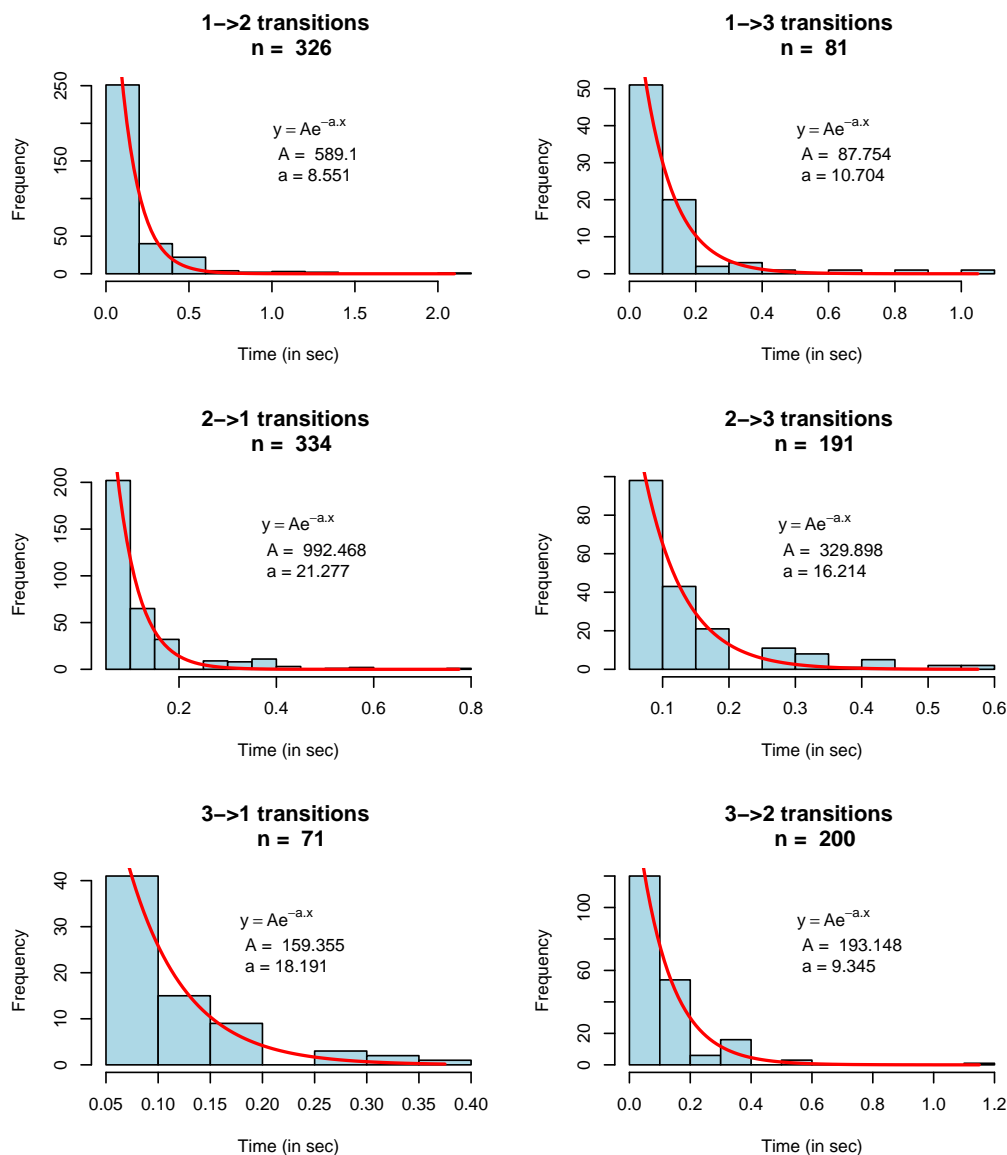


FIGURE 6.32 – Histogrammes des temps de résidences.

Ces histogrammes permettent de calculer le temps de résidence moyen (en sec^{-1}) dans un état donné du système avant transition vers un autre état

forme non-fluorescente lorsque la protéine est liée. Les temps de diffusions observés en présence d'ODN sont plus élevés que les temps correspondants aux acides nucléiques seuls, ce qui montre qu'une partie des complexes émet également de la fluorescence. Le rapport des courbes d'autocorrélation des complexes NCp7/(14-39)cTAR marqué et non-marqué par un accepteur de fluorescence permet de déterminer le temps de relaxation moyen des fluctuations dues à l'interaction entre le TMR et la Cy5 (Figure 6.33b). Des premiers résultats

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

| | N_{app} | τ_{Diff} | f_t | τ_{fluct} | $N_{relatif}$ | I_{norm} |
|----------------------------|-----------|---------------|-------|----------------|---------------|------------|
| NCp7 | 89 | 0.156 | - | - | 1 | 1 |
| NCp7/5'Cy5-(14-39)cTAR | 17 | 0.123 | 0.37 | 1.4 | 0.19 | 0.15 |
| NCp7/(14-39)cTAR | 59 | 0.128 | 0.40 | 1.0 | 0.66 | 0.77 |
| NCp7/5'Cy5-SL3 | 16 | 0.103 | 0.35 | 1.0 | 0.18 | 0.17 |
| NCp7/5'Cy5- ΔP PBS | 20 | 0.099 | 0.47 | 2.0 | 0.22 | 0.15 |

TABLE 6.2 – Paramètres de FCS de Rh-NCp7 libre ou complexée à différents ODN. Les concentrations de Rh-NCp7 ou des complexes équimolaires Rh-NCp7/ODN sont de 350nM. Les mesures sont enregistrées dans un tampon TRIS HCl 25mM, 100mM NaCl, 0.2mM MgCl₂.

N_{app} : nombre moyen de molécules apparentes dans le volume focal

τ_{Diff} : temps de diffusion

f_t fraction correspondant aux molécules dont le signal change au cours de la traversée du volume focal

τ_{fluct} temps des fluctuations additionelles

$N_{relatif}$: nombre moyen de molécules apparentes normalisé par rapport à Rh-NCp7 seule

I_{norm} intensité de l'émission de fluorescence mesurée en stationnaire normalisé par rapport à la fluorescence de Rh-NCp7

à confirmer montrent que la courbe $G^*(\tau)$ suit un déclin selon une courbe exponentielle étirée, selon :

$$\langle \tau_r \rangle = \frac{\tau_r}{\beta} \Gamma \left(\frac{1}{\beta} \right) \quad (6.3)$$

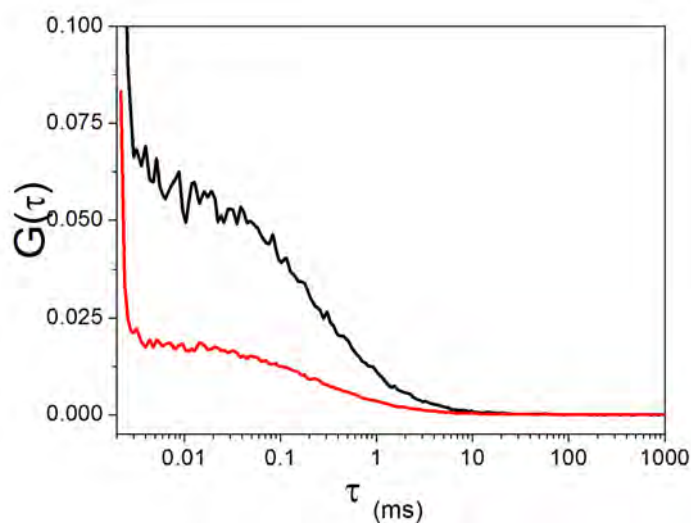
La faible valeur du paramètre β montre que la distribution des temps de relaxation est large, laissant supposer que la dynamique de l'interaction entre les sondes est un phénomène complexe. Le temps de relaxation moyen vaut environ 125ms.

Pour en revenir au calcul de temps de résidence, et puisque l'hypothèse initiale selon laquelle les formes liées de la NCp7 ne sont pas fluorescentes, contrairement aux formes libres, s'avère inexacte, nous avons estimé les temps de résidence de la Rh-(11-55)NCp7 sur le 5'Cy5-(14-39)cTAR en faisant l'hypothèse que la population associée à la fluorescence intermédiaire correspond à l'état lié. Il vient que $k_{on} \sim 22.0 \pm 0.8 \times 10^{-7} M^{-1}.s^{-1}$ et $k_{off} \sim 2.0 \pm 0.5 s^{-1}$. Ceci permet de calculer la constante d'affinité selon

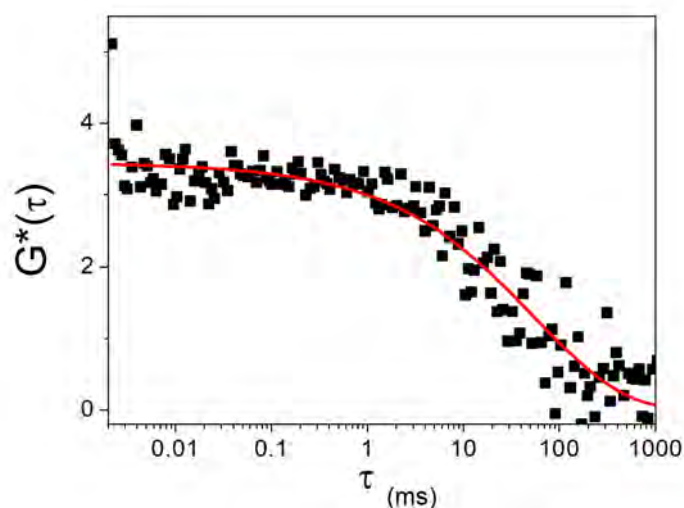
$$K_{app} = \frac{k_{on}}{k_{off}}$$

On trouve $K_{app} = 1.1 \pm 0.4 \times 10^7 M^{-1}$, une valeur beaucoup plus proche de celle déjà décrite.

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR



(a)



(b)

FIGURE 6.33 – (a) Courbe de FCS des complexes Rh-NCp7/5'Cy5(14-39)cTAR (noir) et Rh-NCp7/(14-39)cTAR (rouge).

(b) Ratio entre les deux courbes présentées en (a). La courbe rouge correspond au fit à l'aide de l'équation 6.3 dans laquelle $\beta = 0.48$, $\tau_r = 58$ ms, $A=0$ et $B=3.4$

Ces résultats illustrent la difficulté de dissocier, en l'absence de signal de FRET, l'origine des fluctuations de la fluorescence de la rhodamine. Si la liaison de la NCp7 sur l'oligonucléotide constitue probablement l'origine principale de ces fluctuations, la NCp7 liée n'est pas nécessairement éteinte et sa fluorescence peut également fluctuer. Ceci pourrait traduire l'existence d'autres sites de liaisons plus distants de la sonde en 5' de l'ODN, ou encore

6.6 Interaction de la Rh-NCp7 avec (14-39)cTAR

l'existence de modes de liaison différents. Dans tous les cas, et même si il n'est possible de caractériser que de manière approximative les constantes d'association et de dissociation en absence de FRET, la liaison de la NCp7 sur les oligonucléotides apparait, dans nos conditions de mesures, transitoire et en échange rapide entre les formes liées et libres.

Conclusions et perspectives

7.1 Conclusions et perspectives

La NCp7 fait partie de la famille des protéines chaperonnes d'acides nucléiques. Si l'existence et la justification biologique des protéines chaperonnes d'ARN trouve une explication intellectuellement satisfaisante, leurs mécanismes d'action s'avèrent relativement complexes et leur connaissance reste parcellaire. Nous avons essayé au cours de cette thèse, par des méthodes de fluorescence, d'apporter quelques réponses quant à la caractérisation de ces mécanismes et de leurs dynamiques.

Dans un premier temps, nous avons exploité les propriétés de la FCS qui constitue une méthode adaptée à l'étude des interactions entre molécules faiblement concentrées. Nous avons ainsi pu montrer l'existence d'homodimères de courtes séquences d'acides nucléiques induits par NCp7 et caractérisé la dynamique de la sonde liée aux ODN. Nous avons également pu montrer la validité d'une approche par corrélation croisée. Enfin, nous avons développé une méthode de FCS à deux foyers (2fFCS) à l'aide d'un interféromètre de Michelson. La maîtrise de la technique de la FCS apporte un avantage double : tout d'abord, elle donne accès à la dynamique des molécules sur une échelle de temps assez large (de la microseconde à la dizaine de seconde), ce qui la rend complémentaire aux résolutions temporelles obtenues par l'approche en molécule unique. Ensuite, et c'est sans doute un avantage majeur de cette méthode, il sera possible de passer à l'étude de biomolécules au sein de systèmes plus complexes, et en particulier dans un environnement cellulaire, notamment grâce à l'excitation biphotonique, peu destructrice pour les cellules.

Dans un second temps, nous avons caractérisé le mécanisme de l'hybridation de TAR avec cTAR en utilisant des signaux de FRET. Nous avons montré que la séquence cTAR s'hybride à TAR, en présence de NCp7, *via* sa partie basse, et nous en avons caractérisé les vitesses d'hybridation. L'étude détaillée de la réaction d'hybridation de la séquence cTAR avec les séquences TAR ou dTAR a permis de disséquer le mécanisme de cette réaction et de mieux comprendre le rôle de l'activité chaperonne de NCp7 dans cette réaction. Depuis la publication de ces travaux, la littérature concernant le mécanisme de la réaction d'hybridation cTAR/dTAR et l'action de NCp7 sur celle-ci s'est lar-

7.1 Conclusions et perspectives

gement étoffée. Des études détaillées de l'hybridation de mini-cTAR avec miniTAR, de cTAR avec TAR dans différentes conditions et par différentes techniques ont été rapportées [431, 432, 433, 434, 435, 282, 436, 437]. Il apparait maintenant clairement que plusieurs sites d'initiation de l'hybridation de cTAR avec TAR existent. Alors qu'en absence de NCp7, l'hybridation se produit majoritairement *via* une interaction boucle-boucle impliquant 17 nt, l'ajout de NCp7 induit un *switch* mécanistique et la réaction se produit très majoritairement par le bas des tiges. L'étude détaillée de cette réaction d'hybridation a permis d'utiliser ce système comme modèle d'étude de l'activité chaperonne d'autres protéines (NCp10 [385], Tat [438] ou la protéine du core de HCV [439]). Ce travail a également apporté un éclairage important pour la compréhension du mécanisme d'hybridation d'autres systèmes comme celui de la séquence PBS impliquée dans le second saut de brin [281]. Enfin, ce système permet de tester l'activité de molécules inhibant potentiellement la NCp7. Dans la suite de ces travaux, nous sommes impliqués dans un projet couplant microfluidique et fluorescence résolue en temps visant à caractériser les états intermédiaires de la réaction d'hybridation de cTAR avec TAR afin de caractériser plus précisément l'action de la NCp7 (et de Tat) (projet FLIM-fluidique programme interdisciplinaire-interface PCB - CNRS 2009)

Nous nous sommes également intéressés aux propriétés d'une sonde structurale, la 2-Aminopurine. L'étude des déclins de la fluorescence de la 2-Ap nous ont permis de caractériser les dynamiques des acides nucléiques au sein des complexes ADN/protéine. La résolution en pico-nanoseconde est particulièrement adaptée pour sonder les dynamiques locales des acides nucléiques (mobilité des bases et mouvements segmentaux). Nous avons ainsi pu montrer que l'action de déstabilisation de la NCp7 résulte principalement, au niveau moléculaire, de l'immobilisation locale de la séquence d'ADN sur laquelle la protéine se lie. Ces modifications de mobilité, et particulièrement du *stacking* des bases, modifient largement les signaux de fluorescence émis par la 2-Ap. Il nous a ainsi été possible de caractériser des affinités différentielles, de manière site-spécifique, au sein de courts ODN capables de lier deux NCp7. Nous avons également pu montrer que les modifications des dynamiques de la boucle de PBS en réponse à la liaison de la NCp7 sont largement responsables du changement de mécanisme réactionnel au cours de la réaction d'hybridation

7.1 Conclusions et perspectives

de (+)PBS avec (-)PBS. Enfin, nous avons étudié de manière site-spécifique la liaison de la NCp7 sur la séquence cTAR, ce qui nous a permis de mettre en évidence deux sites de liaisons préférentiels dont la localisation permet d'affiner le mécanisme de déstabilisation de cette séquence par la NCp7.

On peut bien sûr imaginer décliner l'approche qui consiste à caractériser la liaison de manière spécifique à de nombreuses autres séquences du génome du VIH-1 comme DIS, SL3 (en cours), etc ... Récemment, il a été montré que la NCp7 est capable de plier les hélices double-brin d'ADN [440]. Les auteurs ont proposé un mécanisme impliquant une "bulle" de fusion du double brin (déstabilisation) couplée à une attraction inter-brin apportant l'énergie nécessaire au rapprochement des extrémités. Un mécanisme purement électrostatique, apparenté à celui produit par des petits cations mobiles, pourrait également expliquer ces observations [441]. La fluorescence de la 2-Ap pourrait permettre de caractériser les mécanismes par lesquels la NCp7 produit ces effets qui interviennent probablement dans la compaction de l'ADN viral. De plus, nous avons observé que la majorité des conformations de la 2-Ap au sein des acides nucléiques correspondent à des formes fortement *stackées*, dont l'émission de fluorescence associée est rapide (quelques pico-secondes). Du fait de la résolution temporelle de notre approche, la caractérisation de ces émissions nous était inaccessible. La spectroscopie de fluorescence ultra-rapide (femtoseconde) permet d'accéder à ces gammes temporelles de désexcitation et donc d'apporter des informations précieuses sur les dynamiques de ces événements. Un projet de recherche a été conçu dans cette optique, en collaboration avec l'équipe du Pr S. Haacke, spécialiste de la spectroscopie ultra-rapide. Ce projet a reçu un financement ANR (Projet FemtoStack (ANR 2010) : Nouveaux outils de spectroscopie ultra-rapide et de stimulation pour l'étude de la dynamique structurale de complexes protéines chaperonnes/ADN). La spectroscopie de fluorescence de la 2-Ap constitue également un complément extrêmement intéressant des données obtenues par cristallographie ou par RMN et s'avère un outil intéressant pour valider les approches par modélisation moléculaire. Notre expertise dans le domaine de la fluorescence de la 2-Ap pourra ainsi s'inscrire dans des projets multidisciplinaires visant à caractériser les aspects structuraux et dynamiques de complexes nucléiques ou nucléo-protéiques.

7.1 Conclusions et perspectives

Enfin, nous avons développé et validé une approche par spectroscopie de fluorescence à l'échelle de la molécule unique. Cette méthode permet de réaliser des mesures de FRET à l'échelle d'un couple de donneur et d'accepteur de fluorescence, permettant de caractériser la dynamique des interactions et les changements conformationnels qui en résultent. À très court terme, cette méthode sera utilisée pour compléter les premières observations réalisées avec la NCp7 afin de caractériser la dynamique de la liaison de la NCp7. La méthode d'immobilisation des molécules à l'aide de liposomes permettra également d'étudier cette interaction dans des conditions plus "peuplées" en protéine, notamment en ajoutant des quantités significatives de NCp7 non-marquée afin d'étudier l'influence du *molecular crowding* sur la liaison. Outre les interactions NCp7/acides nucléiques, cette méthode permettra également l'étude de l'influence de la NCp7 sur l'activité de la RT, dans le but d'expliquer à l'échelle moléculaire les observations faites notamment dans Grohmann & al. [332]. Ces travaux sont déjà en cours dans le cadre d'un projet ANRS 2008 intitulé : "Caractérisation de l'interaction entre la protéine de la nucléocapside et la rétrotranscriptase du VIH-1 au cours de l'initiation de la rétrotranscription et du premier saut de brin : étude en molécules uniques". L'objectif de ces travaux est de caractériser à l'échelle de la molécule unique les effets de la NCp7 sur l'orientation de la RT sur son substrat [442] puisqu'il semblerait que cette orientation définisse son activité polymérase ou RNase H. L'objectif est également d'évaluer l'action de la NCp7 sur les pauses de la RT au cours de la phase d'initiation de la reverse transcription. L'expression et la purification d'une RT marquée à déjà été réalisée (travaux réalisés dans le laboratoire de T. Restlé (Lübeck), ne figurant pas dans ce manuscrit). Nous avons également obtenus des résultats préliminaires qui montrent que la NCp7 favorise l'orientation polymérase de la RT sur un substrat d'ADN.

Bien évidemment, l'approche en molécule unique peut être utilisée pour l'étude d'autres complexes moléculaires. Des travaux dans le cadre d'une collaboration sont déjà engagés, dans la lignée des travaux de la publication IX, pour caractériser l'interaction de récepteurs nucléaires à l'acide rétinoïque sur leur ADN cible, en absence et en présence de cofacteurs. De manière intéressante, ces complexes de taille bien plus grande que les complexes ODN/NCp7 se montrent plus simples à étudier, notamment par le fait de l'absence d'interaction directe sonde-sonde ou de tout autre perturbation résultant d'une faible distance

7.1 Conclusions et perspectives

inter-sonde, ainsi que par des dynamiques “ralenties” du fait de la taille des complexes. Ceci permet d’espérer raisonnablement pouvoir travailler au sein de complexes beaucoup plus intégrés pour étudier les propriétés chaperonnes de NCp7. On peut par exemple imaginer caractériser l’effet de la NCp7 sur la dynamique et le repliement de la partie 5’UTR complète du génome du VIH ou encore s’intéresser à l’effet de la NCp7 sur l’aspect dynamique du *DNA flap* qui joue un rôle critique dans l’import nucléaire du génome du VIH-1. Ces travaux pourraient apporter des informations précieuses quant à la compréhension de l’activité chaperonne de la NCp7.

