UNIVERSITÉ DE STRASBOURG

Université de Strasbourg

Ecole Doctorale des Sciences de la Vie et de la Santé

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université de Strasbourg

Discipline : Sciences du Vivant

Domaine : Aspects Moléculaires et Cellulaires de la Biologie

par

Viktoriia POSTUPALENKO

Fluorophores Ratiométriques pour le Marquage de Peptides et d'Oligonucléotides: Applications à la Protéine de la Nucléocapside de VIH-1

Soutenue le 26 septembre 2011 devant la commission d'examen :

Dr. Agnès DELMAS	**Rapporteur externe**
Pr. Solange LAVIELLE	**Rapporteur externe**
Pr. Burkhard BECHINGER	**Rapporteur interne**
Pr. Yves MELY	**Directeur de thèse**
Pr. Vasyl PIVOVARENKO	**Co-Directeur de thèse**

UMR CNRS 7213, Faculté de Pharmacie, ILLKIRCH

This doctoral thesis, involving various fields, would not be possible without the help of many people. I would like to take this opportunity to acknowledge the persons who provided great help in my study.

First and foremost, I would like to express my heartfelt gratitude to my supervisors Prof. Yves Mély and Prof. Vasyl Pivovarenko for their invaluable guidance, support and thoughtful advices during these years of my PhD.

I am also thankful to Dr. Agnès Delmas, Prof. Solange Lavielle and Prof. Burkhard Bechinger who kindly agreed to be jury members of my thesis.

I would like to thank group members that I have worked with over the years notably: Dr. Andrey Klymchenko for his great advices and helpful discussions, Dr. Hugues de Rocquigny for his help with peptide synthesis, Dr. Youri Arntz for his help with AFM studies, Dr. Ludovic Richert and Dr. Pascal Didier for fluorescence imaging studies and Dr. Guy Duportail for his help and support during my research.

I am also thankful to the group of Prof. Alain Burger (Université de Nice Sophia Antipolis) for their collaboration during my research.

I would like to thank all my friends in the lab Sasha, Namrata, Vova, Kamal, Beata, Zeinab, Julien, Vanille, Armelle and other members of the lab for their friendships and support. I would also like to thank master's students Noemie, Mariane and Ziad with whom I worked. I will miss all of you! Good luck to each of you in your future endeavours.

And thank to all the people around me who encouraged and supported me during my PhD.

Ms Viktoriia POSTUPALENKO was a member of the European Doctoral College of the University of Strasbourg during the preparation of her PhD, from 2008 to 2011, class name Rosa Parks. She has benefited from specific financial supports offered by the College and, along with her mainstream research, has followed a special course on topics of general European interests presented by international experts. This PhD research project has been led with the collaboration of two universities: Taras Shevchenko National University of Kyiv, Ukraine and the University of Strasbourg, France.

ABBREVIATIONS

1,8-ANS	1-Anilinonaphthalene-8-Sulfonic Acid
3HC	3-Hyrdoxychromone
3HF	3-Hydroxyflavone
AFM	Atomic Force Microscope
bp	Base pair
CT-DNA	Calf Thymus DNA
DIEA	N,N-Diisopropylethylamine
DLS	Dynamic Light Scattering
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DOPC	Dioleoylphosphatidylcholine
DOPE	Dioleoylphosphatidylethanolamine
DOPS	Dioleoylphosphatidylserine
ESIPT	Excited-State Intramolecular Proton Transfer
GUVs	Giant Unilamellar Vesicles
HCl	Hydrochloric acid
HIV	Human Immunodeficiency Virus
HOBt	1-hydroxybenzotriazole
LUVs	Large Unilamellar Vesicles
LC-MS	Liquid chromatography-mass spectrometry
MS	Mass Spectrometry
NMP	N-methylpyrrolidone
NMR	Nuclear Magnetic Resonance
N*/T*	Intensity ratio of the N* and T* bands
QY	Quantum Yield of fluorescence
RT	Room Temperature
SPPS	Solid Phase Peptide Synthesis
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Tm	Melting (transition) temperature
λ_{abs}	Position of absorbtion maximum
λ_{em}	Position of fluorescence maximum (i.e. emission maximum)

CHAPTER 1

BIBLIOGRAPHICAL REVIEW

1. BIBLIOGRAPHICAL REVIEW

1.1. Peptide-nucleic acid interactions

Proteins are responsible for a large variety of functions, such as enzymatic catalysis, metabolic reactions, protection against infection, homeostasis and transport. All these functions rely on interactions of proteins with other biomolecules. The characterization of such biomolecular interactions is thus of key importance to understand these functions. Moreover, the experimental determination and prediction of protein interactions with biomolecules like other proteins, small molecules, sugars, DNAs/RNAs or membranes is a key requirement for the design of new drugs, e.g. for the prediction of possible side-effects of drugs interacting with more than one protein, in immunobiology - for the understanding of cross-reactions of antibodies used in medical diagnosis and medical treatment, or in cytology - for understanding the cell-cell communication and cell differentiation, etc (Meyer and Schomburg 1999).

Nucleic acids are universal in living organisms, being found notably in all cells and viruses. In nature, they exist in two similar chemical forms: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Within the cell, DNA commonly adopts a double helix structure, stabilized by the hydrogen bonding between the nucleobases and by π-π stacking interactions that occur between aromatic planes of the nucleobase units (Hecht 1996). The specific formation of hydrogen bonds between the two antiparallel strands of DNA is known as Watson-Crick base pairs (Figure 1.1) (Watson and Crick 1953). Purine nucleobases form hydrogen bonds to pyrimidines, with A bonding to T, and G bonding to C. Moreover, the overall structure of DNA comprises two distinct helical grooves: the minor and the major grooves with specific structural parameters (Figure 1.1). DNA can adopt different structures including single stranded, super-coiled, hairpin DNA or multiple stranded triplexes and quadruplexes structures (Bloomfield, Crothers et al. 2000).

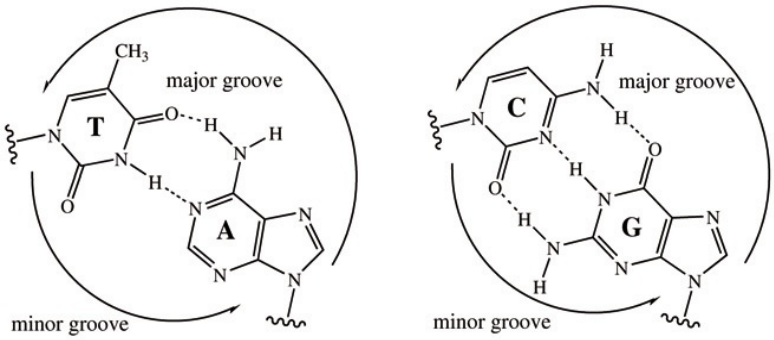


Figure 1.1. Watson-Crick base pairing for T・A (left) and C・G (right) and grooves (major and minor) of nucleic acids.

Protein-nucleic acid interactions play important roles in cellular functions. Gene expression, transcription, replication, recombination, packaging and repairing are notably controlled by these interactions. The main function of DNA is the long-term storage and carrying of genetic information (Avery, Macleod et al. 1944; Hershey and Chase 1952), whereas RNA has a variety of