Bibliographie

- [1] Jacob F. La logique du vivant. Tel ed. Gallimard, editor ; 1976.
- [2] Monod J. Le hasard et la nécessité - Essai sur la philosophie naturelle de la biologie moderne. du Seuil E, editor ; 1970.
- [3] Klatzmann D, Champagne E, Chamaret S, Gruest J, Guetard D, Hercend T, et al. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature*. 1984;312(5996) :767–768.
- [4] Barré-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 1983 May ;220(4599) :868–871.
- [5] Clavel F, Guyader M, Guétard D, Sallé M, Montagnier L, Alizon M. Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature*. 1986;324(6098) :691–695. Available from : <http://dx.doi.org/10.1038/324691a0>.
- [6] Murphy E, Korber B, Georges-Courbot MC, You B, Pinter A, Cook D, et al. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the central African Republic. *AIDS Res Hum Retroviruses*. 1993 Oct ;9(10) :997–1006.
- [7] Goldschmidt V. La rétrotranscription de HIV-1 : étude du complexe d'initiation et des mécanismes de résistance aux inhibiteurs nucléosidiques. Université Louis Pasteur, Strasbourg ; 2004.
- [8] Berkhout B. Structure and function of the human immunodeficiency virus leader RNA. *Prog Nucleic Acid Res Mol Biol*. 1996;54 :1–34.
- [9] Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Swanstrom R, et al. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*. 2009 Aug ;460(7256) :711–716. Available from : <http://dx.doi.org/10.1038/nature08237>.
- [10] Karn J. Tackling Tat. *J Mol Biol*. 1999 Oct ;293(2) :235–254. Available from : <http://dx.doi.org/10.1006/jmbi.1999.3060>.
- [11] Harrich D, Ulich C, Gaynor RB. A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription. *J Virol*. 1996 Jun ;70(6) :4017–4027.
- [12] Clever JL, Eckstein DA, Parslow TG. Genetic dissociation of the encapsidation and reverse transcription functions in the 5' R region of human immunodeficiency virus type 1. *J Virol*. 1999 Jan ;73(1) :101–109.
- [13] Klasens BI, Thiesen M, Virtanen A, Berkhout B. The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure. *Nucleic Acids Res*. 1999 Jan ;27(2) :446–454.
- [14] Paillart JC, Shehu-Xhilaga M, Marquet R, Mak J. Dimerization of retroviral RNA genomes : an inseparable pair. *Nat Rev Microbiol*. 2004 Jun ;2(6) :461–472. Available from : <http://dx.doi.org/10.1038/nrmicro903>.
- [15] O'Reilly MM, McNally MT, Beemon KL. Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology*. 1995 Nov ;213(2) :373–385. Available from : <http://dx.doi.org/10.1006/viro.1995.0010>.
- [16] Purcell DF, Martin MA. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol*. 1993 Nov ;67(11) :6365–6378.
- [17] Spriggs S, Garyu L, Connor R, Summers MF. Potential intra- and intermolecular interactions involving the unique-5' region of the HIV-1 5'-UTR. *Biochemistry*. 2008 Dec ;47(49) :13064–13073. Available from : <http://dx.doi.org/10.1021/bi8014373>.

- [18] Malim MH, Böhnlein S, Fenrick R, Le SY, Maizel JV, Cullen BR. Functional comparison of the Rev trans-activators encoded by different primate immunodeficiency virus species. *Proc Natl Acad Sci U S A*. 1989 Nov ;86(21) :8222–8226.
- [19] Zimmel RW, Kelley AC, Karn J, Butler PJ. Flexible regions of RNA structure facilitate co-operative Rev assembly on the Rev-response element. *J Mol Biol*. 1996 May ;258(5) :763–777.
- [20] Reicin AS, Kalpana G, Paik S, Marmon S, Goff S. Sequences in the human immunodeficiency virus type 1 U3 region required for in vivo and in vitro integration. *J Virol*. 1995 Sep ;69(9) :5904–5907.
- [21] Aloia RC, Tian H, Jensen FC. Lipid composition and fluidity of the human immunodeficiency virus envelope and host cell plasma membranes. *Proc Natl Acad Sci U S A*. 1993 Jun ;90(11) :5181–5185.
- [22] Ono A, Freed EO. Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proc Natl Acad Sci U S A*. 2001 Nov ;98(24) :13925–13930. Available from : <http://dx.doi.org/10.1073/pnas.241320298>.
- [23] Ott DE. Potential roles of cellular proteins in HIV-1. *Rev Med Virol*. 2002 ;12(6) :359–374. Available from : <http://dx.doi.org/10.1002/rmv.367>.
- [24] Roux KH, Taylor KA. AIDS virus envelope spike structure. *Curr Opin Struct Biol*. 2007 Apr ;17(2) :244–252. Available from : <http://dx.doi.org/10.1016/j.sbi.2007.03.008>.
- [25] chin Huang C, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, et al. Structure of a V3-containing HIV-1 gp120 core. *Science*. 2005 Nov ;310(5750) :1025–1028. Available from : <http://dx.doi.org/10.1126/science.1118398>.
- [26] Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J Biol Chem*. 1990 Jun ;265(18) :10373–10382.
- [27] Zhu X, Borchers C, Bienstock RJ, Tomer KB. Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. *Biochemistry*. 2000 Sep ;39(37) :11194–11204.
- [28] Polzer S, Dittmar MT, Schmitz H, Schreiber M. The N-linked glycan g15 within the V3 loop of the HIV-1 external glycoprotein gp120 affects coreceptor usage, cellular tropism, and neutralization. *Virology*. 2002 Dec ;304(1) :70–80.
- [29] Thordsen I, Polzer S, Schreiber M. Infection of cells expressing CXCR4 mutants lacking N-glycosylation at the N-terminal extracellular domain is enhanced for R5X4-dualtropic human immunodeficiency virus type-1. *BMC Infect Dis*. 2002 Dec ;2 :31.
- [30] Pinter A. Roles of HIV-1 Env variable regions in viral neutralization and vaccine development. *Curr HIV Res*. 2007 Nov ;5(6) :542–553.
- [31] Profy AT, Salinas PA, Eckler LI, Dunlop NM, Nara PL, Putney SD. Epitopes recognized by the neutralizing antibodies of an HIV-1-infected individual. *J Immunol*. 1990 Jun ;144(12) :4641–4647.
- [32] Zwick MB, Burton DR. HIV-1 neutralization : mechanisms and relevance to vaccine design. *Curr HIV Res*. 2007 Nov ;5(6) :608–624.
- [33] Prabakaran P, Dimitrov AS, Fouts TR, Dimitrov DS. Structure and function of the HIV envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors. *Adv Pharmacol*. 2007 ;55 :33–97. Available from : [http://dx.doi.org/10.1016/S1054-3589\(07\)55002-7](http://dx.doi.org/10.1016/S1054-3589(07)55002-7).
- [34] Hill CP, Worthylake D, Bancroft DP, Christensen AM, Sundquist WI. Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein : implications for membrane association and assembly. *Proc Natl Acad Sci U S A*. 1996 Apr ;93(7) :3099–3104.
- [35] Adamson CS, Freed EO. Human immunodeficiency virus type 1 assembly, release, and maturation. *Adv Pharmacol*. 2007 ;55 :347–387. Available from : [http://dx.doi.org/10.1016/S1054-3589\(07\)55010-6](http://dx.doi.org/10.1016/S1054-3589(07)55010-6).
- [36] Bukrinskaya A. HIV-1 matrix protein : a mysterious regulator of the viral life cycle. *Virus Res*. 2007 Mar ;124(1-2) :1–11. Available from : <http://dx.doi.org/10.1016/j.virusres.2006.07.001>.
- [37] Hearps AC, Jans DA. Regulating the functions of the HIV-1 matrix protein. *AIDS Res Hum Retroviruses*. 2007 Mar ;23(3) :341–346. Available from : <http://dx.doi.org/10.1089/aid.2006.0108>.

- [38] Massiah MA, Starich MR, Paschall C, Summers MF, Christensen AM, Sundquist WI. Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. *J Mol Biol.* 1994 Nov ;244(2) :198–223. Available from : <http://dx.doi.org/10.1006/jmbi.1994.1719>.
- [39] Ono A, Ablan SD, Lockett SJ, Nagashima K, Freed EO. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc Natl Acad Sci U S A.* 2004 Oct ;101(41) :14889–14894. Available from : <http://dx.doi.org/10.1073/pnas.0405596101>.
- [40] Saad JS, Loeliger E, Luncsford P, Liriano M, Tai J, Kim A, et al. Point mutations in the HIV-1 matrix protein turn off the myristyl switch. *J Mol Biol.* 2007 Feb ;366(2) :574–585. Available from : <http://dx.doi.org/10.1016/j.jmb.2006.11.068>.
- [41] Spearman P, Horton R, Ratner L, Kuli-Zade I. Membrane binding of human immunodeficiency virus type 1 matrix protein in vivo supports a conformational myristyl switch mechanism. *J Virol.* 1997 Sep ;71(9) :6582–6592.
- [42] Lopez-Vergès S, Camus G, Blot G, Beauvoir R, Benarous R, Berlioz-Torrent C. Tail-interacting protein TIP47 is a connector between Gag and Env and is required for Env incorporation into HIV-1 virions. *Proc Natl Acad Sci U S A.* 2006 Oct ;103(40) :14947–14952. Available from : <http://dx.doi.org/10.1073/pnas.0602941103>.
- [43] Wu Z, Alexandratos J, Ericksen B, Lubkowski J, Gallo RC, Lu W. Total chemical synthesis of N-myristoylated HIV-1 matrix protein p17 : structural and mechanistic implications of p17 myristoylation. *Proc Natl Acad Sci U S A.* 2004 Aug ;101(32) :11587–11592. Available from : <http://dx.doi.org/10.1073/pnas.0404649101>.
- [44] Purohit P, Dupont S, Stevenson M, Green MR. Sequence-specific interaction between HIV-1 matrix protein and viral genomic RNA revealed by in vitro genetic selection. *RNA.* 2001 Apr ;7(4) :576–584.
- [45] Gamble TR, Vajdos FF, Yoo S, Worthylake DK, Houseweart M, Sundquist WI, et al. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell.* 1996 Dec ;87(7) :1285–1294.
- [46] Gamble TR, Yoo S, Vajdos FF, von Schwedler UK, Worthylake DK, Wang H, et al. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science.* 1997 Oct ;278(5339) :849–853.
- [47] Gitti RK, Lee BM, Walker J, Summers MF, Yoo S, Sundquist WI. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science.* 1996 Jul ;273(5272) :231–235.
- [48] Momany C, Kovari LC, Prongay AJ, Keller W, Gitti RK, Lee BM, et al. Crystal structure of dimeric HIV-1 capsid protein. *Nat Struct Biol.* 1996 Sep ;3(9) :763–770.
- [49] von Schwedler UK, Stray KM, Garrus JE, Sundquist WI. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J Virol.* 2003 May ;77(9) :5439–5450.
- [50] Worthylake DK, Wang H, Yoo S, Sundquist WI, Hill CP. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr D Biol Crystallogr.* 1999 Jan ;55(Pt 1) :85–92. Available from : <http://dx.doi.org/10.1107/S0907444998007689>.
- [51] Byeon IJL, Meng X, Jung J, Zhao G, Yang R, Ahn J, et al. Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell.* 2009 Nov ;139(4) :780–790. Available from : <http://dx.doi.org/10.1016/j.cell.2009.10.010>.
- [52] Ganser-Pornillos BK, Cheng A, Yeager M. Structure of full-length HIV-1 CA : a model for the mature capsid lattice. *Cell.* 2007 Oct ;131(1) :70–79. Available from : <http://dx.doi.org/10.1016/j.cell.2007.08.018>.
- [53] Ganser-Pornillos BK, Yeager M, Sundquist WI. The structural biology of HIV assembly. *Curr Opin Struct Biol.* 2008 Apr ;18(2) :203–217. Available from : <http://dx.doi.org/10.1016/j.sbi.2008.02.001>.
- [54] Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, et al. X-ray structures of the hexameric building block of the HIV capsid. *Cell.* 2009 Jun ;137(7) :1282–1292. Available from : <http://dx.doi.org/10.1016/j.cell.2009.04.063>.

- [55] Sterling TR, Hoover DR, Astemborski J, Vlahov D, Bartlett JG, Schupbach J. Heat-denatured human immunodeficiency virus type 1 protein 24 antigen : prognostic value in adults with early-stage disease. *J Infect Dis.* 2002 Oct ;186(8) :1181–1185. Available from : <http://dx.doi.org/10.1086/343807>.
- [56] Freed EO. Viral late domains. *J Virol.* 2002 May ;76(10) :4679–4687.
- [57] Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase : implications for drug resistance. *Science.* 1998 ;282(5394) :1669–75.
- [58] Strack B, Calistri A, Craig S, Popova E, Göttlinger HG. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell.* 2003 Sep ;114(6) :689–699.
- [59] VerPlank L, Bouamr F, LaGrassa TJ, Agresta B, Kikonyogo A, Leis J, et al. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc Natl Acad Sci U S A.* 2001 Jul ;98(14) :7724–7729. Available from : <http://dx.doi.org/10.1073/pnas.131059198>.
- [60] Gottwein E, Jäger S, Habermann A, Kräusslich HG. Cumulative mutations of ubiquitin acceptor sites in human immunodeficiency virus type 1 gag cause a late budding defect. *J Virol.* 2006 Jul ;80(13) :6267–6275. Available from : <http://dx.doi.org/10.1128/JVI.02177-05>.
- [61] Kondo E, Mammano F, Cohen EA, Göttlinger HG. The p6gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of Vpr into heterologous viral particles. *J Virol.* 1995 May ;69(5) :2759–2764.
- [62] Paxton W, Connor RI, Landau NR. Incorporation of Vpr into human immunodeficiency virus type 1 virions : requirement for the p6 region of gag and mutational analysis. *J Virol.* 1993 Dec ;67(12) :7229–7237.
- [63] Pettit SC, Simsic J, Loeb DD, Everitt L, Hutchison CA, Swanstrom R. Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. *J Biol Chem.* 1991 Aug ;266(22) :14539–14547.
- [64] Kaplan AH, Manchester M, Swanstrom R. The activity of the protease of human immunodeficiency virus type 1 is initiated at the membrane of infected cells before the release of viral proteins and is required for release to occur with maximum efficiency. *J Virol.* 1994 Oct ;68(10) :6782–6786.
- [65] Suguna K, Padlan EA, Smith CW, Carlson WD, Davies DR. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis* : implications for a mechanism of action. *Proc Natl Acad Sci U S A.* 1987 Oct ;84(20) :7009–7013.
- [66] Prabu-Jeyabalan M, Nalivaika E, Schiffer CA. Substrate shape determines specificity of recognition for HIV-1 protease : analysis of crystal structures of six substrate complexes. *Structure.* 2002 Mar ;10(3) :369–381.
- [67] You L, Garwicz D, Rögnvaldsson T. Comprehensive bioinformatic analysis of the specificity of human immunodeficiency virus type 1 protease. *J Virol.* 2005 Oct ;79(19) :12477–12486. Available from : <http://dx.doi.org/10.1128/JVI.79.19.12477-12486.2005>.
- [68] Michel F, Crucifix C, Granger F, Eiler S, Mouscadet JF, Korolev S, et al. Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. *EMBO J.* 2009 Apr ;28(7) :980–991. Available from : <http://dx.doi.org/10.1038/emboj.2009.41>.
- [69] Delelis O, Carayon K, Saïb A, Deprez E, Mouscadet JF. Integrase and integration : biochemical activities of HIV-1 integrase. *Retrovirology.* 2008 ;5 :114. Available from : <http://dx.doi.org/10.1186/1742-4690-5-114>.
- [70] Wielens J, Crosby IT, Chalmers DK. A three-dimensional model of the human immunodeficiency virus type 1 integration complex. *J Comput Aided Mol Des.* 2005 May ;19(5) :301–317. Available from : <http://dx.doi.org/10.1007/s10822-005-5256-2>.
- [71] Ciuffi A, Llano M, Poeschla E, Hoffmann C, Leipzig J, Shinn P, et al. A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med.* 2005 Dec ;11(12) :1287–1289. Available from : <http://dx.doi.org/10.1038/nm1329>.
- [72] Boyer PL, Ferris AL, Clark P, Whitmer J, Frank P, Tantillo C, et al. Mutational analysis of the fingers and palm subdomains of human immunodeficiency virus type-1 (HIV-1) reverse transcriptase. *J Mol Biol.* 1994 ;243(3) :472–83.

- [73] Jacobo-Molina A, Ding J, Nanni RG, Clark J A D, Lu X, Tantillo C, et al. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc Natl Acad Sci U S A*. 1993;90(13) :6320–4.
- [74] Kew Y, Olsen LR, Japour AJ, Prasad VR. Insertions into the beta3-beta4 hairpin loop of HIV-1 reverse transcriptase reveal a role for fingers subdomain in processive polymerization. *J Biol Chem*. 1998 Mar ;273(13) :7529–7537.
- [75] Boyer PL, Ferris AL, Hughes SH. Mutational analysis of the fingers domain of human immunodeficiency virus type 1 reverse transcriptase. *J Virol*. 1992 Dec ;66(12) :7533–7537.
- [76] Johnson MS, McClure MA, Feng DF, Gray J, Doolittle RF. Computer analysis of retroviral pol genes : assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. *Proc Natl Acad Sci U S A*. 1986 Oct ;83(20) :7648–7652.
- [77] Patel PH, Jacobo-Molina A, Ding J, Tantillo C, Clark J A D, Raag R, et al. Insights into DNA polymerization mechanisms from structure and function analysis of HIV-1 reverse transcriptase. *Biochemistry*. 1995;34(16) :5351–63.
- [78] Ding J, Hughes SH, Arnold E. Protein-nucleic acid interactions and DNA conformation in a complex of human immunodeficiency virus type 1 reverse transcriptase with a double-stranded DNA template-primer. *Biopolymers*. 1997 ;44(2) :125–38.
- [79] Ghosh M, Jacques PS, Rodgers DW, Ottman M, Darlix JL, Grice SFL. Alterations to the primer grip of p66 HIV-1 reverse transcriptase and their consequences for template-primer utilization. *Biochemistry*. 1996 Jul ;35(26) :8553–8562. Available from : <http://dx.doi.org/10.1021/bi952773j>.
- [80] Ghosh M, Williams J, Powell MD, Levin JG, Grice SFL. Mutating a conserved motif of the HIV-1 reverse transcriptase palm subdomain alters primer utilization. *Biochemistry*. 1997 May ;36(19) :5758–5768. Available from : <http://dx.doi.org/10.1021/bi963045e>.
- [81] Powell MD, Ghosh M, Jacques PS, Howard KJ, Grice SFL, Levin JG. Alanine-scanning mutations in the "primer grip" of p66 HIV-1 reverse transcriptase result in selective loss of RNA priming activity. *J Biol Chem*. 1997 May ;272(20) :13262–13269.
- [82] Hermann T, Meier T, Gotte M, Heumann H. The 'helix clamp' in HIV-1 reverse transcriptase : a new nucleic acid binding motif common in nucleic acid polymerases. *Nucleic Acids Res*. 1994 ;22(22) :4625–33.
- [83] Huang H, Harrison SC, Verdine GL. Trapping of a catalytic HIV reverse transcriptase*template :primer complex through a disulfide bond. *Chem Biol*. 2000 May ;7(5) :355–364.
- [84] Lin S, Long S, Ramirez SM, Cotter RJ, Woods AS. Characterization of the "helix clamp" motif of HIV-1 reverse transcriptase using MALDI-TOF MS and surface plasmon resonance. *Anal Chem*. 2000 ;72(11) :2635–40.
- [85] Divita G, Restle T, Goody RS. Characterization of the dimerization process of HIV-1 reverse transcriptase heterodimer using intrinsic protein fluorescence. *FEBS Lett*. 1993 Jun ;324(2) :153–158.
- [86] Wang J, Smerdon SJ, Jäger J, Kohlstaedt LA, Rice PA, Friedman JM, et al. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. *Proc Natl Acad Sci U S A*. 1994 Jul ;91(15) :7242–7246.
- [87] Davies JF, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science*. 1991 Apr ;252(5002) :88–95.
- [88] Hostomska Z, Matthews DA, Davies JF, Nides BR, Hostomsky Z. Proteolytic release and crystallization of the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase. *J Biol Chem*. 1991 Aug ;266(22) :14697–14702.
- [89] Maier G, Dietrich U, Panhans B, Schröder B, RübSamen-Waigmann H, Cellai L, et al. Mixed reconstitution of mutated subunits of HIV-1 reverse transcriptase coexpressed in *Escherichia coli* - two tags tie it up. *Eur J Biochem*. 1999 Apr ;261(1) :10–18.
- [90] Wintersberger U. Ribonucleases H of retroviral and cellular origin. *Pharmacol Ther*. 1990 ;48(2) :259–280.

- [91] Wisniewski M, Balakrishnan M, Palaniappan C, Fay PJ, Bambara RA. The sequential mechanism of HIV reverse transcriptase RNase H. *J Biol Chem.* 2000 Dec ;275(48) :37664–37671. Available from : <http://dx.doi.org/10.1074/jbc.M007381200>.
- [92] Wisniewski M, Chen Y, Balakrishnan M, Palaniappan C, Roques BP, Fay PJ, et al. Substrate requirements for secondary cleavage by HIV-1 reverse transcriptase RNase H. *J Biol Chem.* 2002 ;277(32) :28400–10.
- [93] Gopalakrishnan V, Peliska JA, Benkovic SJ. Human immunodeficiency virus type 1 reverse transcriptase : spatial and temporal relationship between the polymerase and RNase H activities. *Proc Natl Acad Sci U S A.* 1992 ;89(22) :10763–7.
- [94] Peliska JA, Benkovic SJ. Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science.* 1992 ;258(5085) :1112–8.
- [95] Diamond TL, Roshal M, Jamburuthugoda VK, Reynolds HM, Merriam AR, Lee KY, et al. Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. *J Biol Chem.* 2004 Dec ;279(49) :51545–51553. Available from : <http://dx.doi.org/10.1074/jbc.M408573200>.
- [96] Harrich D, Ulich C, García-Martínez LF, Gaynor RB. Tat is required for efficient HIV-1 reverse transcription. *EMBO J.* 1997 Mar ;16(6) :1224–1235. Available from : <http://dx.doi.org/10.1093/emboj/16.6.1224>.
- [97] Chiu YL, Ho CK, Saha N, Schwer B, Shuman S, Rana TM. Tat stimulates cotranscriptional capping of HIV mRNA. *Mol Cell.* 2002 Sep ;10(3) :585–597.
- [98] Berro R, Kehn K, de la Fuente C, Pumfery A, Adair R, Wade J, et al. Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. *J Virol.* 2006 Apr ;80(7) :3189–3204. Available from : <http://dx.doi.org/10.1128/JVI.80.7.3189-3204.2006>.
- [99] Howcroft TK, Strebel K, Martin MA, Singer DS. Repression of MHC class I gene promoter activity by two-exon Tat of HIV. *Science.* 1993 May ;260(5112) :1320–1322.
- [100] Pollard VW, Malim MH. The HIV-1 Rev protein. *Annu Rev Microbiol.* 1998 ;52 :491–532. Available from : <http://dx.doi.org/10.1146/annurev.micro.52.1.491>.
- [101] Felber BK, Hadzopoulou-Cladaras M, Cladaras C, Copeland T, Pavlakis GN. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc Natl Acad Sci U S A.* 1989 Mar ;86(5) :1495–1499.
- [102] Müller B, Tessmer U, Schubert U, Kräusslich HG. Human immunodeficiency virus type 1 Vpr protein is incorporated into the virion in significantly smaller amounts than gag and is phosphorylated in infected cells. *J Virol.* 2000 Oct ;74(20) :9727–9731.
- [103] Welker R, Hohenberg H, Tessmer U, Huckhagel C, Kräusslich HG. Biochemical and structural analysis of isolated mature cores of human immunodeficiency virus type 1. *J Virol.* 2000 Feb ;74(3) :1168–1177.
- [104] Zhang S, Pointer D, Singer G, Feng Y, Park K, Zhao LJ. Direct binding to nucleic acids by Vpr of human immunodeficiency virus type 1. *Gene.* 1998 Jun ;212(2) :157–166.
- [105] Rouzic EL, Benichou S. The Vpr protein from HIV-1 : distinct roles along the viral life cycle. *Retrovirology.* 2005 ;2 :11. Available from : <http://dx.doi.org/10.1186/1742-4690-2-11>.
- [106] Muthumani K, Choo AY, Premkumar A, Hwang DS, Thieu KP, Desai BM, et al. Human immunodeficiency virus type 1 (HIV-1) Vpr-regulated cell death : insights into mechanism. *Cell Death Differ.* 2005 Aug ;12 Suppl 1 :962–970. Available from : <http://dx.doi.org/10.1038/sj.cdd.4401583>.
- [107] Planelles V, Jowett JB, Li QX, Xie Y, Hahn B, Chen IS. Vpr-induced cell cycle arrest is conserved among primate lentiviruses. *J Virol.* 1996 Apr ;70(4) :2516–2524.
- [108] Jacotot E, Ravagnan L, Loeffler M, Ferri KF, Vieira HL, Zamzami N, et al. The HIV-1 viral protein R induces apoptosis via a direct effect on the mitochondrial permeability transition pore. *J Exp Med.* 2000 Jan ;191(1) :33–46.
- [109] Benson RE, Sanfridson A, Ottinger JS, Doyle C, Cullen BR. Downregulation of cell-surface CD4 expression by simian immunodeficiency virus Nef prevents viral super infection. *J Exp Med.* 1993 Jun ;177(6) :1561–1566.