cellular functions in addition (Blackburn, Gait et al. 2006). For example, messenger RNAs (mRNAs) in the transcription process bring information on protein sequences from DNA to the translation machinery, composed of mRNA in complex with transfer RNA (tRNA) and ribosome. In this process, each tRNA serves as the adaptor (codon-anticodon interaction) between the mRNA and a given amino acid, serving as a building block of the future protein (Bloomfield, Crothers et al. 2000; Blackburn, Gait et al. 2006). All the tRNAs bind appropriate amino acids, and then transfer them to the translation machinery (Nomanbhoy, Morales et al. 2001). At all stages (from transcription to translation), an RNA transcript interacts with many different RNA-binding proteins. The composition of this supramolecular assembly, known as a ribonucleoprotein particle, is diverse and highly dynamic. The association of RNA-binding proteins (RBPs) with RNA transcripts begins during transcription. Some of these early-binding RBPs remain bound to the RNA until it is degraded, whereas others recognize and transiently bind to RNA at later stages for specific processes such as splicing, processing, transport and localization (Dreyfuss, Kim et al. 2002). RNA binding proteins coordinate translation, mRNA localization and stability, and pre-mRNA splicing through the association with defined sequences in target transcripts. Some RBPs function as RNA chaperones (Lorsch 2002) by helping the RNA, which is initially single-stranded, to form various secondary or tertiary structures (for example, nucleocapsid proteins (Rein, Henderson et al. 1998)). When folded, these structured RNAs, together with specific RNA sequences, act as a signal for other RBPs that mediate gene regulation. In viruses and retroviruses, RNA can carry genetic information, which is then copied into DNA by reverse transcriptase (Hecht 1996). Apart from transcription, translation, regulation and transport functions, RNA molecules are involved in catalytic processes (Kruger, Grabowski et al. 1982; Cech 1987; Cech 1989; Qin and Pyle 1998; Cech 2000; Takagi, Warashina et al. 2001). RNAs with enzymatic (specifically, catalytic) activity, such as the self-splicing molecules, are commonly referred to ribozymes (Serganov and Patel 2007). The diversity of functions of RNA-binding proteins would suggest a correspondingly large diversity in the structures that are responsible for RNA recognition. However, most RNA-binding proteins are built from few RNA-binding modules. Instead, the large structural diversity of substrates is accommodated by the presence of multiple copies of these RNA-binding domains (RBDs) presented in various structural arrangements to expand the functional repertoire of these proteins. There are different domains such as the RNA-recognition motif (RRM), the K-homology (KH) domain (which can bind both single-stranded RNA and DNA), the dsRBD (a sequence-independent dsRNA-binding module) and RNA binding zinc-finger (ZnF) domains and others (Lunde, Moore et al. 2007). Some examples of how RNA-binding modules recognize RNA are represented on Figure 1.2.