- [110] Schwartz O, Maréchal V, Gall SL, Lemonnier F, Heard JM. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat Med.* 1996 Mar;2(3) :338–342.
- [111] Yang OO, Nguyen PT, Kalams SA, Dorfman T, Göttlinger HG, Stewart S, et al. Nef-mediated resistance of human immunodeficiency virus type 1 to antiviral cytotoxic T lymphocytes. *J Virol.* 2002 Feb;76(4) :1626–1631.
- [112] Henriët S, Richer D, Bernacchi S, Decroly E, Vigne R, Ehresmann B, et al. Cooperative and specific binding of Vif to the 5' region of HIV-1 genomic RNA. *J Mol Biol.* 2005 Nov;354(1) :55–72. Available from : <http://dx.doi.org/10.1016/j.jmb.2005.09.025>.
- [113] Bouyac M, Courcoul M, Bertoia G, Baudat Y, Gabuzda D, Blanc D, et al. Human immunodeficiency virus type 1 Vif protein binds to the Pr55Gag precursor. *J Virol.* 1997 Dec;71(12) :9358–9365.
- [114] Henriët S, Sinck L, Bec G, Gorelick RJ, Marquet R, Paillart JC. Vif is a RNA chaperone that could temporarily regulate RNA dimerization and the early steps of HIV-1 reverse transcription. *Nucleic Acids Res.* 2007;35(15) :5141–5153. Available from : <http://dx.doi.org/10.1093/nar/gkm542>.
- [115] Willey RL, Maldarelli F, Martin MA, Strebel K. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. *J Virol.* 1992 Dec;66(12) :7193–7200.
- [116] Göttlinger HG, Dorfman T, Cohen EA, Haseltine WA. Vpu protein of human immunodeficiency virus type 1 enhances the release of capsids produced by gag gene constructs of widely divergent retroviruses. *Proc Natl Acad Sci U S A.* 1993 Aug;90(15) :7381–7385.
- [117] Matarrese P, Malorni W. Human immunodeficiency virus (HIV)-1 proteins and cytoskeleton : partners in viral life and host cell death. *Cell Death Differ.* 2005 Aug;12 Suppl 1 :932–941. Available from : <http://dx.doi.org/10.1038/sj.cdd.4401582>.
- [118] Campbell EM, Nunez R, Hope TJ. Disruption of the actin cytoskeleton can complement the ability of Nef to enhance human immunodeficiency virus type 1 infectivity. *J Virol.* 2004 Jun;78(11) :5745–5755. Available from : <http://dx.doi.org/10.1128/JVI.78.11.5745-5755.2004>.
- [119] Silva FSD, Venturini DS, Wagner E, Shank PR, Sharma S. CD4-independent infection of human B cells with HIV type 1 : detection of unintegrated viral DNA. *AIDS Res Hum Retroviruses.* 2001 Nov;17(17) :1585–1598. Available from : <http://dx.doi.org/10.1089/088922201753341997>.
- [120] Livingstone WJ, Moore M, Innes D, Bell JE, Simmonds P. Frequent infection of peripheral blood CD8-positive T-lymphocytes with HIV-1. Edinburgh Heterosexual Transmission Study Group. *Lancet.* 1996 Sep;348(9028) :649–654.
- [121] Al-Harthi L, Guilbert LJ, Hoxie JA, Landay A. Trophoblasts are productively infected by CD4-independent isolate of HIV type 1. *AIDS Res Hum Retroviruses.* 2002 Jan;18(1) :13–17. Available from : <http://dx.doi.org/10.1089/088922202753394673>.
- [122] Fortin JF, Cantin R, Lamontagne G, Tremblay M. Host-derived ICAM-1 glycoproteins incorporated on human immunodeficiency virus type 1 are biologically active and enhance viral infectivity. *J Virol.* 1997 May;71(5) :3588–3596.
- [123] Saphire AC, Bobardt MD, Galloway PA. Human immunodeficiency virus type 1 hijacks host cyclophilin A for its attachment to target cells. *Immunol Res.* 2000;21(2-3) :211–217. Available from : <http://dx.doi.org/10.1385/IR:21:2-3:211>.
- [124] Saphire AC, Bobardt MD, Galloway PA. Host cyclophilin A mediates HIV-1 attachment to target cells via heparans. *EMBO J.* 1999 Dec;18(23) :6771–6785. Available from : <http://dx.doi.org/10.1093/emboj/18.23.6771>.
- [125] Braaten D, Franke EK, Luban J. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. *J Virol.* 1996 Jun;70(6) :3551–3560.
- [126] Sokolskaja E, Sayah DM, Luban J. Target cell cyclophilin A modulates human immunodeficiency virus type 1 infectivity. *J Virol.* 2004 Dec;78(23) :12800–12808. Available from : <http://dx.doi.org/10.1128/JVI.78.23.12800-12808.2004>.
- [127] Fassati A, Goff SP. Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol.* 2001 Apr;75(8) :3626–3635. Available from : <http://dx.doi.org/10.1128/JVI.75.8.3626-3635.2001>.

- [128] Fitzon T, Leschonsky B, Bieler K, Paulus C, Schröder J, Wolf H, et al. Proline residues in the HIV-1 NH2-terminal capsid domain : structure determinants for proper core assembly and subsequent steps of early replication. *Virology*. 2000 Mar ;268(2) :294–307. Available from : <http://dx.doi.org/10.1006/viro.1999.0178>.
- [129] Zhang H, Dornadula G, Orenstein J, Pomerantz RJ. Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription : evidence indicating that reverse transcription may not take place within the intact viral core. *J Hum Virol*. 2000 ;3(3) :165–172.
- [130] Arhel N, Munier S, Souque P, Mollier K, Charneau P. Nuclear import defect of human immunodeficiency virus type 1 DNA flap mutants is not dependent on the viral strain or target cell type. *J Virol*. 2006 Oct ;80(20) :10262–10269. Available from : <http://dx.doi.org/10.1128/JVI.00974-06>.
- [131] Arhel NJ, Souquere-Besse S, Munier S, Souque P, Guadagnini S, Rutherford S, et al. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO J*. 2007 Jun ;26(12) :3025–3037. Available from : <http://dx.doi.org/10.1038/sj.emboj.7601740>.
- [132] McDonald D, Vodicka MA, Lucero G, Svitkina TM, Borisy GG, Emerman M, et al. Visualization of the intracellular behavior of HIV in living cells. *J Cell Biol*. 2002 Nov ;159(3) :441–452. Available from : <http://dx.doi.org/10.1083/jcb.200203150>.
- [133] Warrilow D, Harrich D. HIV-1 replication from after cell entry to the nuclear periphery. *Curr HIV Res*. 2007 May ;5(3) :293–299.
- [134] Warrilow D, Stenzel D, Harrich D. Isolated HIV-1 core is active for reverse transcription. *Retrovirology*. 2007 ;4 :77. Available from : <http://dx.doi.org/10.1186/1742-4690-4-77>.
- [135] Strebel K. Virus-host interactions : role of HIV proteins Vif, Tat, and Rev. *AIDS*. 2003 ;17 Suppl 4 :S25–S34.
- [136] Yap MW, Nisole S, Stoye JP. A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Curr Biol*. 2005 Jan ;15(1) :73–78. Available from : <http://dx.doi.org/10.1016/j.cub.2004.12.042>.
- [137] Arhel NJ, Souquere-Besse S, Charneau P. Wild-type and central DNA flap defective HIV-1 lentiviral vector genomes : intracellular visualization at ultrastructural resolution levels. *Retrovirology*. 2006 ;3 :38. Available from : <http://dx.doi.org/10.1186/1742-4690-3-38>.
- [138] Farnet CM, Bushman FD. HIV-1 cDNA integration : requirement of HMG I(Y) protein for function of preintegration complexes in vitro. *Cell*. 1997 Feb ;88(4) :483–492.
- [139] Chen H, Engelman A. The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc Natl Acad Sci U S A*. 1998 Dec ;95(26) :15270–15274.
- [140] Sherman MP, Greene WC. Slipping through the door : HIV entry into the nucleus. *Microbes Infect*. 2002 Jan ;4(1) :67–73.
- [141] Zennou V, Petit C, Guetard D, Nerhbass U, Montagnier L, Charneau P. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*. 2000 Apr ;101(2) :173–185. Available from : [http://dx.doi.org/10.1016/S0092-8674\(00\)80828-4](http://dx.doi.org/10.1016/S0092-8674(00)80828-4).
- [142] Rivière L, Darlix JL, Cimarelli A. Analysis of the viral elements required in the nuclear import of HIV-1 DNA. *J Virol*. 2010 Jan ;84(2) :729–739. Available from : <http://dx.doi.org/10.1128/JVI.01952-09>.
- [143] Engelman A, Mizuuchi K, Craigie R. HIV-1 DNA integration : mechanism of viral DNA cleavage and DNA strand transfer. *Cell*. 1991 Dec ;67(6) :1211–1221.
- [144] Yoder KE, Bushman FD. Repair of gaps in retroviral DNA integration intermediates. *J Virol*. 2000 Dec ;74(23) :11191–11200.
- [145] Carteau S, Gorelick RJ, Bushman FD. Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro : stimulation by the viral nucleocapsid protein. *J Virol*. 1999 Aug ;73(8) :6670–6679.
- [146] Hindmarsh P, Ridky T, Reeves R, Andrade M, Skalka AM, Leis J. HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration in vitro. *J Virol*. 1999 Apr ;73(4) :2994–3003.

- [147] Pandey KK, Sinha S, Grandgenett DP. Transcriptional coactivator LEDGF/p75 modulates human immunodeficiency virus type 1 integrase-mediated concerted integration. *J Virol*. 2007 Apr;81(8) :3969–3979. Available from : <http://dx.doi.org/10.1128/JVI.02322-06>.
- [148] Jung A, Maier R, Vartanian JP, Bocharov G, Jung V, Fischer U, et al. Recombination : Multiply infected spleen cells in HIV patients. *Nature*. 2002 Jul;418(6894) :144. Available from : <http://dx.doi.org/10.1038/418144a>.
- [149] Schröder ARW, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*. 2002 Aug;110(4) :521–529.
- [150] Cullen BR. Regulation of HIV-1 gene expression. *FASEB J*. 1991 Jul;5(10) :2361–2368.
- [151] Jeang KT, Xiao H, Rich EA. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *J Biol Chem*. 1999 Oct;274(41) :28837–28840.
- [152] Gatignol A. Transcription of HIV : Tat and cellular chromatin. *Adv Pharmacol*. 2007;55 :137–159. Available from : [http://dx.doi.org/10.1016/S1054-3589\(07\)55004-0](http://dx.doi.org/10.1016/S1054-3589(07)55004-0).
- [153] Veschambre P, Roisin A, Jalinot P. Biochemical and functional interaction of the human immunodeficiency virus type 1 Tat transactivator with the general transcription factor TFIIB. *J Gen Virol*. 1997 Sep;78 (Pt 9) :2235–2245.
- [154] Kashanchi F, Piras G, Radonovich MF, Duvall JF, Fattaey A, Chiang CM, et al. Direct interaction of human TFIID with the HIV-1 transactivator tat. *Nature*. 1994 Jan;367(6460) :295–299. Available from : <http://dx.doi.org/10.1038/367295a0>.
- [155] Marzio G, Tyagi M, Gutierrez MI, Giacca M. HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc Natl Acad Sci U S A*. 1998 Nov;95(23) :13519–13524.
- [156] Parada CA, Roeder RG. Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature*. 1996 Nov;384(6607) :375–378. Available from : <http://dx.doi.org/10.1038/384375a0>.
- [157] Schwartz S, Felber BK, Fenyö EM, Pavlakis GN. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *J Virol*. 1990 Nov;64(11) :5448–5456.
- [158] Malim MH, Tiley LS, McCarn DF, Rusche JR, Hauber J, Cullen BR. HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell*. 1990 Feb;60(4) :675–683.
- [159] Freed EO, Myers DJ, Risser R. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc Natl Acad Sci U S A*. 1990 Jun;87(12) :4650–4654.
- [160] Brasey A, Lopez-Lastra M, Ohlmann T, Beerens N, Berkhout B, Darlix JL, et al. The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J Virol*. 2003 Apr;77(7) :3939–3949.
- [161] Buck CB, Shen X, Egan MA, Pierson TC, Walker CM, Siliciano RF. The human immunodeficiency virus type 1 gag gene encodes an internal ribosome entry site. *J Virol*. 2001 Jan;75(1) :181–191. Available from : <http://dx.doi.org/10.1128/JVI.75.1.181-191.2001>.
- [162] Gheysen D, Jacobs E, de Foresta F, Thiriart C, Francotte M, Thines D, et al. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell*. 1989 Oct;59(1) :103–112.
- [163] Cimarelli A, Sandin S, Höglund S, Luban J. Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA. *J Virol*. 2000 Apr;74(7) :3046–3057.
- [164] Grigorov B, Décimo D, Smagulova F, Péchoux C, Mougél M, Muriaux D, et al. Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers. *Retrovirology*. 2007;4 :54. Available from : <http://dx.doi.org/10.1186/1742-4690-4-54>.
- [165] Darlix JL, Garrido JL, Morellet N, Mély Y, de Rocquigny H. Properties, functions, and drug targeting of the multifunctional nucleocapsid protein of the human immunodeficiency virus. *Adv Pharmacol*. 2007;55 :299–346. Available from : [http://dx.doi.org/10.1016/S1054-3589\(07\)55009-X](http://dx.doi.org/10.1016/S1054-3589(07)55009-X).
- [166] Kaye JF, Lever AM. Human immunodeficiency virus types 1 and 2 differ in the predominant mechanism used for selection of genomic RNA for encapsidation. *J Virol*. 1999 Apr;73(4) :3023–3031.

- [167] Linial ML, Miller AD. Retroviral RNA packaging : sequence requirements and implications. *Curr Top Microbiol Immunol.* 1990;157 :125–152.
- [168] Poole E, Strappe P, Mok HP, Hicks R, Lever AML. HIV-1 Gag-RNA interaction occurs at a perinuclear/centrosomal site; analysis by confocal microscopy and FRET. *Traffic.* 2005 Sep;6(9) :741–755. Available from : <http://dx.doi.org/10.1111/j.1600-0854.2005.00312.x>.
- [169] Briggs JAG, Simon MN, Gross I, Kräusslich HG, Fuller SD, Vogt VM, et al. The stoichiometry of Gag protein in HIV-1. *Nat Struct Mol Biol.* 2004 Jul;11(7) :672–675. Available from : <http://dx.doi.org/10.1038/nsmb785>.
- [170] Klein KC, Reed JC, Lingappa JR. Intracellular destinies : degradation, targeting, assembly, and endocytosis of HIV Gag. *AIDS Rev.* 2007;9(3) :150–161.
- [171] Grigorov B, Arcanger F, Roingeard P, Darlix JL, Muriaux D. Assembly of infectious HIV-1 in human epithelial and T-lymphoblastic cell lines. *J Mol Biol.* 2006 Jun;359(4) :848–862. Available from : <http://dx.doi.org/10.1016/j.jmb.2006.04.017>.
- [172] Nydegger S, Foti M, Derdowski A, Spearman P, Thali M. HIV-1 egress is gated through late endosomal membranes. *Traffic.* 2003 Dec;4(12) :902–910.
- [173] Ono A, Freed EO. Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body. *J Virol.* 2004 Feb;78(3) :1552–1563.
- [174] Pelchen-Matthews A, Kramer B, Marsh M. Infectious HIV-1 assembles in late endosomes in primary macrophages. *J Cell Biol.* 2003 Aug;162(3) :443–455. Available from : <http://dx.doi.org/10.1083/jcb.200304008>.
- [175] Gousset K, Ablan SD, Coren LV, Ono A, Soheilian F, Nagashima K, et al. Real-time visualization of HIV-1 GAG trafficking in infected macrophages. *PLoS Pathog.* 2008 Mar;4(3) :e1000015. Available from : <http://dx.doi.org/10.1371/journal.ppat.1000015>.
- [176] Derdowski A, Ding L, Spearman P. A novel fluorescence resonance energy transfer assay demonstrates that the human immunodeficiency virus type 1 Pr55Gag I domain mediates Gag-Gag interactions. *J Virol.* 2004 Feb;78(3) :1230–1242.
- [177] Finzi A, Orthwein A, Mercier J, Cohen EA. Productive human immunodeficiency virus type 1 assembly takes place at the plasma membrane. *J Virol.* 2007 Jul;81(14) :7476–7490. Available from : <http://dx.doi.org/10.1128/JVI.00308-07>.
- [178] Hübner W, Chen P, Portillo AD, Liu Y, Gordon RE, Chen BK. Sequence of human immunodeficiency virus type 1 (HIV-1) Gag localization and oligomerization monitored with live confocal imaging of a replication-competent, fluorescently tagged HIV-1. *J Virol.* 2007 Nov;81(22) :12596–12607. Available from : <http://dx.doi.org/10.1128/JVI.01088-07>.
- [179] Jouvenet N, Bieniasz PD, Simon SM. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature.* 2008 Jul;454(7201) :236–240. Available from : <http://dx.doi.org/10.1038/nature06998>.
- [180] Jouvenet N, Neil SJD, Bess C, Johnson MC, Virgen CA, Simon SM, et al. Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol.* 2006 Dec;4(12) :e435. Available from : <http://dx.doi.org/10.1371/journal.pbio.0040435>.
- [181] Deneka M, Pelchen-Matthews A, Byland R, Ruiz-Mateos E, Marsh M. In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspansins CD81, CD9, and CD53. *J Cell Biol.* 2007 Apr;177(2) :329–341. Available from : <http://dx.doi.org/10.1083/jcb.200609050>.
- [182] Welsch S, Keppler OT, Habermann A, Allespach I, Krijnse-Locker J, Kräusslich HG. HIV-1 buds predominantly at the plasma membrane of primary human macrophages. *PLoS Pathog.* 2007 Mar;3(3) :e36. Available from : <http://dx.doi.org/10.1371/journal.ppat.0030036>.
- [183] Shehu-Xhilaga M, Crowe SM, Mak J. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J Virol.* 2001 Feb;75(4) :1834–1841. Available from : <http://dx.doi.org/10.1128/JVI.75.4.1834-1841.2001>.
- [184] Shehu-Xhilaga M, Hill M, Marshall JA, Kappes J, Crowe SM, Mak J. The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of pol protein. *J Virol.* 2002 May;76(9) :4331–4340.