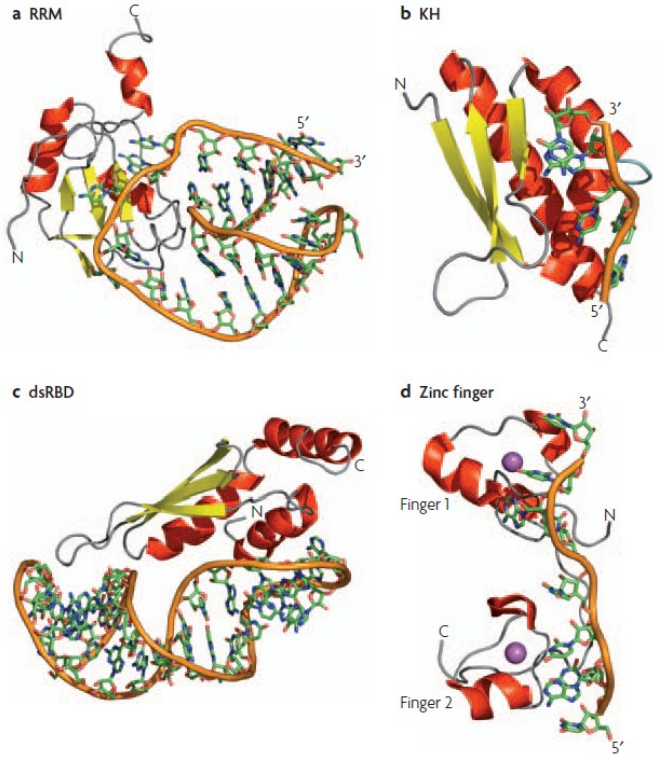


Figure 1.2. Recognition of RNA by RNA-binding modules (Lunde, Moore et al. 2007). The RNA backbone is represented with an orange ribbon, α-helices are in red and β-sheets are in yellow; the zinc atom in the TIS11d structure is in magenta. a) Structure of the N-terminal RNA-recognition motif (RRM) of human U1A bound to RNA (Oubridge, Ito et al. 1994). In this structure, and in many other RRM–RNA complexes, single-stranded bases are specifically recognized through the protein β-sheet and through two loops that connect the secondary structure elements. b) K-homology-3 (KH3) domain of Nova-2 bound to 5′-AUCAC-3′ (Lewis, Musunuru et al. 2000). KH domains bind to both single-stranded DNA and RNA through a conserved GXXG sequence located in an exposed loop (light blue). c) Yeast Rnt1 double-stranded RNA-binding domain (dsRBD) bound to an RNA helix capped by an AGNN tetraloop (Wu, Henras et al. 2004). A conserved protein loop (left-most part of the structure) interacts with 2′-OH groups in the RNA minor groove, whereas highly conserved Lys and Arg residues at the end of the longer helix recognize the position of phosphate atoms that are characteristic of an A-form helix. d) Zinc fingers of TIS11d bound to an AU-rich RNA element (Hudson, Martinez-Yamout et al. 2004). The identity of the single-stranded RNA is recognized by the protein backbone through hydrogen bonds with the Watson-Crick face of each base.

Myriad of cellular processes require extensive interaction of nucleic acids with various classes of proteins: for example, polymerases, helicases, ligases, transcription factors, histones, and ribosomal proteins. Some proteins recognize and bind only single-stranded nucleic acids, whereas other proteins recognize and bind only double-stranded nucleic acids with a defined sequence motif or they bind nucleic acids with a particular three dimensional structure. DNA-binding proteins are responsible for replicating the genome, for transcribing active genes, and for repairing damaged DNA. One of the largest and most diverse class of DNA-binding proteins is the transcription factor family (Pabo and Sauer 1992) that regulates cell development, differentiation, and cell growth by binding to specific DNA sites (or set of sites) and regulates gene expression. Certain human

transcription factors can recognize and bind to short DNA sequences that are found in only a few copies in a genome that consists of billions of nucleic acid bases. Well-established families of transcription factors include the helix-turn-helix (HTH) proteins (the Escherichia coli CAP protein (McKay and Steitz 1981), Lac repressor (Kaptein, Zuiderweg et al. 1985), Trp repressor (Schevitz, Otwinowski et al. 1985)), the homeodomains (Drosophila Antennapedia (Antp) homeodomain (Billeter, Qian et al. 1990)), zinc finger proteins (Xenopus transcription factor IIIA (TFIIIA) (Brown, Sander et al. 1985)), the steroid receptors, leucine zipper proteins, and the helix-loop-helix proteins (Rohs, Jin et al. 2010).

Investigation of the interaction between DNA and its binding partners (from small molecules to large proteins) are of interest for the understanding processes such as unspecific or specific DNA binding as well as transcription. Increasing amounts of data on the structures of protein-DNA complexes provided us detailed information about the features of interactions between DNA bases and amino acids. However, inspection of these features reveals that there is no clear one-to-one correspondence between amino acids and bases and that their interactions show a large diversity by their spatial structure, i.e., the same pair may interact using a variety of geometries. This is in marked contrast with the rather strict complementarity of base pairing in double-stranded DNA. It is critically important to understand the basic mechanisms by which DNA- and RNA-binding proteins recognize and bind appropriate target sequences with high selectivity and specificity.

Structural studies of protein-DNA complexes have suggested two underlying mechanisms by which proteins discriminate between potential target sites. One is “direct readout,” in which specific amino acid residues of the protein make contacts (e.g., hydrogen bonds) with specific bases (A, T, G, or C). Proteins that use direct readout typically bind to the major groove of DNA where the pattern of hydrogen-bond donors and acceptors is distinct for each Watson-Crick base pair (Rice 2008). Thus, members of this class of proteins recognize their binding sites as a result of direct examination of the hydrogen-bonding pattern exposed in the major groove, and which is unique for each kind of base pair. Examples of protein motifs with DNA-binding domains that primarily employ direct readout are the helix-turn-helix motif in, for example, the 434 repressor (Aggarwal, Rodgers et al. 1988), the basic leucine zipper (bZIP) motif in yeast GCN4 protein (Ellenberger, Brandl et al. 1992), and the zinc finger motifs in TFIIIA, a transcriptional regulator from Xenopus (Foster, Wuttke et al. 1997).