- [185] Corbin A, Grigorov B, Roingeard P, Darlix JL, Muriaux D. [Revisiting HIV-1 assembly]. *Med Sci (Paris)*. 2008 Jan;24(1) :49–55.
- [186] Swanson CM, Malim MH. SnapShot : HIV-1 proteins. *Cell*. 2008 May;133(4) :742, 742.e1. Available from : <http://dx.doi.org/10.1016/j.cell.2008.05.005>.
- [187] Wolf D, Goff SP. Host restriction factors blocking retroviral replication. *Annu Rev Genet*. 2008;42 :143–163. Available from : <http://dx.doi.org/10.1146/annurev.genet.42.110807.091704>.
- [188] Mergener K, Fäcke M, Welker R, Brinkmann V, Gelderblom HR, Kräusslich HG. Analysis of HIV particle formation using transient expression of subviral constructs in mammalian cells. *Virology*. 1992 Jan;186(1) :25–39.
- [189] Muriaux D, Darlix JL, Cimarelli A. Targeting the assembly of the human immunodeficiency virus type I. *Curr Pharm Des*. 2004;10(30) :3725–3739.
- [190] Cosson P. Direct interaction between the envelope and matrix proteins of HIV-1. *EMBO J*. 1996 Nov;15(21) :5783–5788.
- [191] Khorchid A, Javanbakht H, Wise S, Halwani R, Parniak MA, Wainberg MA, et al. Sequences within Pr160gag-pol affecting the selective packaging of primer tRNA(Lys3) into HIV-1. *J Mol Biol*. 2000 May;299(1) :17–26. Available from : <http://dx.doi.org/10.1006/jmbi.2000.3709>.
- [192] Kleiman L. tRNA(Lys3) : the primer tRNA for reverse transcription in HIV-1. *IUBMB Life*. 2002 Feb;53(2) :107–114. Available from : <http://dx.doi.org/10.1080/15216540211469>.
- [193] Marsh M, Thali M. HIV's great escape. *Nat Med*. 2003 Oct;9(10) :1262–1263. Available from : <http://dx.doi.org/10.1038/nm1003-1262>.
- [194] Nguyen DH, Hildreth JE. Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. *J Virol*. 2000 Apr;74(7) :3264–3272.
- [195] Grigorov B, Attuil-Audenis V, Perugi F, Nedelec M, Watson S, Pique C, et al. A role for CD81 on the late steps of HIV-1 replication in a chronically infected T cell line. *Retrovirology*. 2009;6 :28. Available from : <http://dx.doi.org/10.1186/1742-4690-6-28>.
- [196] Nydegger S, Khurana S, Kremontsov DN, Foti M, Thali M. Mapping of tetraspanin-enriched microdomains that can function as gateways for HIV-1. *J Cell Biol*. 2006 Jun;173(5) :795–807. Available from : <http://dx.doi.org/10.1083/jcb.200508165>.
- [197] Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. *Nature*. 2003 Mar;422(6929) :307–312. Available from : <http://dx.doi.org/10.1038/nature01470>.
- [198] Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med*. 2009 Dec;361(23) :2209–2220. Available from : <http://dx.doi.org/10.1056/NEJMoa0908492>.
- [199] Thomas JA, Gorelick RJ. Nucleocapsid protein function in early infection processes. *Virus Res*. 2008 Jun;134(1-2) :39–63. Available from : <http://dx.doi.org/10.1016/j.virusres.2007.12.006>.
- [200] Shehu-Xhilaga M, Kraeusslich HG, Pettit S, Swanstrom R, Lee JY, Marshall JA, et al. Proteolytic processing of the p2/nucleocapsid cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation. *J Virol*. 2001 Oct;75(19) :9156–9164. Available from : <http://dx.doi.org/10.1128/JVI.75.19.9156-9164.2001>.
- [201] Mirambeau G, Lyonais S, Coulaud D, Hameau L, Lafosse S, Jeusset J, et al. HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner. *PLoS ONE*. 2007;2(7) :e669. Available from : <http://dx.doi.org/10.1371/journal.pone.0000669>.
- [202] Sheng N, Pettit SC, Tritch RJ, Ozturk DH, Rayner MM, Swanstrom R, et al. Determinants of the human immunodeficiency virus type 1 p15NC-RNA interaction that affect enhanced cleavage by the viral protease. *J Virol*. 1997 Aug;71(8) :5723–5732.
- [203] Coren LV, Thomas JA, Chertova E, Sowder RC, Gagliardi TD, Gorelick RJ, et al. Mutational analysis of the C-terminal gag cleavage sites in human immunodeficiency virus type 1. *J Virol*. 2007 Sep;81(18) :10047–10054. Available from : <http://dx.doi.org/10.1128/JVI.02496-06>.

- [204] Henderson LE, Bowers MA, Sowder RC, Serabyn SA, Johnson DG, Bess JW, et al. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1 : posttranslational modifications, proteolytic processings, and complete amino acid sequences. *J Virol.* 1992 Apr ;66(4) :1856–1865.
- [205] Tanchou V, Decimo D, Péchoux C, Lener D, Rogemond V, Berthoux L, et al. Role of the N-terminal zinc finger of human immunodeficiency virus type 1 nucleocapsid protein in virus structure and replication. *J Virol.* 1998 May ;72(5) :4442–4447.
- [206] Morellet N, Jullian N, Rocquigny HD, Maignet B, Darlix JL, Roques BP. Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR. *EMBO J.* 1992 Aug ;11(8) :3059–3065.
- [207] Morellet N, de Rocquigny H, Mély Y, Jullian N, Déméné H, Ottmann M, et al. Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR. *J Mol Biol.* 1994 Jan ;235(1) :287–301.
- [208] de Rocquigny H, Ficheux D, Gabus C, Fournie-Zaluski MC, Darlix JL, Roques BP. First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem Biophys Res Commun.* 1991 ;180(2) :1010–8.
- [209] Gorelick RJ, Gagliardi TD, Bosche WJ, Wilttrout TA, Coren LV, Chabot DJ, et al. Strict conservation of the retroviral nucleocapsid protein zinc finger is strongly influenced by its role in viral infection processes : characterization of HIV-1 particles containing mutant nucleocapsid zinc-coordinating sequences. *Virology.* 1999 Mar ;256(1) :92–104. Available from : <http://dx.doi.org/10.1006/viro.1999.9629>.
- [210] Cruceanu M, Urbaneja MA, Hixson CV, Johnson DG, Datta SA, Fivash MJ, et al. Nucleic acid binding and chaperone properties of HIV-1 Gag and nucleocapsid proteins. *Nucleic Acids Res.* 2006 ;34(2) :593–605. Available from : <http://dx.doi.org/10.1093/nar/gkj458>.
- [211] Berg JM. Potential metal-binding domains in nucleic acid binding proteins. *Science.* 1986 Apr ;232(4749) :485–487.
- [212] Bess JW, Powell PJ, Issaq HJ, Schumack LJ, Grimes MK, Henderson LE, et al. Tightly bound zinc in human immunodeficiency virus type 1, human T-cell leukemia virus type I, and other retroviruses. *J Virol.* 1992 Feb ;66(2) :840–847.
- [213] Bombarda E, Grell E, Roques BP, Mély Y. Molecular mechanism of the Zn²⁺-induced folding of the distal CCHC finger motif of the HIV-1 nucleocapsid protein. *Biophys J.* 2007 Jul ;93(1) :208–217. Available from : <http://dx.doi.org/10.1529/biophysj.106.101378>.
- [214] Bombarda E, Morellet N, Cherradi H, Spiess B, Bouaziz S, Grell E, et al. Determination of the pK(a) of the four Zn²⁺-coordinating residues of the distal finger motif of the HIV-1 nucleocapsid protein : consequences on the binding of Zn²⁺. *J Mol Biol.* 2001 Jul ;310(3) :659–672. Available from : <http://dx.doi.org/10.1006/jmbi.2001.4770>.
- [215] Mély Y, Rocquigny HD, Morellet N, Roques BP, Gérard D. Zinc binding to the HIV-1 nucleocapsid protein : a thermodynamic investigation by fluorescence spectroscopy. *Biochemistry.* 1996 Apr ;35(16) :5175–5182. Available from : <http://dx.doi.org/10.1021/bi952587d>.
- [216] South TL, Blake PR, Sowder RC, Arthur LO, Henderson LE, Summers MF. The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers. *Biochemistry.* 1990 Aug ;29(34) :7786–7789.
- [217] Green LM, Berg JM. A retroviral Cys-Xaa2-Cys-Xaa4-His-Xaa4-Cys peptide binds metal ions : spectroscopic studies and a proposed three-dimensional structure. *Proc Natl Acad Sci U S A.* 1989 Jun ;86(11) :4047–4051.
- [218] Omichinski JG, Clore GM, Sakaguchi K, Appella E, Gronenborn AM. Structural characterization of a 39-residue synthetic peptide containing the two zinc binding domains from the HIV-1 p7 nucleocapsid protein by CD and NMR spectroscopy. *FEBS Lett.* 1991 Nov ;292(1-2) :25–30.
- [219] Summers MF, Henderson LE, Chance MR, Bess JW, South TL, Blake PR, et al. Nucleocapsid zinc fingers detected in retroviruses : EXAFS studies of intact viruses and the solution-state structure

- of the nucleocapsid protein from HIV-1. *Protein Sci.* 1992 May ;1(5) :563–574. Available from : <http://dx.doi.org/10.1002/pro.5560010502>.
- [220] Déméné H, Dong CZ, Ottmann M, Rouyez MC, Jullian N, Morellet N, et al. 1H NMR structure and biological studies of the His23→Cys mutant nucleocapsid protein of HIV-1 indicate that the conformation of the first zinc finger is critical for virus infectivity. *Biochemistry.* 1994 Oct ;33(39) :11707–11716.
- [221] Gorelick RJ, Chabot DJ, Rein A, Henderson LE, Arthur LO. The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J Virol.* 1993 Jul ;67(7) :4027–4036.
- [222] Gorelick RJ, Fu W, Gagliardi TD, Bosche WJ, Rein A, Henderson LE, et al. Characterization of the block in replication of nucleocapsid protein zinc finger mutants from moloney murine leukemia virus. *J Virol.* 1999 Oct ;73(10) :8185–8195.
- [223] De Rocquigny H, Gabus C, Vincent A, Fournie-Zaluski MC, Roques B, Darlix JL. Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc Natl Acad Sci U S A.* 1992 ;89(14) :6472–6.
- [224] Mély Y, Jullian N, Morellet N, Rocquigny HD, Dong CZ, Piémont E, et al. Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer. *Biochemistry.* 1994 Oct ;33(40) :12085–12091.
- [225] Amarasinghe GK, Guzman RND, Turner RB, Chancellor KJ, Wu ZR, Summers MF. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J Mol Biol.* 2000 Aug ;301(2) :491–511. Available from : <http://dx.doi.org/10.1006/jmbi.2000.3979>.
- [226] Beltz H, Clauss C, Piémont E, Ficheux D, Gorelick RJ, Roques B, et al. Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis. *J Mol Biol.* 2005 May ;348(5) :1113–1126. Available from : <http://dx.doi.org/10.1016/j.jmb.2005.02.042>.
- [227] Bourbigot S, Ramalanjaona N, Boudier C, Salgado GFJ, Roques BP, Mély Y, et al. How the HIV-1 nucleocapsid protein binds and destabilises the (-)primer binding site during reverse transcription. *J Mol Biol.* 2008 Nov ;383(5) :1112–1128. Available from : <http://dx.doi.org/10.1016/j.jmb.2008.08.046>.
- [228] De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science.* 1998 ;279(5349) :384–8.
- [229] Morellet N, Déméné H, Teilleux V, Huynh-Dinh T, de Rocquigny H, Fournié-Zaluski MC, et al. Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J Mol Biol.* 1998 Oct ;283(2) :419–434. Available from : <http://dx.doi.org/10.1006/jmbi.1998.2098>.
- [230] Lam WC, Maki AH, Casas-Finet JR, Erickson JW, Kane BP, Sowder RC, et al. Phosphorescence and optically detected magnetic resonance investigation of the binding of the nucleocapsid protein of the human immunodeficiency virus type 1 and related peptides to RNA. *Biochemistry.* 1994 Sep ;33(35) :10693–10700.
- [231] Mély Y, Piémont E, Sorinas-Jimeno M, de Rocquigny H, Jullian N, Morellet N, et al. Structural and dynamic characterization of the aromatic amino acids of the human immunodeficiency virus type I nucleocapsid protein zinc fingers and their involvement in heterologous tRNA(Phe) binding : a steady-state and time-resolved fluorescence study. *Biophys J.* 1993 Oct ;65(4) :1513–1522. Available from : [http://dx.doi.org/10.1016/S0006-3495\(93\)81222-0](http://dx.doi.org/10.1016/S0006-3495(93)81222-0).
- [232] Dorfman T, Luban J, Goff SP, Haseltine WA, Gottlinger HG. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol.* 1993 ;67(10) :6159–69.
- [233] Amarasinghe GK, De Guzman RN, Turner RB, Summers MF. NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform. *J Mol Biol.* 2000 ;299(1) :145–56.

- [234] Guzman RND, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science*. 1998 Jan;279(5349) :384–388.
- [235] Anfinsen CB. Principles that govern the folding of protein chains. *Science*. 1973 Jul;181(96) :223–230.
- [236] Levinthal C. Are there pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie Biologique*. 1968;65 :44–45.
- [237] Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding : a synthesis. *Proteins*. 1995 Mar;21(3) :167–195. Available from : <http://dx.doi.org/10.1002/prot.340210302>.
- [238] Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat Struct Biol*. 1997 Jan;4(1) :10–19.
- [239] James LC, Tawfik DS. Conformational diversity and protein evolution—a 60-year-old hypothesis revisited. *Trends Biochem Sci*. 2003 Jul;28(7) :361–368.
- [240] Zwanzig R. Two-state models of protein folding kinetics. *Proc Natl Acad Sci U S A*. 1997 Jan;94(1) :148–150.
- [241] Chen SJ, Dill KA. RNA folding energy landscapes. *Proc Natl Acad Sci U S A*. 2000 Jan;97(2) :646–651.
- [242] Treiber DK, Williamson JR. Exposing the kinetic traps in RNA folding. *Curr Opin Struct Biol*. 1999 Jun;9(3) :339–345. Available from : [http://dx.doi.org/10.1016/S0959-440X\(99\)80045-1](http://dx.doi.org/10.1016/S0959-440X(99)80045-1).
- [243] Russell R, Zhuang X, Babcock HP, Millett IS, Doniach S, Chu S, et al. Exploring the folding landscape of a structured RNA. *Proc Natl Acad Sci U S A*. 2002 Jan;99(1) :155–160. Available from : <http://dx.doi.org/10.1073/pnas.221593598>.
- [244] Herschlag D. RNA chaperones and the RNA folding problem. *J Biol Chem*. 1995;270(36) :20871–4.
- [245] Fisher RJ, Rein A, Fivash M, Urbaneja MA, Casas-Finet JR, Medaglia M, et al. Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J Virol*. 1998 Mar;72(3) :1902–1909.
- [246] Mély Y, de Rocquigny H, Sorinas-Jimeno M, Keith G, Roques BP, Marquet R, et al. Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3Lys), in vitro, is essentially not specific. *J Biol Chem*. 1995 Jan;270(4) :1650–1656.
- [247] Shubsda MF, Paoletti AC, Hudson BS, Borer PN. Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry*. 2002 Apr;41(16) :5276–5282.
- [248] Urbaneja MA, Kane BP, Johnson DG, Gorelick RJ, Henderson LE, Casas-Finet JR. Binding properties of the human immunodeficiency virus type 1 nucleocapsid protein p7 to a model RNA : elucidation of the structural determinants for function. *J Mol Biol*. 1999 Mar;287(1) :59–75. Available from : <http://dx.doi.org/10.1006/jmbi.1998.2521>.
- [249] Vuilleumier C, Bombarda E, Morellet N, Gerard D, Roques BP, Mely Y. Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7 : a fluorescence study. *Biochemistry*. 1999;38(51) :16816–25.
- [250] Lapadat-Tapolsky M, Pernelle C, Borie C, Darlix JL. Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1. *Nucleic Acids Res*. 1995;23(13) :2434–41.
- [251] Berkowitz R, Fisher J, Goff SP. RNA packaging. *Curr Top Microbiol Immunol*. 1996;214 :177–218.
- [252] Urbaneja MA, Wu M, Casas-Finet JR, Karpel RL. HIV-1 nucleocapsid protein as a nucleic acid chaperone : spectroscopic study of its helix-destabilizing properties, structural binding specificity, and annealing activity. *J Mol Biol*. 2002 May;318(3) :749–764. Available from : [http://dx.doi.org/10.1016/S0022-2836\(02\)00043-8](http://dx.doi.org/10.1016/S0022-2836(02)00043-8).
- [253] Wu JQ, Ozarowski A, Maki AH, Urbaneja MA, Henderson LE, Casas-Finet JR. Binding of the nucleocapsid protein of type 1 human immunodeficiency virus to nucleic acids studied using phosphorescence and optically detected magnetic resonance. *Biochemistry*. 1997 Oct;36(41) :12506–12518. Available from : <http://dx.doi.org/10.1021/bi970676f>.

- [254] Berglund JA, Charpentier B, Rosbash M. A high affinity binding site for the HIV-1 nucleocapsid protein. *Nucleic Acids Res.* 1997 Mar ;25(5) :1042–1049.
- [255] D'Souza V, Dey A, Habib D, Summers MF. NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus. *J Mol Biol.* 2004 Mar ;337(2) :427–442. Available from : <http://dx.doi.org/10.1016/j.jmb.2004.01.037>.
- [256] Avilov S, Godet J, Piemont E, Ly YM. Site-specific characterization of HIV-1 nucleocapsid protein binding to oligonucleotides with two binding sites. *Biochemistry.* 2009 Feb ;Available from : <http://dx.doi.org/10.1021/bi8022366>.
- [257] Hagan N, Fabris D. Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 Psi-recognition element. *Biochemistry.* 2003 Sep ;42(36) :10736–10745. Available from : <http://dx.doi.org/10.1021/bi0348922>.
- [258] Maki AH, Ozarowski A, Misra A, Urbaneja MA, Casas-Finet JR. Phosphorescence and optically detected magnetic resonance of HIV-1 nucleocapsid protein complexes with stem-loop sequences of the genomic Psi-recognition element. *Biochemistry.* 2001 Feb ;40(5) :1403–1412.
- [259] Paoletti AC, Shubsda MF, Hudson BS, Borer PN. Affinities of the nucleocapsid protein for variants of SL3 RNA in HIV-1. *Biochemistry.* 2002 Dec ;41(51) :15423–15428.
- [260] Johnson PE, Turner RB, Wu ZR, Hairston L, Guo J, Levin JG, et al. A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry.* 2000 Aug ;39(31) :9084–9091.
- [261] Ramboarina S, Srividya N, Atkinson RA, Morellet N, Roques BP, Lefèvre JF, et al. Effects of temperature on the dynamic behaviour of the HIV-1 nucleocapsid NCp7 and its DNA complex. *J Mol Biol.* 2002 Feb ;316(3) :611–627. Available from : <http://dx.doi.org/10.1006/jmbi.2001.5379>.
- [262] Stote RH, Kellenberger E, Muller H, Bombarda E, Roques BP, Kieffer B, et al. Structure of the His44 → Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein : a combined NMR, molecular dynamics simulation, and fluorescence study. *Biochemistry.* 2004 Jun ;43(24) :7687–7697. Available from : <http://dx.doi.org/10.1021/bi036137u>.
- [263] Tisne C, Roques BP, Dardel F. The annealing mechanism of HIV-1 reverse transcription primer onto the viral genome. *J Biol Chem.* 2004 ;279(5) :3588–95.
- [264] Tisne C, Roques BP, Dardel F. Heteronuclear NMR studies of the interaction of tRNA(Lys)3 with HIV-1 nucleocapsid protein. *J Mol Biol.* 2001 ;306(3) :443–54.
- [265] Tisné C, Roques BP, Dardel F. Specific recognition of primer tRNA Lys 3 by HIV-1 nucleocapsid protein : involvement of the zinc fingers and the N-terminal basic extension. *Biochimie.* 2003 May ;85(5) :557–561.
- [266] Williams MC, Gorelick RJ, Musier-Forsyth K. Specific zinc-finger architecture required for HIV-1 nucleocapsid protein's nucleic acid chaperone function. *Proc Natl Acad Sci U S A.* 2002 Jun ;99(13) :8614–8619. Available from : <http://dx.doi.org/10.1073/pnas.132128999>.
- [267] Azoulay J, Clamme JP, Darlix JL, Roques BP, Mély Y. Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations. *J Mol Biol.* 2003 Feb ;326(3) :691–700.
- [268] Beltz H, Azoulay J, Bernacchi S, Clamme JP, Ficheux D, Roques B, et al. Impact of the terminal bulges of HIV-1 cTAR DNA on its stability and the destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol.* 2003 Apr ;328(1) :95–108.
- [269] Beltz H, Piémont E, Schaub E, Ficheux D, Roques B, Darlix JL, et al. Role of the structure of the top half of HIV-1 cTAR DNA on the nucleic acid destabilizing activity of the nucleocapsid protein NCp7. *J Mol Biol.* 2004 May ;338(4) :711–723. Available from : <http://dx.doi.org/10.1016/j.jmb.2004.03.019>.
- [270] Williams MC, Rouzina I, Wenner JR, Gorelick RJ, Musier-Forsyth K, Bloomfield VA. Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching. *Proc Natl Acad Sci U S A.* 2001 May ;98(11) :6121–6126. Available from : <http://dx.doi.org/10.1073/pnas.101033198>.

- [271] Chan B, Weidemaier K, Yip WT, Barbara PF, Musier-Forsyth K. Intra-tRNA distance measurements for nucleocapsid protein-independent tRNA unwinding during priming of HIV reverse transcription. *Proc Natl Acad Sci U S A*. 1999 Jan;96(2) :459–464.
- [272] Gregoire CJ, Gautheret D, Loret EP. No tRNA³Lys unwinding in a complex with HIV NCp7. *J Biol Chem*. 1997 Oct;272(40) :25143–25148.
- [273] Hargittai MR, Mangla AT, Gorelick RJ, Musier-Forsyth K. HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing. *J Mol Biol*. 2001 Oct;312(5) :985–997. Available from : <http://dx.doi.org/10.1006/jmbi.2001.5021>.
- [274] Barat C, Schatz O, Le Grice S, Darlix JL. Analysis of the interactions of HIV1 replication primer tRNA(Lys,3) with nucleocapsid protein and reverse transcriptase. *J Mol Biol*. 1993;231(2) :185–90.
- [275] Darlix JL, Lapadat-Tapolsky M, de Rocquigny H, Roques BP. First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J Mol Biol*. 1995;254(4) :523–37.
- [276] Dib-Hajj F, Khan R, Giedroc DP. Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. *Protein Sci*. 1993 Feb;2(2) :231–243. Available from : <http://dx.doi.org/10.1002/pro.5560020212>.
- [277] Tsuchihashi Z, Brown PO. DNA strand exchange and selective DNA annealing promoted by the human immunodeficiency virus type 1 nucleocapsid protein. *J Virol*. 1994 Sep;68(9) :5863–5870.
- [278] You JC, McHenry CS. Human immunodeficiency virus nucleocapsid protein accelerates strand transfer of the terminally redundant sequences involved in reverse transcription. *J Biol Chem*. 1994 Dec;269(50) :31491–31495.
- [279] Heath MJ, Derebail SS, Gorelick RJ, DeStefano JJ. Differing roles of the N- and C-terminal zinc fingers in human immunodeficiency virus nucleocapsid protein-enhanced nucleic acid annealing. *J Biol Chem*. 2003 Aug;278(33) :30755–30763. Available from : <http://dx.doi.org/10.1074/jbc.M303819200>.
- [280] Godet J, de Rocquigny H, Raja C, Glasser N, Ficheux D, Darlix JL, et al. During the early phase of HIV-1 DNA synthesis, nucleocapsid protein directs hybridization of the TAR complementary sequences via the ends of their double-stranded stem. *J Mol Biol*. 2006 Mar;356(5) :1180–1192. Available from : <http://dx.doi.org/10.1016/j.jmb.2005.12.038>.
- [281] Ramalanjaona N, de Rocquigny H, Millet A, Ficheux D, Darlix JL, Mély Y. Investigating the mechanism of the nucleocapsid protein chaperoning of the second strand transfer during HIV-1 DNA synthesis. *J Mol Biol*. 2007 Dec;374(4) :1041–1053. Available from : <http://dx.doi.org/10.1016/j.jmb.2007.10.001>.
- [282] Vo MN, Barany G, Rouzina I, Musier-Forsyth K. Mechanistic studies of mini-TAR RNA/DNA annealing in the absence and presence of HIV-1 nucleocapsid protein. *J Mol Biol*. 2006 Oct;363(1) :244–261. Available from : <http://dx.doi.org/10.1016/j.jmb.2006.08.039>.
- [283] Krishnamoorthy G, Roques B, Darlix JL, Mély Y. DNA condensation by the nucleocapsid protein of HIV-1 : a mechanism ensuring DNA protection. *Nucleic Acids Res*. 2003 Sep;31(18) :5425–5432.
- [284] Le Cam E, Coulaud D, Delain E, Petitjean P, Roques BP, Gerard D, et al. Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein : an electron microscopy investigation. *Biopolymers*. 1998;45(3) :217–29.
- [285] Stoylov SP, Vuilleumier C, Stoylova E, De Rocquigny H, Roques BP, Gerard D, et al. Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein. *Biopolymers*. 1997;41(3) :301–12.
- [286] Haddrick M, Lear AL, Cann AJ, Heaphy S. Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J Mol Biol*. 1996 May;259(1) :58–68. Available from : <http://dx.doi.org/10.1006/jmbi.1996.0301>.
- [287] Laughrea M, Jetté L. A 19-nucleotide sequence upstream of the 5' major splice donor is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry*. 1994 Nov;33(45) :13464–13474.

- [288] Muriaux D, Girard PM, Bonnet-Mathonière B, Paoletti J. Dimerization of HIV-1Lai RNA at low ionic strength. An autocomplementary sequence in the 5' leader region is evidenced by an antisense oligonucleotide. *J Biol Chem.* 1995 Apr ;270(14) :8209–8216.
- [289] Skripkin E, Paillart JC, Marquet R, Ehresmann B, Ehresmann C. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proc Natl Acad Sci U S A.* 1994 May ;91(11) :4945–4949.
- [290] Laughrea M, Jetté L. Kissing-loop model of HIV-1 genome dimerization : HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation. *Biochemistry.* 1996 Feb ;35(5) :1589–1598. Available from : <http://dx.doi.org/10.1021/bi951838f>.
- [291] Muriaux D, Fossé P, Paoletti J. A kissing complex together with a stable dimer is involved in the HIV-1Lai RNA dimerization process in vitro. *Biochemistry.* 1996 Apr ;35(15) :5075–5082. Available from : <http://dx.doi.org/10.1021/bi952822s>.
- [292] Muriaux D, Rocquigny HD, Roques BP, Paoletti J. NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop-loop complex into a stable dimer. *J Biol Chem.* 1996 Dec ;271(52) :33686–33692.
- [293] Ennifar E, Paillart JC, Bernacchi S, Walter P, Pale P, Decout JL, et al. A structure-based approach for targeting the HIV-1 genomic RNA dimerization initiation site. *Biochimie.* 2007 Oct ;89(10) :1195–1203. Available from : <http://dx.doi.org/10.1016/j.biochi.2007.03.003>.
- [294] Ennifar E, Paillart JC, Marquet R, Ehresmann B, Ehresmann C, Dumas P, et al. HIV-1 RNA dimerization initiation site is structurally similar to the ribosomal A site and binds aminoglycoside antibiotics. *J Biol Chem.* 2003 Jan ;278(4) :2723–2730. Available from : <http://dx.doi.org/10.1074/jbc.M205726200>.
- [295] Ennifar E, Yusupov M, Walter P, Marquet R, Ehresmann B, Ehresmann C, et al. The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure.* 1999 Nov ;7(11) :1439–1449.
- [296] Andersen ES, Contera SA, Knudsen B, Damgaard CK, Besenbacher F, Kjems J. Role of the transactivation response element in dimerization of HIV-1 RNA. *J Biol Chem.* 2004 May ;279(21) :22243–22249. Available from : <http://dx.doi.org/10.1074/jbc.M314326200>.
- [297] Abbink TEM, Berkhout B. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J Biol Chem.* 2003 Mar ;278(13) :11601–11611. Available from : <http://dx.doi.org/10.1074/jbc.M210291200>.
- [298] Dirac AMG, Huthoff H, Kjems J, Berkhout B. Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. *Nucleic Acids Res.* 2002 Jun ;30(12) :2647–2655.
- [299] Lanchy JM, Rentz CA, Ivanovitch JD, Lodmell JS. Elements located upstream and downstream of the major splice donor site influence the ability of HIV-2 leader RNA to dimerize in vitro. *Biochemistry.* 2003 Mar ;42(9) :2634–2642. Available from : <http://dx.doi.org/10.1021/bi0271190>.
- [300] Clever J, Sasseti C, Parslow TG. RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J Virol.* 1995 ;69(4) :2101–9.
- [301] Clever JL, Parslow TG. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J Virol.* 1997 May ;71(5) :3407–3414.
- [302] Gorelick RJ, Nigida SM, Bess JW, Arthur LO, Henderson LE, Rein A. Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J Virol.* 1990 Jul ;64(7) :3207–3211.
- [303] Amarasinghe GK, Zhou J, Miskimon M, Chancellor KJ, McDonald JA, Matthews AG, et al. Stem-loop SL4 of the HIV-1 psi RNA packaging signal exhibits weak affinity for the nucleocapsid protein. structural studies and implications for genome recognition. *J Mol Biol.* 2001 Dec ;314(5) :961–970. Available from : <http://dx.doi.org/10.1006/jmbi.2000.5182>.
- [304] McBride MS, Panganiban AT. Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. *J Virol.* 1997 Mar ;71(3) :2050–2058.

- [305] McBride MS, Schwartz MD, Panganiban AT. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation. *J Virol*. 1997 Jun ;71(6) :4544–4554.
- [306] Cruceanu M, Gorelick RJ, Musier-Forsyth K, Rouzina I, Williams MC. Rapid kinetics of protein-nucleic acid interaction is a major component of HIV-1 nucleocapsid protein's nucleic acid chaperone function. *J Mol Biol*. 2006 Nov ;363(5) :867–877. Available from : <http://dx.doi.org/10.1016/j.jmb.2006.08.070>.
- [307] Mak J, Khorchid A, Cao Q, Huang Y, Lowy I, Parniak MA, et al. Effects of mutations in Pr160gag-pol upon tRNA(Lys3) and Pr160gag-plo incorporation into HIV-1. *J Mol Biol*. 1997 Jan ;265(4) :419–431.
- [308] Marquet R, Isel C, Ehresmann C, Ehresmann B. tRNAs as primer of reverse transcriptases. *Biochimie*. 1995 ;77(1-2) :113–124.
- [309] Cen S, Huang Y, Khorchid A, Darlix JL, Wainberg MA, Kleiman L. The role of Pr55(gag) in the annealing of tRNA^{Lys} to human immunodeficiency virus type 1 genomic RNA. *J Virol*. 1999 ;73(5) :4485–8.
- [310] Feng YX, Campbell S, Harvin D, Ehresmann B, Ehresmann C, Rein A. The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity : possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site. *J Virol*. 1999 May ;73(5) :4251–4256.
- [311] Feng YX, Copeland TD, Henderson LE, Gorelick RJ, Bosche WJ, Levin JG, et al. HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro. *Proc Natl Acad Sci U S A*. 1996 Jul ;93(15) :7577–7581.
- [312] Hargittai MRS, Gorelick RJ, Rouzina I, Musier-Forsyth K. Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site. *J Mol Biol*. 2004 Apr ;337(4) :951–968. Available from : <http://dx.doi.org/10.1016/j.jmb.2004.01.054>.
- [313] Prats AC, Billy GD, Wang P, Darlix JL. CUG initiation codon used for the synthesis of a cell surface antigen coded by the murine leukemia virus. *J Mol Biol*. 1989 Jan ;205(2) :363–372.
- [314] Prats AC, Sarih L, Gabus C, Litvak S, Keith G, Darlix JL. Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA. *EMBO J*. 1988 Jun ;7(6) :1777–1783.
- [315] Guo F, Saadatmand J, Niu M, Kleiman L. Roles of Gag and NCp7 in facilitating tRNA(Lys)(3) Annealing to viral RNA in human immunodeficiency virus type 1. *J Virol*. 2009 Aug ;83(16) :8099–8107. Available from : <http://dx.doi.org/10.1128/JVI.00488-09>.
- [316] Khan R, Chang HO, Kaluarachchi K, Giedroc DP. Interaction of retroviral nucleocapsid proteins with transfer RNAPhe : a lead ribozyme and 1H NMR study. *Nucleic Acids Res*. 1996 Sep ;24(18) :3568–3575.
- [317] Khan R, Giedroc DP. Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. *J Biol Chem*. 1992 Apr ;267(10) :6689–6695.
- [318] Baudin F, Marquet R, Isel C, Darlix JL, Ehresmann B, Ehresmann C. Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J Mol Biol*. 1993 Jan ;229(2) :382–397. Available from : <http://dx.doi.org/10.1006/jmbi.1993.1041>.
- [319] Goldschmidt V, Rigourde M, Ehresmann C, Grice SFJL, Ehresmann B, Marquet R. Direct and indirect contributions of RNA secondary structure elements to the initiation of HIV-1 reverse transcription. *J Biol Chem*. 2002 Nov ;277(45) :43233–43242. Available from : <http://dx.doi.org/10.1074/jbc.M205295200>.
- [320] Isel C, Marquet R, Keith G, Ehresmann C, Ehresmann B. Modified nucleotides of tRNA(3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. *J Biol Chem*. 1993 Dec ;268(34) :25269–25272.
- [321] Lanchy JM, Ehresmann C, Le Grice SF, Ehresmann B, Marquet R. Binding and kinetic properties of HIV-1 reverse transcriptase markedly differ during initiation and elongation of reverse transcription. *Embo J*. 1996 ;15(24) :7178–87.

- [322] Beerens N, Berkhout B. Switching the in vitro tRNA usage of HIV-1 by simultaneous adaptation of the PBS and PAS. *RNA*. 2002 Mar ;8(3) :357–369.
- [323] Beerens N, Berkhout B. The tRNA primer activation signal in the human immunodeficiency virus type 1 genome is important for initiation and processive elongation of reverse transcription. *J Virol*. 2002 Mar ;76(5) :2329–2339.
- [324] Beerens N, Groot F, Berkhout B. Initiation of HIV-1 reverse transcription is regulated by a primer activation signal. *J Biol Chem*. 2001 Aug ;276(33) :31247–31256. Available from : <http://dx.doi.org/10.1074/jbc.M102441200>.
- [325] Huthoff H, Bugala K, Barciszewski J, Berkhout B. On the importance of the primer activation signal for initiation of tRNA(lys3)-primed reverse transcription of the HIV-1 RNA genome. *Nucleic Acids Res*. 2003 Sep ;31(17) :5186–5194.
- [326] Ooms M, Cupac D, Abbink TEM, Huthoff H, Berkhout B. The availability of the primer activation signal (PAS) affects the efficiency of HIV-1 reverse transcription initiation. *Nucleic Acids Res*. 2007 ;35(5) :1649–1659. Available from : <http://dx.doi.org/10.1093/nar/gkm046>.
- [327] Goldschmidt V, Ehresmann C, Ehresmann B, Marquet R. Does the HIV-1 primer activation signal interact with tRNA3(Lys) during the initiation of reverse transcription? *Nucleic Acids Res*. 2003 Feb ;31(3) :850–859.
- [328] Goldschmidt V, Paillart JC, Rigourd M, Ehresmann B, Aubertin AM, Ehresmann C, et al. Structural variability of the initiation complex of HIV-1 reverse transcription. *J Biol Chem*. 2004 Aug ;279(34) :35923–35931. Available from : <http://dx.doi.org/10.1074/jbc.M404473200>.
- [329] Liang C, Rong L, Gotte M, Li X, Quan Y, Kleiman L, et al. Mechanistic studies of early pausing events during initiation of HIV-1 reverse transcription. *J Biol Chem*. 1998 ;273(33) :21309–15.
- [330] Rong L, Liang C, Hsu M, Kleiman L, Petitjean P, de Rocquigny H, et al. Roles of the human immunodeficiency virus type 1 nucleocapsid protein in annealing and initiation versus elongation in reverse transcription of viral negative-strand strong-stop DNA. *J Virol*. 1998 ;72(11) :9353–8.
- [331] Bampi C, Bibillo A, Wendeler M, Divita G, Gorelick RJ, Grice SFJL, et al. Nucleotide excision repair and template-independent addition by HIV-1 reverse transcriptase in the presence of nucleocapsid protein. *J Biol Chem*. 2006 Apr ;281(17) :11736–11743. Available from : <http://dx.doi.org/10.1074/jbc.M600290200>.
- [332] Grohmann D, Godet J, Mély Y, Darlix JL, Restle T. HIV-1 nucleocapsid traps reverse transcriptase on nucleic acid substrates. *Biochemistry*. 2008 Nov ;47(46) :12230–12240. Available from : <http://dx.doi.org/10.1021/bi801386r>.
- [333] Iwatani Y, Rosen AE, Guo J, Musier-Forsyth K, Levin JG. Efficient initiation of HIV-1 reverse transcription in vitro. Requirement for RNA sequences downstream of the primer binding site abrogated by nucleocapsid protein-dependent primer-template interactions. *J Biol Chem*. 2003 Apr ;278(16) :14185–14195. Available from : <http://dx.doi.org/10.1074/jbc.M211618200>.
- [334] Gilboa E, Mitra SW, Goff S, Baltimore D. A detailed model of reverse transcription and tests of crucial aspects. *Cell*. 1979 Sep ;18(1) :93–100.
- [335] Luo GX, Chao M, Hsieh SY, Sureau C, Nishikura K, Taylor J. A specific base transition occurs on replicating hepatitis delta virus RNA. *J Virol*. 1990 ;64(3) :1021–7.
- [336] Panganiban AT, Fiore D. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science*. 1988 Aug ;241(4869) :1064–1069.
- [337] Hu WS, Temin HM. Retroviral recombination and reverse transcription. *Science*. 1990 Nov ;250(4985) :1227–1233.
- [338] Jones JS, Allan RW, Temin HM. One retroviral RNA is sufficient for synthesis of viral DNA. *J Virol*. 1994 Jan ;68(1) :207–216.
- [339] van Wamel JL, Berkhout B. The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. *Virology*. 1998 May ;244(2) :245–251. Available from : <http://dx.doi.org/10.1006/viro.1998.9096>.

- [340] Drummond JE, Mounts P, Gorelick RJ, Casas-Finet JR, Bosche WJ, Henderson LE, et al. Wild-type and mutant HIV type 1 nucleocapsid proteins increase the proportion of long cDNA transcripts by viral reverse transcriptase. *AIDS Res Hum Retroviruses*. 1997 May ;13(7) :533–543.
- [341] Ji X, Klarmann GJ, Preston BD. Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro. *Biochemistry*. 1996 ;35(1) :132–43.
- [342] Tanchou V, Gabus C, Rogemond V, Darlix JL. Formation of stable and functional HIV-1 nucleoprotein complexes in vitro. *J Mol Biol*. 1995 ;252(5) :563–71.
- [343] Ben-Artzi H, Shemesh J, Zeelon E, Amit B, Kleiman L, Gorecki M, et al. Molecular analysis of the second template switch during reverse transcription of the HIV RNA template. *Biochemistry*. 1996 Aug ;35(32) :10549–10557. Available from : <http://dx.doi.org/10.1021/bi960439x>.
- [344] Burnett BP, McHenry CS. Posttranscriptional modification of retroviral primers is required for late stages of DNA replication. *Proc Natl Acad Sci U S A*. 1997 Jul ;94(14) :7210–7215.
- [345] Auxilien S, Keith G, Le Grice SF, Darlix JL. Role of post-transcriptional modifications of primer tRNA^{Lys,3} in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *J Biol Chem*. 1999 ;274(7) :4412–20.
- [346] Smith CM, Smith JS, Roth MJ. RNase H requirements for the second strand transfer reaction of human immunodeficiency virus type 1 reverse transcription. *J Virol*. 1999 ;73(8) :6573–81.
- [347] Wu T, Guo J, Bess J, Henderson LE, Levin JG. Molecular requirements for human immunodeficiency virus type 1 plus-strand transfer : analysis in reconstituted and endogenous reverse transcription systems. *J Virol*. 1999 ;73(6) :4794–805.
- [348] Egelé C, Schaub E, Ramalanjaona N, Piémont E, Ficheux D, Roques B, et al. HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences and chaperones their kissing interactions. *J Mol Biol*. 2004 Sep ;342(2) :453–466. Available from : <http://dx.doi.org/10.1016/j.jmb.2004.07.059>.
- [349] Dudding LR, Nkabinde NC, Mizrahi V. Analysis of the RNA- and DNA-dependent DNA polymerase activities of point mutants of HIV-1 reverse transcriptase lacking ribonuclease H activity. *Biochemistry*. 1991 Oct ;30(43) :10498–10506.
- [350] Guo J, Henderson LE, Bess J, Kane B, Levin JG. Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA. *J Virol*. 1997 Jul ;71(7) :5178–5188.
- [351] Guo J, Wu W, Yuan ZY, Post K, Crouch RJ, Levin JG. Defects in primer-template binding, processive DNA synthesis, and RNase H activity associated with chimeric reverse transcriptases having the murine leukemia virus polymerase domain joined to Escherichia coli RNase H. *Biochemistry*. 1995 Apr ;34(15) :5018–5029.
- [352] Klarmann GJ, Schaub CA, Preston BD. Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequences in vitro. *J Biol Chem*. 1993 May ;268(13) :9793–9802.
- [353] Charneau P, Clavel F. A single-stranded gap in human immunodeficiency virus unintegrated linear DNA defined by a central copy of the polypurine tract. *J Virol*. 1991 May ;65(5) :2415–2421.
- [354] Hungnes O, Tjøtta E, Grinde B. Mutations in the central polypurine tract of HIV-1 result in delayed replication. *Virology*. 1992 Sep ;190(1) :440–442.
- [355] Klarmann GJ, Yu H, Chen X, Dougherty JP, Preston BD. Discontinuous plus-strand DNA synthesis in human immunodeficiency virus type 1-infected cells and in a partially reconstituted cell-free system. *J Virol*. 1997 Dec ;71(12) :9259–9269.
- [356] Charneau P, Mirambeau G, Roux P, Paulous S, Buc H, Clavel F. HIV-1 reverse transcription. A termination step at the center of the genome. *J Mol Biol*. 1994 Sep ;241(5) :651–662. Available from : <http://dx.doi.org/10.1006/jmbi.1994.1542>.
- [357] Fuentes GM, Palaniappan C, Fay PJ, Bambara RA. Strand displacement synthesis in the central polypurine tract region of HIV-1 promotes DNA to DNA strand transfer recombination. *J Biol Chem*. 1996 Nov ;271(47) :29605–29611.

- [358] Hameau L, Jeusset J, Lafosse S, Coulaud D, Delain E, Unge T, et al. Human immunodeficiency virus type 1 central DNA flap : dynamic terminal product of plus-strand displacement dna synthesis catalyzed by reverse transcriptase assisted by nucleocapsid protein. *J Virol.* 2001 Apr ;75(7) :3301–3313. Available from : <http://dx.doi.org/10.1128/JVI.75.7.3301-3313.2001>.
- [359] Maele BV, Rijck JD, Clercq ED, Debyser Z. Impact of the central polypurine tract on the kinetics of human immunodeficiency virus type 1 vector transduction. *J Virol.* 2003 Apr ;77(8) :4685–4694.
- [360] Rijck JD, Debyser Z. The central DNA flap of the human immunodeficiency virus type 1 is important for viral replication. *Biochem Biophys Res Commun.* 2006 Oct ;349(3) :1100–1110. Available from : <http://dx.doi.org/10.1016/j.bbrc.2006.08.141>.
- [361] Rijck JD, Maele BV, Debyser Z. Positional effects of the central DNA flap in HIV-1-derived lentiviral vectors. *Biochem Biophys Res Commun.* 2005 Mar ;328(4) :987–994. Available from : <http://dx.doi.org/10.1016/j.bbrc.2005.01.052>.
- [362] Kankia BI, Musier-Forsyth K. The HIV-1 central DNA flap region contains a "flapping" third strand. *Biophys Chem.* 2007 Apr ;127(1-2) :64–68. Available from : <http://dx.doi.org/10.1016/j.bpc.2006.12.005>.
- [363] Lyonais S, Gorelick RJ, Mergny JL, Cam EL, Mirambeau G. G-quartets direct assembly of HIV-1 nucleocapsid protein along single-stranded DNA. *Nucleic Acids Res.* 2003 Oct ;31(19) :5754–5763.
- [364] Lyonais S, Hounsou C, Teulade-Fichou MP, Jeusset J, Cam EL, Mirambeau G. G-quartets assembly within a G-rich DNA flap. A possible event at the center of the HIV-1 genome. *Nucleic Acids Res.* 2002 Dec ;30(23) :5276–5283.
- [365] Allain B, Lapadat-Tapolsky M, Berlioz C, Darlix JL. Transactivation of the minus-strand DNA transfer by nucleocapsid protein during reverse transcription of the retroviral genome. *Embo J.* 1994 ;13(4) :973–81.
- [366] Barat C, Lullien V, Schatz O, Keith G, Nugeyre MT, Gruninger-Leitch F, et al. HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *Embo J.* 1989 ;8(11) :3279–85.
- [367] Cameron CE, Ghosh M, Le Grice SF, Benkovic SJ. Mutations in HIV reverse transcriptase which alter RNase H activity and decrease strand transfer efficiency are suppressed by HIV nucleocapsid protein. *Proc Natl Acad Sci U S A.* 1997 ;94(13) :6700–5.
- [368] Darlix JL, Vincent A, Gabus C, de Rocquigny H, Roques B. Trans-activation of the 5' to 3' viral DNA strand transfer by nucleocapsid protein during reverse transcription of HIV1 RNA. *C R Acad Sci III.* 1993 ;316(8) :763–71.
- [369] DeStefano JJ. Interaction of human immunodeficiency virus nucleocapsid protein with a structure mimicking a replication intermediate. Effects on stability, reverse transcriptase binding, and strand transfer. *J Biol Chem.* 1996 Jul ;271(27) :16350–16356.
- [370] Li X, Quan Y, Arts EJ, Li Z, Preston BD, de Rocquigny H, et al. Human immunodeficiency virus Type 1 nucleocapsid protein (NCp7) directs specific initiation of minus-strand DNA synthesis primed by human tRNA(Lys3) in vitro : studies of viral RNA molecules mutated in regions that flank the primer binding site. *J Virol.* 1996 ;70(8) :4996–5004.
- [371] Wu W, Henderson LE, Copeland TD, Gorelick RJ, Bosche WJ, Rein A, et al. Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract. *J Virol.* 1996 Oct ;70(10) :7132–7142.
- [372] Druillenec S, Caneparo A, de Rocquigny H, Roques BP. Evidence of interactions between the nucleocapsid protein NCp7 and the reverse transcriptase of HIV-1. *J Biol Chem.* 1999 ;274(16) :11283–8.
- [373] Lener D, Tanchou V, Roques BP, Grice SFL, Darlix JL. Involvement of HIV-I nucleocapsid protein in the recruitment of reverse transcriptase into nucleoprotein complexes formed in vitro. *J Biol Chem.* 1998 Dec ;273(50) :33781–33786.
- [374] Gratton E, Limkeman M. A continuously variable frequency cross-correlation phase fluorometer with picosecond resolution. *Biophys J.* 1983 Dec ;44(3) :315–324. Available from : [http://dx.doi.org/10.1016/S0006-3495\(83\)84305-7](http://dx.doi.org/10.1016/S0006-3495(83)84305-7).

- [375] Lakowicz JR, Maliwal BP. Construction and performance of a variable-frequency phase-modulation fluorometer. *Biophys Chem.* 1985 Jan;21(1) :61–78.
- [376] Badea MG, Brand L. Time-resolved fluorescence measurements. *Methods Enzymol.* 1979;61 :378–425.
- [377] Livesey AK, Brochon JC. Analyzing the Distribution of Decay Constants in Pulse-Fluorimetry Using the Maximum Entropy Method. *Biophys J.* 1987 Nov;52(5) :693–706.
- [378] Förster T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik.* 1948;437(1-2) :55–75. Available from : <http://dx.doi.org/10.1002/andp.19484370105>.
- [379] Summers MF. Zinc finger motif for single-stranded nucleic acids? Investigations by nuclear magnetic resonance. *J Cell Biochem.* 1991 Jan;45(1) :41–48. Available from : <http://dx.doi.org/10.1002/jcb.240450110>.
- [380] Aldovini A, Young RA. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J Virol.* 1990 May;64(5) :1920–1926.
- [381] Rein A, Henderson LE, Levin JG. Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins : significance for viral replication. *Trends Biochem Sci.* 1998;23(8) :297–301.
- [382] Clamme JP, Azoulay J, Mély Y. Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys J.* 2003 Mar;84(3) :1960–1968. Available from : [http://dx.doi.org/10.1016/S0006-3495\(03\)75004-8](http://dx.doi.org/10.1016/S0006-3495(03)75004-8).
- [383] Palmer AG, Thompson NL. Theory of sample translation in fluorescence correlation spectroscopy. *Biophys J.* 1987 Feb;51(2) :339–343. Available from : [http://dx.doi.org/10.1016/S0006-3495\(87\)83340-4](http://dx.doi.org/10.1016/S0006-3495(87)83340-4).
- [384] Widengren J, Rigler R. Fluorescence correlation spectroscopy as a tool to investigate chemical reactions in solutions and on cell surfaces. *Cell Mol Biol (Noisy-le-grand).* 1998 Jul;44(5) :857–879.
- [385] Egelé C, Piémont E, Didier P, Ficheux D, Roques B, Darlix JL, et al. The single-finger nucleocapsid protein of moloney murine leukemia virus binds and destabilizes the TAR sequences of HIV-1 but does not promote efficiently their annealing. *Biochemistry.* 2007 Dec;46(50) :14650–14662. Available from : <http://dx.doi.org/10.1021/bi7012239>.
- [386] Egelé C, Schaub E, Piémont E, de Rocquigny H, Mély Y. Investigation by fluorescence correlation spectroscopy of the chaperoning interactions of HIV-1 nucleocapsid protein with the viral DNA initiation sequences. *C R Biol.* 2005 Dec;328(12) :1041–1051. Available from : <http://dx.doi.org/10.1016/j.crvi.2005.06.005>.
- [387] Eggeling C, Fries JR, Brand L, Günther R, Seidel CA. Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proc Natl Acad Sci U S A.* 1998 Feb;95(4) :1556–1561.
- [388] Neubauer H, Gaiko N, Berger S, Schaffer J, Eggeling C, Tuma J, et al. Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy. *J Am Chem Soc.* 2007 Oct;129(42) :12746–12755. Available from : <http://dx.doi.org/10.1021/ja0722574>.
- [389] Dertinger T, Pacheco V, von der Hocht I, Hartmann R, Gregor I, Enderlein J. Two-focus fluorescence correlation spectroscopy : a new tool for accurate and absolute diffusion measurements. *Chemphyschem.* 2007 Feb;8(3) :433–443. Available from : <http://dx.doi.org/10.1002/cphc.200600638>.
- [390] Garcés J, Wittek R. Reverse-transcriptase-associated RNaseH activity mediates template switching during reverse transcription in vitro. *Proc Biol Sci.* 1991 Mar;243(1308) :235–239. Available from : <http://dx.doi.org/10.1098/rspb.1991.0037>.
- [391] Driscoll MD, Hughes SH. Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences. *J Virol.* 2000 Oct;74(19) :8785–8792.
- [392] Guo J, Wu T, Anderson J, Kane BF, Johnson DG, Gorelick RJ, et al. Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer. *J Virol.* 2000;74(19) :8980–8.

- [393] Bernacchi S, Stoylov S, Piémont E, Ficheux D, Roques BP, Darlix JL, et al. HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence. *J Mol Biol.* 2002 Mar;317(3) :385–399. Available from : <http://dx.doi.org/10.1006/jmbi.2002.5429>.
- [394] Heilman-Miller SL, Wu T, Levin JG. Alteration of nucleic acid structure and stability modulates the efficiency of minus-strand transfer mediated by the HIV-1 nucleocapsid protein. *J Biol Chem.* 2004 Oct ;279(42) :44154–44165. Available from : <http://dx.doi.org/10.1074/jbc.M401646200>.
- [395] Bernacchi S, Mely Y. Exciton interaction in molecular beacons : a sensitive sensor for short range modifications of the nucleic acid structure. *Nucleic Acids Res.* 2001 ;29(13) :E62–2.
- [396] Bernacchi S, Piémont E, Potier N, van Dorsselaer A, Mély Y. Excitonic heterodimer formation in an HIV-1 oligonucleotide labeled with a donor-acceptor pair used for fluorescence resonance energy transfer. *Biophys J.* 2003 Jan ;84(1) :643–654.
- [397] Pontius BW, Berg P. Renaturation of complementary DNA strands mediated by purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein : implications for a mechanism for rapid molecular assembly. *Proc Natl Acad Sci U S A.* 1990 Nov ;87(21) :8403–8407.
- [398] Kanevsky I, Chaminade F, Ficheux D, Moumen A, Gorelick R, Negroni M, et al. Specific interactions between HIV-1 nucleocapsid protein and the TAR element. *J Mol Biol.* 2005 May ;348(5) :1059–1077. Available from : <http://dx.doi.org/10.1016/j.jmb.2005.03.046>.
- [399] Avilov SV, Piemont E, Shvadchak V, de Rocquigny H, Mely Y. Probing dynamics of HIV-1 nucleocapsid protein/target hexanucleotide complexes by 2-aminopurine. *Nucleic Acids Res.* 2008 ;36(3) :885–96.
- [400] Cosa G, Harbron EJ, Zeng Y, Liu HW, O'Connor DB, Eta-Hosokawa C, et al. Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. *Biophys J.* 2004 Oct ;87(4) :2759–2767. Available from : <http://dx.doi.org/10.1529/biophysj.104.043083>.
- [401] Cosa G, Zeng Y, Liu HW, Landes CF, Makarov DE, Musier-Forsyth K, et al. Evidence for non-two-state kinetics in the nucleocapsid protein chaperoned opening of DNA hairpins. *J Phys Chem B.* 2006 Feb ;110(5) :2419–2426. Available from : <http://dx.doi.org/10.1021/jp054189i>.
- [402] Jiao Y, Stringfellow S, Yu H. Distinguishing "looped-out" and "stacked-in" DNA bulge conformation using fluorescent 2-aminopurine replacing a purine base. *J Biomol Struct Dyn.* 2002 Apr ;19(5) :929–934.
- [403] Breslauer KJ. Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions. *Methods Enzymol.* 1995 ;259 :221–242.
- [404] Nordlund TM, Xu D, Evans KO. Excitation energy transfer in DNA : duplex melting and transfer from normal bases to 2-aminopurine. *Biochemistry.* 1993 ;32(45) :12090–5.
- [405] O'Neill MA, Barton JK. 2-Aminopurine : a probe of structural dynamics and charge transfer in DNA and DNA :RNA hybrids. *J Am Chem Soc.* 2002 ;124(44) :13053–66.
- [406] Xu DG, Nordlund TM. Sequence dependence of energy transfer in DNA oligonucleotides. *Biophys J.* 2000 ;78(2) :1042–58.
- [407] Xu D, Evans KO, Nordlund TM. Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry.* 1994 ;33(32) :9592–9.
- [408] Kenfack CA, Piémont E, Gaiéd NB, Burger A, Mély Y. Time-resolved fluorescent properties of 8-vinyl-deoxyadenosine and 2-amino-deoxyribosylpurine exhibit different sensitivity to their opposite base in duplexes. *J Phys Chem B.* 2008 Aug ;112(32) :9736–9745. Available from : <http://dx.doi.org/10.1021/jp8028243>.
- [409] Berquand A, Mazeran PE, Pantigny J, Proux-Delrouyre V, Laval JM, Bourdillon C. Two-Step Formation of Streptavidin-Supported Lipid Bilayers by PEG-Triggered Vesicle Fusion. Fluorescence and Atomic Force Microscopy Characterization. *Langmuir.* 2003 Mar ;19(5) :1700–1707. Available from : <http://dx.doi.org/10.1021/la0260180>.
- [410] Didier P, Godet J, Mély Y. Two-photon two-focus fluorescence correlation spectroscopy with a tunable distance between the excitation volumes. *J Fluoresc.* 2009 May ;19(3) :561–565. Available from : <http://dx.doi.org/10.1007/s10895-008-0424-0>.

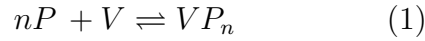
- [411] Sorzano COS, Thévenaz P, Unser M. Elastic registration of biological images using vector-spline regularization. *IEEE Trans Biomed Eng.* 2005 Apr;52(4) :652–663. Available from : <http://dx.doi.org/10.1109/TBME.2005.844030>.
- [412] R Development Core Team. *R : A Language and Environment for Statistical Computing*. Vienna, Austria; 2009. Available from : <http://www.R-project.org>.
- [413] Knight AE, Veigel C, Chambers C, Molloy JE. Analysis of single-molecule mechanical recordings : application to acto-myosin interactions. *Prog Biophys Mol Biol.* 2001;77(1) :45–72.
- [414] Carter BC, Vershinin M, Gross SP. A comparison of step-detection methods : how well can you do? *Biophys J.* 2008 Jan;94(1) :306–319. Available from : <http://dx.doi.org/10.1529/biophysj.107.110601>.
- [415] Qin F, Auerbach A, Sachs F. Hidden Markov modeling for single channel kinetics with filtering and correlated noise. *Biophys J.* 2000 Oct;79(4) :1928–1944. Available from : [http://dx.doi.org/10.1016/S0006-3495\(00\)76442-3](http://dx.doi.org/10.1016/S0006-3495(00)76442-3).
- [416] Qin F, Auerbach A, Sachs F. A direct optimization approach to hidden Markov modeling for single channel kinetics. *Biophys J.* 2000 Oct;79(4) :1915–1927. Available from : [http://dx.doi.org/10.1016/S0006-3495\(00\)76441-1](http://dx.doi.org/10.1016/S0006-3495(00)76441-1).
- [417] Andrec M, Levy RM, Talaga DS. Direct Determination of Kinetic Rates from Single-Molecule Photon Arrival Trajectories Using Hidden Markov Models. *J Phys Chem A.* 2003 Sep;107(38) :7454–7464. Available from : <http://dx.doi.org/10.1021/jp035514+>.
- [418] Yang H, Luo G, Karnchanaphanurach P, Louie TM, Rech I, Cova S, et al. Protein conformational dynamics probed by single-molecule electron transfer. *Science.* 2003 Oct ;302(5643) :262–266. Available from : <http://dx.doi.org/10.1126/science.1086911>.
- [419] McKinney SA, Joo C, Ha T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys J.* 2006 Sep;91(5) :1941–1951. Available from : <http://dx.doi.org/10.1529/biophysj.106.082487>.
- [420] Dahan M, Deniz A, Ha TJ, Chemla DS, Schultz PG, Weiss S. Ratiometric measurement and identification of single diffusing molecules. *Chemical Physics.* 1999;247 :85–106.
- [421] Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on.* 1967 Apr ;13(2) :260–269.
- [422] Levinson S, Rabiner R, Sondhi M. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell Systems Technical Journal.* 1983;62 :1035–1074.
- [423] Baum LE, Petrie T, Soules G, Weiss N. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics.* 1970;41(1) :164–171. Available from : <http://dx.doi.org/10.2307/2239727>.
- [424] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J Royal Statist Society B.* 1977;39 :1–38.
- [425] Bronson JE, Fei J, Hofman JM, Gonzalez RL, Wiggins CH. Learning rates and states from biophysical time series : a Bayesian approach to model selection and single-molecule FRET data. *Biophys J.* 2009 Dec;97(12) :3196–3205. Available from : <http://dx.doi.org/10.1016/j.bpj.2009.09.031>.
- [426] Liu Y, Park J, Dahmen KA, Chemla YR, Ha T. A comparative study of multivariate and univariate hidden Markov modelings in time-binned single-molecule FRET data analysis. *J Phys Chem B.* 2010 Apr ;114(16) :5386–5403. Available from : <http://dx.doi.org/10.1021/jp9057669>.
- [427] Akaike H. New look at statistical model identification. *IEEE Transactions On Automatic Control.* 1974;19 :716–723.
- [428] Schwarz G. Estimating dimension of a model. *Annals of Statistics.* 1978;6 :461–464.
- [429] Ha T, Ting AY, Liang J, Caldwell WB, Deniz AA, Chemla DS, et al. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc Natl Acad Sci U S A.* 1999 Feb ;96(3) :893–898.

- [430] Sabanayagam CR, Eid JS, Meller A. Using fluorescence resonance energy transfer to measure distances along individual DNA molecules : corrections due to nonideal transfer. *J Chem Phys.* 2005 Feb;122(6) :061103. Available from : <http://dx.doi.org/10.1063/1.1854120>.
- [431] Landes CF, Zeng Y, Liu HW, Musier-Forsyth K, Barbara PF. Single-molecule study of the inhibition of HIV-1 transactivation response region DNA/DNA annealing by argininamide. *J Am Chem Soc.* 2007 Aug;129(33) :10181–10188. Available from : <http://dx.doi.org/10.1021/ja071491r>.
- [432] Liu HW, Cosa G, Landes CF, Zeng Y, Kovaleski BJ, Mullen DG, et al. Single-molecule FRET studies of important intermediates in the nucleocapsid-protein-chaperoned minus-strand transfer step in HIV-1 reverse transcription. *Biophys J.* 2005 Nov;89(5) :3470–3479. Available from : <http://dx.doi.org/10.1529/biophysj.105.065326>.
- [433] Liu HW, Zeng Y, Landes CF, Kim YJ, Zhu Y, Ma X, et al. Insights on the role of nucleic acid/protein interactions in chaperoned nucleic acid rearrangements of HIV-1 reverse transcription. *Proc Natl Acad Sci U S A.* 2007 Mar;104(13) :5261–5267. Available from : <http://dx.doi.org/10.1073/pnas.0700166104>.
- [434] Vo MN, Barany G, Rouzina I, Musier-Forsyth K. Effect of Mg(2+) and Na(+) on the nucleic acid chaperone activity of HIV-1 nucleocapsid protein : implications for reverse transcription. *J Mol Biol.* 2009 Feb;386(3) :773–788. Available from : <http://dx.doi.org/10.1016/j.jmb.2008.12.073>.
- [435] Vo MN, Barany G, Rouzina I, Musier-Forsyth K. HIV-1 nucleocapsid protein switches the pathway of transactivation response element RNA/DNA annealing from loop-loop "kissing" to "zipper". *J Mol Biol.* 2009 Feb;386(3) :789–801. Available from : <http://dx.doi.org/10.1016/j.jmb.2008.12.070>.
- [436] Zeng Y, Liu HW, Landes CF, Kim YJ, Ma X, Zhu Y, et al. Probing nucleation, reverse annealing, and chaperone function along the reaction path of HIV-1 single-strand transfer. *Proc Natl Acad Sci U S A.* 2007 Jul;104(31) :12651–12656. Available from : <http://dx.doi.org/10.1073/pnas.0700350104>.
- [437] Wu T, Heilman-Miller SL, Levin JG. Effects of nucleic acid local structure and magnesium ions on minus-strand transfer mediated by the nucleic acid chaperone activity of HIV-1 nucleocapsid protein. *Nucleic Acids Res.* 2007;35(12) :3974–3987. Available from : <http://dx.doi.org/10.1093/nar/gkm375>.
- [438] Boudier C, Storchak R, Sharma KK, Didier P, Follenius-Wund A, Muller S, et al. The mechanism of HIV-1 Tat-directed nucleic acid annealing supports its role in reverse transcription. *J Mol Biol.* 2010 May; Available from : <http://dx.doi.org/10.1016/j.jmb.2010.05.033>.
- [439] Kant Sharma K, Didier P, Darlix JL, de Rocquigny H, Bensikaddour H, Lavergne JP, et al. Kinetic analysis of the nucleic acid chaperone activity of the hepatitis C virus core protein. *Nucleic Acids Res.* 2010 Jun;38(11) :3632–3642. Available from : <http://dx.doi.org/10.1093/nar/gkq094>.
- [440] Wang H, Yeh YS, Barbara PF. HIV-1 nucleocapsid protein bends double-stranded nucleic acids. *J Am Chem Soc.* 2009 Oct;131(42) :15534–15543. Available from : <http://dx.doi.org/10.1021/ja9070046>.
- [441] Rouzina I, Bloomfield VA. DNA bending by small, mobile multivalent cations. *Biophys J.* 1998 Jun;74(6) :3152–3164. Available from : [http://dx.doi.org/10.1016/S0006-3495\(98\)78021-X](http://dx.doi.org/10.1016/S0006-3495(98)78021-X).
- [442] Abbondanzieri EA, Bokinsky G, Rausch JW, Zhang JX, Grice SFJL, Zhuang X. Dynamic binding orientations direct activity of HIV reverse transcriptase. *Nature.* 2008 May;453(7192) :184–189. Available from : <http://dx.doi.org/10.1038/nature06941>.
- [443] Claudon P, Violette A, Lamour K, Decossas M, Fournel S, Heurtault B, et al. Consequences of Isostructural Main-Chain Modifications for the Design of Antimicrobial Foldamers : Helical Mimics of Host-Defense Peptides Based on a Heterogeneous Amide/Urea Backbone. *Angew Chem Int Ed Engl.* 2009 Dec; Available from : <http://dx.doi.org/10.1002/anie.200905591>.

Annexe

8.1 Modèle de la liaison de n ligands sur un récepteur

Le modèle développé ici correspond à l'équilibre suivant :



où n correspond au nombre de sites. Il a servi à modéliser l'interactions de peptides fluorescents avec des vésicules lipidiques illustré par la figure 8.1 [443].

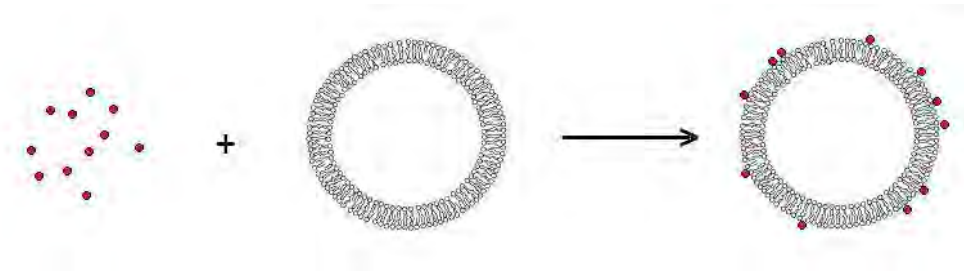


FIGURE 8.1 – Modèle de l'interaction de n peptides fluorescents (P) avec une vésicule lipidique (V)

On peut réécrire l'équilibre (1) selon le système :

$$\left\{ \begin{array}{l} (i) \quad P + V \rightleftharpoons VP \quad K_1 \\ (ii) \quad P + VP \rightleftharpoons VP_2 \quad K_2 \\ \dots \\ (n) \quad P + VP_{n-1} \rightleftharpoons VP_n \quad K_n \end{array} \right.$$

dans lequel on pose :

$$\forall i \geq 1, K_i = \frac{[VP_{i-1}][P]}{[VP_i]}$$

La constante globale de (1) :

$$K_a = \frac{[V][P]^n}{[VP_n]} = \prod_{i=1}^n K_i$$

Le taux de saturation des récepteurs ν s'écrit :

$$\nu = \frac{[P_{bound}]}{[V_{tot}]}$$

La loi de conservation de la matière impose que

$$P_{tot} = P + P_{bound} = P + \sum_{i=1}^n i (VP_i)$$

et

$$V_{tot} = V + V_{bound} = V + \sum_{i=1}^n (VP_i)$$

donc

$$\nu = \frac{\sum_{i=1}^n i (VP_i)}{V + \sum_{i=1}^n (VP_i)}$$

Si l'on considère que les n sites sont identiques :

$$K_a = \frac{n [V] [P]}{[VP]} = \frac{(n-1) [VP] [P]}{2 [VP_2]} = \dots = \frac{(n-i) [VP_i] [P]}{i [VP_{n+i}]} = \dots = \frac{[VP_{n-1}] [P]}{n [VP_n]}$$

On peut donc écrire :

$$V_{bound} = \sum_{i=1}^n (VP_i) = n \frac{[V] [P]}{K_a} + \frac{n(n-1)}{2} [V] \left(\frac{[P]}{K_a} \right)^2 + \dots + \frac{n!}{(n-1)!} [V] \left(\frac{[P]}{K_a} \right)^n$$

$$V_{tot} = V + V_{bound} = [V] \left\{ \sum_{i=1}^n C_n^i \left(\frac{[P]}{K_a} \right)^i \right\} = [V] \left(1 + \frac{[P]}{K_a} \right)^n$$

de même,

$$[P_{bound}] = n \frac{[V] [P]}{K_a} + \frac{n(n-1)}{2} [V] \left(\frac{[P]}{K_a} \right)^2 + \dots + \frac{n!}{(n-1)!} [V] \left(\frac{[P]}{K_a} \right)^n$$

d'où

$$[P_{bound}] = n [V] \frac{[P]}{K_a} \left\{ \sum_{i=1}^n C_{n-1}^i \left(\frac{[P]}{K_a} \right)^i \right\}$$

et

$$\sum_{i=1}^n i (VP_i) = n [V] \frac{[P]}{K_a} \left(1 + \frac{[P]}{K_a} \right)^{n-1}$$

On peut écrire ν en fonction de $[P]$ et K_d puisque :

$$\nu = \frac{n [V] \frac{[P]}{K_a} \left(1 + \frac{[P]}{K_a}\right)^{n-1}}{[V] \left(1 + \frac{[P]}{K_a}\right)^n} = \frac{n [P] / K_a}{1 + [P] / K_a} = \frac{n [P]}{K_d + [P]}$$

or

$$[P] = [P_{tot}] - [P_{bound}] = [P_{tot}] - \nu [V_{tot}]$$

donc :

$$\nu = \frac{n K_a \{[P_{tot}] - \nu [V_{tot}]\}}{1 + K_a \{[P_{tot}] - \nu [V_{tot}]\}} = \frac{n \{[P_{tot}] - \nu [V_{tot}]\}}{K_d + \{[P_{tot}] - \nu [V_{tot}]\}}$$

Réolvons ce polynome :

$$-K_d \times \nu - \{[P_{tot}] - \nu [V_{tot}]\} \times \nu + n \{[P_{tot}] - \nu [V_{tot}]\} = 0$$

$$\Leftrightarrow [V_{tot}] \nu^2 - (K_d + n [V_{tot}] + [P_{tot}]) \nu + n [P_{tot}] = 0$$

dont le discriminant vaut

$$\Delta = (K_d + n [V_{tot}] + [P_{tot}])^2 - 4n [P_{tot}] [V_{tot}]$$

soit finalement

$$\nu = \frac{(K_d + n [V_{tot}] + [P_{tot}]) - \{(K_d + n [V_{tot}] + [P_{tot}])^2 - 4n [P_{tot}] [V_{tot}]\}^{1/2}}{2 [V_{tot}]}$$

La formulation de ce taux de saturation peut être utilisée pour quantifier la force de l'interaction entre un ligand et un récepteur au cours d'expériences de titration basées sur l'intensité ou l'anisotropie de fluorescence.

Titration par intensité de fluorescence

La fluorescence du ligand marqué sous sa forme libre diffère de celle associée à sa forme liée au récepteur. Les concentrations en ligand libre et en ligand marqué sont liées de manière

linéaire à l'intensité de fluorescence mesurée I selon :

$$\nu = \frac{I - I_0 [P_{tot}]}{I_f - I_0 [V_{tot}]}$$

où I_0 et I_f représentent les fluorescences associées à la forme libre et liée du ligand, respectivement.

Cette relation mène tout naturellement à :

$$I = I_0 + (I_f - I_0) \times \nu \times \frac{[V_{tot}]}{[P_{tot}]}$$

ou encore

$$I = I_0 + (I_f - I_0) \frac{(K_d + n [V_{tot}] + [P_{tot}]) - \sqrt{(K_d + n [V_{tot}] + [P_{tot}])^2 - 4n [P_{tot}] [V_{tot}]}}{2 [P_{tot}]}$$

Titration par anisotropie de fluorescence

De même que pour la fluorescence, on observe

$$\nu = \frac{r - r_0 [P_{tot}]}{r_f - r_0 [V_{tot}]}$$

Dans le cas où la fluorescence de la forme liée diffère significativement de la forme libre, il convient de pondérer le signal d'anisotropie r_i par sa fluorescence f_i puisque :

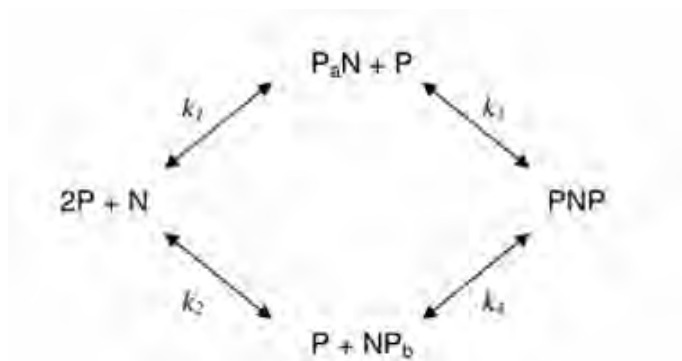
$$r = \sum_i r_i \cdot f_i$$

En pratique, on utilise le ratio R des fluorescences des formes libres et liées. On obtient alors :

$$r = \frac{r_0 + (r_f \cdot R - r_0) \nu [V_{tot}] / [P_{tot}]}{1 + (R - 1) \nu [V_{tot}] / [P_{tot}]}$$

8.2 Liaison de deux ligands sur un récepteur à deux sites de liaisons distincts

Ce modèle a été utilisé dans Avilov & al. [256] afin de déterminer les affinités de la NCp7 pour deux sites de liaisons différents au sein d'un même dodécanucléotide.



La loi d'action de masse permet de définir :

$$k_1 = \frac{[P_aN][P]}{[N][P]^2} = \frac{[P_aN]}{[N][P]}; \quad k_3 = \frac{[PNP]}{[P_aN][P]}; \quad k_1k_3 = \frac{[PNP]}{[N][P]^2}$$

$$k_2 = \frac{[NP_b][P]}{[N][P]^2} = \frac{[NP_b]}{[N][P]}; \quad k_4 = \frac{[PNP]}{[NP_b][P]}; \quad k_2k_4 = \frac{[PNP]}{[N][P]^2}$$

La loi de conservation de la matière impose :

$$[N_{tot}] = [N] + [P_aN] + [NP_b] + [PNP]$$

$$[P_{tot}] = [N] + [P_aN] + [NP_b] + 2[PNP]$$

Si l'on suppose que $k_1 = k_4$ et $k_2 = k_3$, la variation de fluorescence du site *a* traduit la variation du rapport $\frac{[P_aN]+[PNP]}{[N_{tot}]}$ donc :

$$\frac{I_a - I_{a0}}{I_{af} - I_{a0}} = \Delta I_a = \frac{[P_aN] + [PNP]}{[N_{tot}]}$$

$$\Delta I_a = \frac{k_1 [P][N] + k_1k_3 [P]^2 [N]}{[N] + [P_aN] + [NP_b] + [PNP]}$$

$$\Delta I_a = \frac{k_1 [P] + k_1 k_3 [P]^2}{1 + (k_1 + k_2) [P] + (k_1 k_3) [P]^2}$$

de même pour le site b

$$\Delta I_b = \frac{[NP_b] + [PNP]}{[N_{tot}]}$$

$$\Delta I_b = \frac{k_2 [P] + k_1 k_3 [P]^2}{1 + (k_1 + k_2) [P] + (k_1 k_3) [P]^2}$$

On retrouve bien

$$[P_{tot}] = [P] + [N_{tot}] (\Delta I_a + \Delta I_b)$$

Enfin, si la variation de fluorescence traduit la liaison simultanée des deux sites :

$$\Delta I = \Delta I_a + \Delta I_b = \frac{[P_a N] + [NP_b] + [PNP]}{[N_{tot}]}$$

$$\Delta I = \frac{(k_1 + k_2) [P] + k_1 k_3 [P]^2}{1 + (k_1 + k_2) [P] + (k_1 k_3) [P]^2}$$

Autres travaux

9.1

Publication VIII

Solution structures of 3 nuclear receptors heterodimers RXR/RAR, RXR/VDR and RXR/PPAR on DNA

Natacha Rochel¹, Fabrice Ciesielski¹, Julien Godet², Elmiro Moman¹, Manfred Roessle³, Carole Peluso-Iltis¹, Yves Mély², Dmitri I Svergun^{3,4}, Dino Moras^{1*}

¹ Département de Biologie et de Génomique Structurales, IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire) (UMR7104 CNRS, U964 INSERM, UDS), ² Laboratoire de Biophotonique et Pharmacologie, Faculté de Pharmacie, UMR 7213 du CNRS, Université de Strasbourg, 67401 Illkirch, France, ³ European Molecular Biology Laboratory, Hamburg Outstation, 22603 Hamburg, Germany, ⁴ Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia.

* *Corresponding author:* D. Moras, Laboratoire de Biologie et Génomique Structurales, IGBMC, 1 rue Laurent Fries, BP 10142, Illkirch 67404, France. Tel. : +33 388 653 220; Fax: +33 388 653 276; E-mail: moras@igbmc.fr

This study was supported by CNRS, INSERM, Université de Strasbourg, the European Commission Structural Proteomics in Europe (SPINE) (QLG2-CT-220-0098) and SPINE2-Complexes (LSHG-CT-2006-031220) under the integrated program, Quality of Life and Management of Living Resources and by ANR (ANR-06-PCVI-0004-02; NRcoactiv). DI Svergun and M Roessle acknowledge financial support from the EU Framework 6 Programme (Design Study SAXIER, RIDS 011934). F Ciesielski is an ARC fellowship recipient.

Abstract

Nuclear receptors (NRs) control numerous physiological processes through the regulation of gene expression. The present study provides a structural basis for understanding the role of DNA in the spatial organization of NRs heterodimers into functional transcription complexes with coactivators. Using Small Angle X-ray Scattering (SAXS) and Fluorescence Resonance Energy Transfer (FRET) techniques, we have determined the solution structures of three functional heterodimers RXR/RAR, RXR/PPAR and RXR/VDR bound to natural hormone response elements (HREs) forming direct repeats (DR1 or DR5, and DR3) in presence or absence of coactivator domains. DNA-bound complexes adopt a clearly defined conformation specific for each HRE. The structures exhibit a common feature: an extended asymmetric shape with distinct modules containing DNA:DBDs (DNA Binding Domain) and ligand binding domains (LBDs) respectively, the LBD of RXR's partners pointing to the 5' end of DNA. Only one coactivator molecule binds each heterodimer through RXR' partner NR. The structures enlighten the role of HREs and hinge domains in the orientation of LBDs and reveal an unexpected constant in the spatial positioning of cofactors.

9.2

Publication IX

Helical Foldamers

Consequences of Isostructural Main-Chain Modifications for the Design of Antimicrobial Foldamers: Helical Mimics of Host-Defense Peptides Based on a Heterogeneous Amide/Urea Backbone**

Paul Claudon, Aude Violette, Karen Lamour, Marion Decossas, Sylvie Fournel, Béatrice Heurtault, Julien Godet, Yves Mély, Brigitte Jamart-Grégoire, Marie-Christine Averlant-Petit, Jean-Paul Briand, Guy Duportail, Henri Monteil, and Gilles Guichard*

The design of unnatural oligomers with predictable folding patterns (i.e. foldamers) and function has attracted considerable attention over the last ten years.^[1] Applications in biology include the development of antimicrobial agents, cell-penetrating agents, and inhibitors of protein-protein interactions.^[1a-c] Non-natural oligoamides built from α -amino acid residues (e.g. peptoids) and higher homologues (e.g. β -, γ -, and δ -peptides) are the epitomes of peptidomimetic foldamers.^[1,2] A substantial increase in the number of backbones with folding propensity came from the exploration of the β - and γ -peptide families (i.e. by introduction of isosteric or isoelectronic backbone modifications).^[1a] The substitution of

a urea moiety for the $\text{CH}_2\text{-CO-NH}$ units in the γ^1 -peptide backbone is an interesting case of quasi-isostructural replacement.^[3] High-resolution structural studies in solution^[3] and in the solid state^[4] have shown that the oligoureia backbone adopts a remarkably stable helical fold reminiscent of the 14-helix described for the cognate γ -peptides (Figure 1).^[3,4]

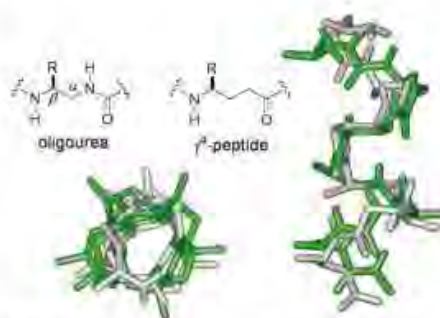


Figure 1. Superimposition of helical structures formed by oligoureias (green)^[4] and γ^1 -peptides (gray).^[3]

How such quasi-isostructural backbones^[4] compare in molecular recognition is presently unknown. We reported previously that oligoureias designed to mimic globally amphiphilic α -helical host-defense peptides display broad antibacterial activity with selectivity for prokaryotic versus mammalian red blood cell membranes.^[7] To better understand structure–property relationships and to gain insight into the mechanisms of membrane disruption, we have now undertaken detailed comparative studies of oligoureias, their γ -peptide counterparts, and various mixed amide/urea congeners (**1–8**) bearing identical side chains (Scheme 1).

Oligoureia **1** was found previously to display both 1) a strong helix propensity in a lipidic environment, and 2) significant antibacterial activity against Gram negative and Gram positive bacteria.^[7] However, further development of **1** was limited by synthetic hurdles. Stepwise elongation of **1** on a Rink resin using monomers protected with *N*-fluorenylmethoxycarbonyl (Fmoc) groups^[3a,7,8] resulted in poor overall yields. The purity of crude product based on C_{18} reverse-phase HPLC was only 17%.^[9] Although chemistry using building blocks protected with the *N*-tert-butoxycarbonyl (Boc) group is more robust, it was found that the urea linkage formed by anchoring the first residue on a 4-methylbenzhydramine

[*] P. Claudon, Dr. A. Violette, K. Lamour, Dr. M. Decossas, Prof. S. Fournel, Dr. J.-P. Briand, Dr. G. Guichard^[1] CNRS, Institut de Biologie Moléculaire et Cellulaire Laboratoire d'Immunologie et Chimie Thérapeutiques 15 rue René Descartes, F-67000 Strasbourg (France) E-mail: g.guichard@iecb.u-bordeaux.fr

Dr. B. Heurtault
Laboratoire de chimie enzymatique et vectorisation
Université Strasbourg 1, Institut Gilbert Laustriat CNRS - UMR 7175, Faculté de Pharmacie
74 route du Rhin, BP 60024, 67401 Illkirch Cedex (France)

Prof. B. Jamart-Grégoire, Dr. M.-C. Averlant-Petit
LCPM, UMR 7568 CNRS, Nancy-Université
7 rue Grandville BP20451, Nancy (France)

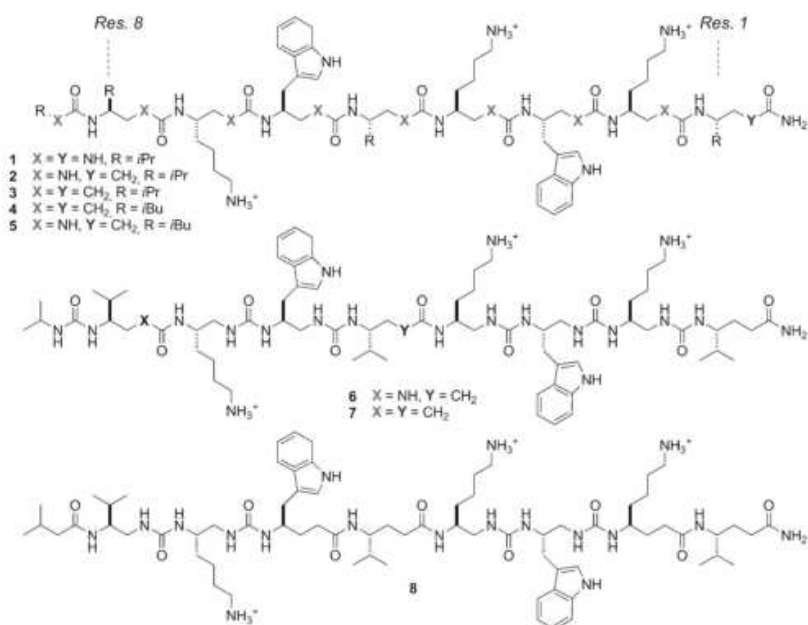
J. Godet, Prof. Y. Mély, Dr. G. Duportail
Laboratoire de Biophotonique et Pharmacologie
CNRS UMR 7213, Faculté de Pharmacie, Université de Strasbourg
BP 60024, 67401 Illkirch Cedex (France)

Prof. H. Monteil
Laboratoire de Physiopathologie et d'Antibiologie des Infections Nosocomiales et Emergentes
Institut de Bactériologie, Faculté de Médecine
3 rue Koeberlé, 67000 Strasbourg (France)

[**] Present address: Institut Européen de Chimie et Biologie
Université de Bordeaux—CNRS UMR 5248, CBMN
2 rue R. Escarpit, 33607 Pessac (France)

[***] This research was supported in part by the CNRS, the Région Alsace and ImmuPharma France and by grants from ANR (project no. 05 PRIB 02-903 and NT05-4-42848). Fellowships from ImmuPharma France, ANRT, and Région Alsace (P.C., A.V., K.L.) are gratefully acknowledged. We thank Daniele Spohner for her help in using the EM-PACT2 apparatus. Access to the RIO Imaging Facility (Plateforme de Microscopie et d'Imagerie Cellulaire Strasbourg Esplanade) is gratefully acknowledged.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.200905591>.



Scheme 1. Globally amphiphilic oligoureas, γ^1 -peptides, and related hybrids 1–8.

(MBHA) resin or the less acid-sensitive benzhydrylamine (BHA) resin does not resist the acidic conditions (trifluoroacetic acid or 3 N HCl in dioxane) required to cleave the Boc group. To overcome this difficulty, we envisaged the preparation of oligoureas bearing a stable terminal amide group by direct attachment of an isosteric γ^1 -amino acid on MBHA resin. This strategy, together with the use of *N*-Boc-protected monomers, proved to be very effective, with purities of crude **2** routinely over 60% (Figure S1 in the Supporting Information).^[9] It is noteworthy that neither 2.5-helix propensity^[34] (Figure S2 and S3 in the Supporting Information)^[9] nor antibacterial activity were affected by this point mutation. The minimum inhibitory concentrations (MICs) obtained for **2** were very similar to those found for melittin, a honeybee toxin, and omiganan,^[10] a cationic peptide currently in phase III clinical trials for the prevention of catheter-related bloodstream infections (Table 1). Whereas omiganan exhibits weak bactericidal activity (MBCs > 256 $\mu\text{g mL}^{-1}$), the minimum bactericidal concentrations (MBCs) measured for **2** were equal to or twice as large as corresponding MICs. Moreover, **2** was found to be equally potent on methicillin-sensitive (ATCC 25923) and methicillin-resistant (clinical IBS 45-501) *S. aureus* strains.

To determine whether the urea backbone is critical for antibacterial activity, we synthesized **3**, the γ^1 -peptide analogue of **2**. Antibacterial activity of γ -peptide **3** was determined against *E. coli*, *P. aeruginosa*, and *S. aureus*. γ -Peptide **3**, although intrinsically less polar than **2**, was virtually inactive (MIC and MBC > 256 $\mu\text{g mL}^{-1}$) on both Gram positive and Gram negative bacteria. From studies with antibacterial α - and β -peptides,^[11] we reasoned that increasing the hydrophobic content of **3** could increase its potency, albeit

at the cost of mammalian versus bacterial cell selectivity. However, substituting *i*Bu for *i*Pr side chains in **3** (to give **4**) led only to a modest improvement (MIC = 128 $\mu\text{g mL}^{-1}$ on the three bacterial strains). In contrast, the same modification on **2** gave the most potent antibacterial oligourea (**5**) reported to date with MIC and MBC values as low as 8 and 32 $\mu\text{g mL}^{-1}$ on *S. aureus*, respectively.

We next examined the susceptibility of host cell membranes to oligoureas and γ -peptides by monitoring lysis of sheep red blood cells. Oligourea **2** ($\text{HD}_{100} = 256 \mu\text{g mL}^{-1}$) compared favorably with melittin, which is strongly toxic towards red blood cells ($\text{HD}_{100} = 32 \mu\text{g mL}^{-1}$). γ^1 -Peptides **3** and **4** display the highest HD_{100} values.

With the aim to investigate further the effects of insertion of γ^1 -amino acid residues on 2.5-helix propensity and antimicrobial activity, we prepared mixed oligourea/ γ^1 -peptides **6–8** containing two to four γ^1 -amino acid residues. Information about the conformational preferences of oligomers **6–8** was gained by monitoring chemical shift differences ($\Delta\delta$) in the ¹H NMR spectra between diastereotopic ¹³C protons of “diaminoethylene” residues (Figures S4–S6 and Table S1 in the Supporting Information)^[9] and by circular dichroism (CD) spectroscopy (Figure S7 in the Supporting Information),^[9] as described previously.^[3] Comparison with oligomer **2** revealed several trends. The insertion of an

Table 1: Antibacterial and haemolytic activities of compounds **2–8** compared to melittin and omiganan.^[a]

| Compound | Bacteria | | | | | | HD_{100} ^[b] |
|----------|--------------------------------|------|--------------------------------------|------|----------------------------------|-------------------|----------------------------------|
| | <i>E. coli</i> (ATCC 25922) | | <i>P. aeruginosa</i> (ATCC 27853) | | <i>S. aureus</i> (ATCC 25923) | | |
| | MIC | MBC | MIC | MBC | MIC | MBC | |
| omiganan | 64 | >256 | 128 | >256 | 32 | >256 | nt ^[c] |
| melittin | 64 | 64 | 128 | 256 | 32 | 64 | 32 |
| 2 | 32 | 32 | 64 | 128 | 32 ^[d] | 64 ^[d] | 256 |
| 3 | >256 | >256 | 256 | >256 | >256 | >256 | >512 |
| 4 | 128 | >256 | 128 | >256 | 128 | >256 | >512 |
| 5 | 16 | 32 | 16 | 64 | 8 | 32 | 64 |
| 6 | 64 | 128 | 32 | 128 | 16 | 64 | 512 |
| 7 | 32 | 128 | 64 | 128 | 32 | 128 | nt |
| 8 | 128 | >256 | 128 | 256 | 128 | >256 | nt |

[a] MIC is the minimum inhibitory concentration and MBC the minimum bactericidal concentration. MIC, MBC, and haemolysis in $\mu\text{g mL}^{-1}$. [b] HD_{100} values are concentrations required to lyse 100% of the cells after 48 h. [c] nt = not tested. [d] Identical MIC and MBC values were obtained on a clinical isolated methicillin-resistant *S. aureus* strain (IBS 45-501).

additional γ^L -amino residue at positions P5 and P8 has a local destabilizing effect on the preceding residues (i.e. residues at P4 and P7), but it does not seem to alter the overall helical conformation. In contrast, the insertion of two γ^L -amino acid residue pairs such as in **8** has a more pronounced effect on the overall structure, with significantly reduced $\Delta\delta$ values. CD spectroscopy experiments show that the positive molar ellipticity at 203 nm per urea moiety decreases significantly as a result of γ^L -amino acid insertion, which may be interpreted as 2.5-helix destabilization.^[9,12] Similarly, antibacterial activity was found to decrease with the number of γ -amino acid residues in the backbone. However, octamer **6**, with a central γ^L -Val residue, compares favorably with **2** in terms of both bactericidal activity and cytotoxicity against mammalian cells. This increased selectivity makes mixed amide/urea oligomer **6** an interesting lead compound for future studies.

Oligomers **6–8** were also evaluated against three ATCC isolates of *Candida*, namely *C. krusei* ATCC 6258, *C. parapsilosis* ATCC 22019, *C. albicans* ATCC 90028.^[13] Again, oligourea **6** was the most active compound in the series, with MICs of 16–64 $\mu\text{g mL}^{-1}$ and minimal fungicidal concentrations (MFCs) against all three *Candida* isolates either equal to or twice as large as the MICs (Table S2 in the Supporting Information).^[9]

Evidence for a direct interaction between oligoureas and bacterial membranes was gained from immunoelectron microscopy experiments using postembedding immunogold labeling experiments.^[9] The ultrastructural localization of the biotinylated analogue of **2** (biot-**2**) on Gram negative bacteria (*E. coli*) reveals the presence of small and well-defined clusters in the vicinity of the inner and outer bacterial membranes (Figure 2). These results may suggest that oligoureas either alone or as aggregates are able to cross bacterial membranes.

To determine whether the functional difference between oligourea and oligoamide backbones results from differential membrane disruption activities, we have undertaken physicochemical investigations using negatively charged phospholipid membranes as model systems. Both oligourea **2** and γ^L -peptide **3** carry indole side chains at positions 3 and 6 in their sequence, thus endowing the molecules with intrinsic fluorescence. We first studied the interaction of **2** and **3** with large unilamellar vesicles (LUV) of egg yolk phosphatidylcholine (EYPC) and egg yolk phosphatidylglycerol (EYPG) in a 70:30 molar ratio by measuring indole emission fluorescence and steady-state fluorescence anisotropy upon addition of an increasing concentration of phospholipids. At a lipid-to-oligomer molar ratio $R_1 = 20$, shifts to lower wavelength (blue shift) of the indole emission peak (λ_{max}) of 23 nm and 19 nm were observed for **2** and **3**, respectively (Figure S8 in the Supporting Information),^[9] thus indicating significant penetration of the indole moieties into the hydrophobic lipid bilayer. The restricted mobility of the indole moiety upon interaction with the membrane was also inferred from an increase in fluorescence anisotropy. Titrations of **2** and **3** with EYPC/EYPG LUVs (Figure S9 in the Supporting Information) were used to quantify the binding of both oligomers to vesicles.^[9] The apparent dissociation constants ($1.4 \times$



Figure 2. Immunoelectron microscopy (biotin, 6 nm) of biot-**2** on *E. coli* after high-pressure freezing and freeze-substitution. Scale bar 50 nm; inset 10 nm.

$10^{-7} \text{ mol L}^{-1}$ for **2** and $9 \times 10^{-7} \text{ mol L}^{-1}$ for **3**) calculated by curve fitting using the binding model given in Equation (13) in the Supporting Information indicate that γ -peptide **3** has a sixfold lower affinity for the membrane than oligourea **2**.

To assess the extent to which peptides **2** and **3** disrupt phospholipid membranes, we measured their abilities to induce efflux of carboxyfluorescein from EYPC/EYPG (70:30 molar ratio) LUVs. The time course of the leakage of encapsulated carboxyfluorescein was detected by the corresponding increase in fluorescence intensity at 518 nm (Figure 3). Whereas γ -peptide **3** was not able to induce significant efflux of carboxyfluorescein at 4.16 μM even after prolonged time, very fast and strong carboxyfluorescein release was observed when LUVs were treated with **2** at the same concentration. This series of experiments further confirms the different membrane-disruption properties exhibited by oligoamide and oligourea backbones.

A number of recent studies have shown that conformational preorganization is not necessarily a prerequisite to design oligomers with antimicrobial activity.^[14,15] Thus, it is unlikely that the difference in helix stability between oligoamides and oligoureas alone accounts for the observed dichotomy in membrane interaction and antibacterial activities. Instead, helix geometry (side-chain projection) and backbone polarity are two factors that vary substantially between the two oligomeric systems and that need to be considered. Moreover, it is tempting to speculate that structural features of helical oligoureas revealed by X-ray diffraction studies,^[4] that is, conformational adaptability (helix plasticity) and directional aggregation, may have

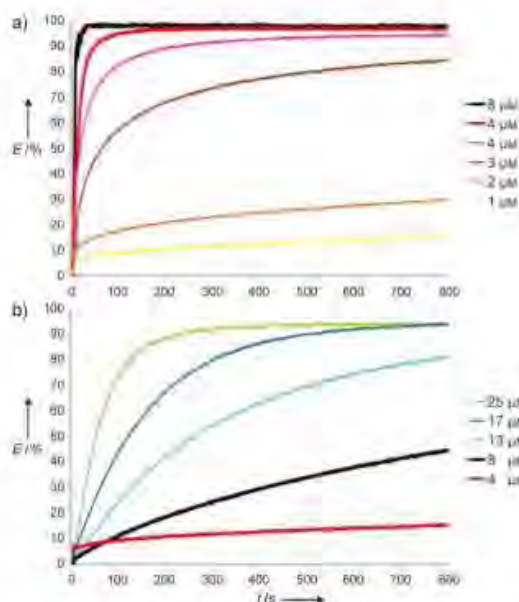


Figure 3. Time course of carboxyfluorescein efflux from LUV EYPC/EYPC (70:30 molar ratio) upon addition of increasing concentrations of a) **2** and b) **3**.

some relevance to the unique antimicrobial activities of oligoureas and related hybrids.

In conclusion, we have documented a case of dichotomy in the foldamer γ -peptide lineage whereby two apparently isostructural helical backbones (i.e. oligoamide and oligourea) bearing identical side chains appear to strongly differ in their antibacterial and biomolecular recognition properties. Our results also point to heterogeneous helical urea/amide backbones, which may become advantageous in the development of more potent yet less cytotoxic antimicrobial helical foldamers for in vivo applications.^[16]

Received: October 7, 2009

Published online: December 2, 2009

Keywords: amphiphiles · antimicrobial agents · helical structures · peptidomimetics · structure–activity relationships

[1] a) *Foldamers: Structure, Properties and Applications* (Eds.: S. Hecht, I. Huc), Wiley-VCH, Weinheim, 2007; b) C. M. Good-

man, S. Choi, S. Sandler, W. F. DeGrado, *Nat. Chem. Biol.* 2007, 3, 252–262; c) D. Seebach, A. K. Beck, D. J. Bierbaum, *Chem. Biodiversity* 2004, 1, 1111–1239; d) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* 2001, 101, 3893–4012; e) S. H. Gellman, *Acc. Chem. Res.* 1998, 31, 173–180.

[2] a) A. D. Bautista, C. J. Craig, E. A. Harker, A. Schepartz, *Curr. Opin. Chem. Biol.* 2007, 11, 685–692; b) A. Czyżewski, A. Barron, *AIChE J.* 2008, 54, 2–8.

[3] a) G. Guichard, A. Violette, G. Chassaing, E. Miclet, *Magn. Reson. Chem.* 2008, 46, 918–924; b) L. Fischer, C. Didierjean, F. Jobibois, V. Semetey, J. Manuel Lozano, J. P. Briand, M. Marraud, R. Potcau, G. Guichard, *Org. Biomol. Chem.* 2008, 6, 2596–2610; c) A. Violette, N. Lancelot, A. Poschako, M. Piotto, J.-P. Briand, J. Rayu, K. Elbayed, A. Bianco, G. Guichard, *Chem. Eur. J.* 2008, 14, 3874–3882; d) A. Violette, M. C. Averlant-Petit, V. Semetey, C. Hemmerlin, R. Casimir, R. Graff, M. Marraud, J.-P. Briand, D. Rognan, G. Guichard, *J. Am. Chem. Soc.* 2005, 127, 2156–2164; e) C. Hemmerlin, M. Marraud, D. Rognan, R. Graff, V. Semetey, J.-P. Briand, G. Guichard, *Helv. Chim. Acta* 2002, 85, 3692–3711.

[4] L. Fischer, P. Claudon, N. Pendem, J.-P. Briand, E. Miclet, C. Didierjean, E. Ennifar, G. Guichard, *Angew. Chem.* 2009, DOI: 10.1002/ange.200905592; *Angew. Chem. Int. Ed.* 2009, DOI: 10.1002/anie.200905592.

[5] a) T. Hintermann, K. Gademann, B. Jain, D. Seebach, *Helv. Chim. Acta* 1998, 81, 983–1002; b) S. Hanessian, X. Luo, R. Schaum, S. Michnick, *J. Am. Chem. Soc.* 1998, 120, 8569–8570.

[6] Average (ϕ , θ_1 , θ_2 , ψ) backbone torsion angles ($^\circ$) for oligourea (Ref. [4]) and γ^2 -peptide (Ref. [5a]) helices are: (–103, +57, +80, –170) and (–127, +66, +64, –140), respectively.

[7] A. Violette, S. Fournel, K. Lamour, O. Chaloin, B. Frisch, J.-P. Briand, H. Monteil, G. Guichard, *Chem. Biol.* 2006, 13, 531–538.

[8] G. Guichard, V. Semetey, M. Rodriguez, J.-P. Briand, *Tetrahedron Lett.* 2000, 41, 1553–1557.

[9] See the Supporting Information.

[10] H. S. Sader, K. A. Fedler, R. P. Rennie, S. Stevens, R. N. Jones, *Antimicrob. Agents Chemother.* 2004, 48, 3112–3118.

[11] D. Liu, W. DeGrado, *J. Am. Chem. Soc.* 2001, 123, 7553–7559.

[12] As mentioned by a referee, care must be taken in comparing CD spectra of oligoureas and related amide/urea hybrids (which contain two different, albeit related, chromophores in their backbone) by using the dichroic maximum at 203 nm.

[13] For recent work by Gellman on cationic 14-helical β -peptides with antifungal activity, see: A. J. Karlsson, W. C. Pomerantz, B. Weisblum, S. H. Gellman, S. P. Palecek, *J. Am. Chem. Soc.* 2006, 128, 12630–12631.

[14] a) B. P. Mowery, S. E. Lee, D. A. Kissounko, R. F. Epand, R. M. Epand, B. Weisblum, S. S. Stahl, S. H. Gellman, *J. Am. Chem. Soc.* 2007, 129, 15474–15476; b) T. L. Raguse, E. A. Porter, B. Weisblum, S. H. Gellman, *J. Am. Chem. Soc.* 2002, 124, 12774–12785.

[15] N. Papo, Y. Shai, *Biochemistry* 2004, 43, 6393–6403.

[16] The potential of heterogeneous peptide backbone foldamers for biomedical applications has been highlighted recently: W. S. Horne, S. H. Gellman, *Acc. Chem. Res.* 2008, 41, 1399–1408.

Absorption and Fluorescence Studies

Concentration determination.

Concentration of oligourea and γ -peptide stock solutions were determined by UV absorbance measurements at 280 nm, using indole molar extinction coefficient $\epsilon_{280} = 5700 \text{ Mol}^{-1} \text{ cm}^{-1}$. Absorption spectra were performed on a Cary 400 spectrophotometer (Varian Inc., Palo Alto, USA)

Vesicles preparation. The phospholipids egg yolk L- α -phosphatidylcholine (EYPC) and L- α -phosphatidyl-DL-glycerol (EYPG) were purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France). A mixture of lipids (EYPC/EYPG 70:30 molar ratio) in chloroform was evaporated and dried under vacuum, using a rotary evaporator. The lipid film was then hydrated with the appropriate buffer, vigorously vortexed during 2 min and then extruded by the classical extrusion method^[2] through 0.2 μm and 0.1 μm Nuclepore polycarbonate filters (Whatman, Maidstone, UK) to finally obtain large unilamellar vesicles (LUVs) of $107 \pm 5 \text{ nm}$ diameter, as determined by dynamic light scattering.

When a gel filtration was needed, the vesicle solution concentration, expressed in lipids, was determined by enzymatic PC titration using phospholipid B test WACO COF (OXOID [Dardilly, France]). All the liposomal formulations were freshly prepared before use.

To calculate the vesicle concentration, the external radius of a vesicle (R) was considered to be 535 \AA , as determined by DLS measurement. The thickness of the lipidic bilayer (t) and the average lipidic density (d) were assumed to be 40 \AA and 70 $\text{\AA}^2/\text{lipid}$, respectively.^[3]

² M. J. Hope, M. B. Bally, G. Webb, P. R. Cullis. *Biochim. Biophys. Acta* **1985**, *812*, 55-65

³ B. A. Lewis, D. M. Engelman, *J. Mol. Biol.* **1983**, *166*, 211-217.

The number of lipids per vesicle was thus determined as :

$$n = 4\pi \frac{(R^2 + (R-t)^2)}{d} = 9,54.10^4 \text{ lipids/vesicle}$$

It was used to infer the vesicle concentration from the lipid concentration

Anisotropy measurements. For these experiments, a 10 mM HEPES, 150 mM NaCl, 3 mM EDTA buffer adjusted to pH 7.4 was used.

Steady-state fluorescence anisotropy,

$$r = (I_{//} - I_{\perp}) / (I_{//} + 2I_{\perp}) \quad (1)$$

was measured with a SLM-8000 spectrofluorimeter (SLM-Aminco, Urbana-Champaign, IL, USA) in the T-format. Excitation was set at 280 nm and the emitted light was monitored through a 342 nm interference filter (Schott, Mainz, Germany). A home-built device ensured automatic rotation of the excitation polarizer, allowing a continuous average measurement of r which includes the G correction factor.

The fluorescence emission spectra were recorded in the wavelength range 300-450 nm with a FluoroMax-3 spectrofluorimeter (Jobin-Yvon Horiba, Longjumeau, France).

For the titration, a 10 μ M oligomer solution was concentrated in lipid by stepwise addition of 200 μ M lipid vesicle solution containing the oligomer at 10 μ M, in order to maintain its concentration constant for all measurements. Both fluorescence emission spectrum and fluorescence anisotropy were recorded for each intermediate solution. The two extreme solutions (without vesicles and with vesicles in excess) were used to determine the two limiting anisotropies (r_0 and r_x) and the ratio of fluorescence intensities (R) of the bound oligomer to the free oligomer in the conditions of fluorescence anisotropy measurements, which means in presence of the 342nm filter.

This determination of R was performed for both the oligourea (R = 2.88) and its gamma peptide equivalent (R = 2.01).

The experimental titration curves (r versus lipid concentration) were fitted assuming an equilibrium reaction for the binding of oligomers with the lipid bilayer of the vesicles, to provide the vesicle-oligomer dissociation constant as well as the number n of oligomer (P) per vesicle (V).



The macroscopic dissociation constant K_d can be defined as:

$$K_d = \frac{[V][P]^n}{[VP_n]} \quad (3)$$

Assuming that all binding sites are independent and equivalent, it is easier to consider the microscopic dissociation constant k which is related to a single binding site, all other sites being identical

$$K_d = k^n \quad (4)$$

with

$$k = \frac{n[V][P]}{[VP]} = \frac{(n-1)[VP][P]}{2[VP_2]} = \dots = \frac{[VP_{n-1}][P]}{n[VP_n]} \quad (5)$$

for the different consecutive equilibria.

By taking into account the first term of the previous equation, and by introducing the total oligomer concentration $[P_T]$ (fixed) and the total vesicle concentration $[V_T]$ (varying) which can be written as:

$$P_T = [P_{\text{free}}] + [P_{\text{bound}}] = P + \sum_{i=1}^n i * [VP_i] \quad (6)$$

$$V_T = [V_{\text{free}}] + [V_{\text{bound}}] = V + \sum_{i=1}^n [VP_i]$$

A mathematical development directly leads to the saturation extent of the vesicles :

$$Y = \frac{[P_{\text{bound}}]}{V_T} = \frac{n[P_{\text{free}}]}{K_d + [P_{\text{free}}]} \quad (7)$$

which can be expressed as a function of V_T :

$$Y = \frac{(K_d + nV_T + P_T) - \sqrt{(K_d + nV_T + P_T)^2 - 4nV_T P_T}}{2V_T} \quad (8)$$

The additive law of anisotropy provides the following equation :

$$r = f_{free} r_0 + f_{bound} r_\infty \quad (9)$$

where f_{free} and f_{bound} represent the fraction of the fluorescence intensity emitted by the free and bound forms of oligomer.

By definition:

$$f_{free} + f_{bound} = 1 \quad (10)$$

Since fluorescence intensities are not the same for both the free and the bound form of oligomer (see fluorescence spectra in Figure S4), we have to consider the ratio R of the fluorescence intensities of both forms in the conditions of the fluorescence anisotropy measurements:

$$R = \frac{I_\infty}{I_0} \quad (11)$$

The anisotropy r can be expressed as :

$$r = \frac{r_0 [P_{free}] + r_\infty R [P_{bound}]}{[P_{free}] + R [P_{bound}]} \quad (12)$$

By using equation (7) we obtain the equation used for the fitting of the experimental titration curves:

$$r = \frac{r_0 + Y \frac{V_T}{P_T} (R r_\infty - r_0)}{1 + Y \frac{V_T}{P_T} (R - 1)} \quad (13)$$

Leakage Experiments. For these experiments, LUVs were prepared in 10mM HEPES, 10mM NaCl, 1mM EDTA, pH 7.4 buffer and containing carboxyfluorescein(CF) (5,6-carboxyfluorescein from Eastman-Kodak, New York, USA, purified on a Sephadex LH-20

column from Pharmacia, Uppsala, Sweden)^[4] at a self-quenching concentration of 50 mM. Non-encapsulated CF was separated from the vesicle suspension through a Sephadex G-75 filtration column (Pharmacia) eluted with an isoosmolar buffer containing 10mM HEPES, 107mM NaCl, 1mM EDTA, pH 7.4. Oligourea and γ -peptide leakage kinetics were measured using the same preparation of CF loaded vesicles.

Leakage kinetics was measured by the increase of CF fluorescence after its release from the inner volume of vesicles. The CF-loaded vesicles (100 μ M in lipids as final concentration) were treated with the appropriate amounts of oligomers (from 7.5 to 75 μ M for the oligourea and from 25 to 150 μ M for the gamma peptide) in a semi-micro quartz cuvette (F4QS from Hellma, Müllheim, Germany) under constant stirring, with a final volume of 1200 μ l (1000 μ l of vesicles + 200 μ l of oligomers). The increase of fluorescence intensity was monitored during 20 minutes on the Fluoromax-3 spectrofluorimeter, with excitation at 480 nm (slit 1 nm) and emission at 518 nm (slit 1 nm). The initial fluorescence intensity F_0 was measured before addition of the 200 μ l oligomer solution. F_∞ , corresponding to 100 % release was determined by adding Triton X-100 to a final concentration of 0.5 % (w/w), causing a complete lysis of vesicles. The fluorescence intensities $F(t)$ and F_∞ were corrected for dilution effects. Leakage kinetics was quantified on a percentage basis according to the equation:

$$E(t) = \frac{F(t) - F_0}{F_\infty - F_0} * 100 \quad (14)$$

$F(t)$ being the value of fluorescence intensity at time t after addition of the oligomer.

⁴E. Ralston, L.M. Hjelmeland, R.D. Klausner, J.N. Weinstein, R. Blumenthal, *Biochim. Biophys. Acta* **1981**, 649, 133-137.

RÉSUMÉ :

Les propriétés chaperonnes de la NCp7 permettent à cette protéine de réarranger les acides nucléiques de manière à favoriser leurs conformations les plus stables. Ces propriétés sont indispensables à la réplication virale du VIH-1. Nous avons utilisé et développé différentes approches expérimentales basées sur des techniques de fluorescence pour approfondir la compréhension de ces propriétés à l'échelle moléculaire. En utilisant la fluorescence résolue en temps de la 2-Aminopurine, nous avons ainsi caractérisé la liaison de la NCp7 de manière site-spécifique. Nous avons également montré que la restriction de la mobilité locale des bases oligonucléotidiques en réponse à la liaison de la NCp7 constitue une composante mécanistique essentielle de l'activité chaperonne. Nos études ont permis de mieux comprendre le rôle de la NCp7 dans les mécanismes réactionnels impliqués lors des étapes du premier et du second saut de brin au cours de la réverse transcription. Enfin, nous avons développé une approche par spectroscopie de fluorescence à l'échelle de la molécule unique afin d'étudier la dynamique de la liaison de la NCp7 sur les acides nucléiques.

Mots clés : NCp7, chaperonne, fluorescence, TCSPC, FCS, Spectroscopie en molécule unique, VIH-1.

ABSTRACT :

The NCp7 chaperone properties constitute a set of features allowing the NCp7-directed folding of nucleic acids into their most stable conformations. These properties are critical for the viral replication of HIV-1. We used and developed different experimental approaches based on fluorescence techniques in order to further characterize these properties at the molecular level. Using time-resolved fluorescence of 2-aminopurine, we characterized site-specifically the binding of NCp7. We demonstrated that the restriction of the local base mobility in response to the NCp7 binding is a key mechanistic component of the NCp7 chaperone activity. Our studies allowed us to further understand the role of the first and second strand transfer involved in the reverse transcription. Finally, we developed a single-molecule fluorescence spectroscopy setup to study the binding kinetics of NCp7 onto oligonucleotides.

Keywords : NCp7, chaperone, fluorescence, TCSPC, FCS, single molecule spectroscopy, HIV-1.