Alternatively, a significant fraction of site-specific DNA-binding proteins recognizes the minor groove of DNA where the pattern of hydrogen bond donors and acceptors is similar for all Watson-Crick base pairs, and symmetric, making it difficult for the protein to distinguish A from T, and G from C (Zhang, Xi et al. 2004). Furthermore, the G:C pair is distinguished from the A:T pair only by one additional centrally located donor group. Thus, for the minor groove-binding proteins, there is a lack of direct sequence information. Such proteins recognize their target sites with high fidelity via “indirect readout.” In this mechanism, sequence-dependent variations in the mechanical properties of DNA, such as local DNA flexibility/“bendability” (Rohs, West et al. 2009) and changes in major or minor groove width, as well as differences in the stacking or twisting parameters, play important roles in the recognition process. Thus, minor groove-binding proteins are believed to “read” DNA sequence by the ease with which they can bend, twist, or deform the potential binding sites in a way to optimize the protein-DNA interface, and in some cases

undergoing large conformational changes as in the opening of the minor groove in the complex formed between TBP and the TATA box (Kim, Nikolov et al. 1993). Quite often, proteins use a combination of both direct and indirect readout to recognize their binding sites.

There are numerous reports in which a mutation of a base not in contact with any amino acid was found to affect the binding affinity (Chen, Gunasekera et al. 2001; Horton, Dorner et al. 2002; Lamoureux, Stuart et al. 2002; Lawson, Swigon et al. 2004). Examples include the 434 repressor (Koudelka and Carlson 1992), catabolite activator protein (CAP) (Chen, Gunasekera et al. 2001), and HincII restriction endonuclease (Horton, Dorner et al. 2002). In this respect, proteins may recognize a DNA sequence via i) water-mediated contacts, ii) specific sequence-dependent conformational features (e.g., bending and local geometry of base pair steps), and/or iii) binding-induced distortion of the DNA (DNA flexibility). It is believed that water molecules play a significant role in protein-DNA interactions. Most water molecules around DNA-protein interfacial areas are released to the bulk solvent upon formation of a protein-DNA complex, which may play a role in the binding thermodynamics, but it is largely unknown whether the process also contributes to binding specificity. A comprehensive analysis of interfacial water molecules within the structures of 109 unique protein-DNA complexes shows that less than 2% of crystallographic water molecules are used to bridge side chains and bases, whereas 76% are used to stabilize either the protein or the DNA separately. The unexpected low frequency of water-mediated interactions, and the high frequency of hydrogen bond formation with either the protein or DNA, suggests that the main role of water is to solvate the protein and DNA atoms at the interface and not to mediate protein-DNA contacts (Reddy, Das et al. 2001). Within the structures of protein-DNA complexes, the DNA molecule is often deformed to some degree, compared with the structure of the free DNA. In this way, the conformation of DNA can contribute to the specificity of protein-DNA recognition by assuming a sequence-specific intrinsic conformation to which a protein preferentially binds. For example, the importance of structural flexibility was recently shown for nonspecific and specific DNA binding of lac repressor (Kalodimos, Biris et al. 2004). The number of available structures of protein-DNA complexes has grown substantially, and several researchers have analyzed the interactions between amino acids and bases in order to find out the general principles of protein-DNA recognition (Suzuki 1994; Pabo and Nekludova 2000; Luscombe, Laskowski et al. 2001). A consensus derived from these studies is that there are no one-to-one correspondences between amino acids and bases (Sarai and Kono 2005). Molecular recognition of DNA by proteins and peptides is usually governed by a combination of electrostatic interactions (salt bridges), dipolar interactions (hydrogen bonding, H-bonds), entropic effects (hydrophobic interactions) and dispersion forces (base stacking) (Rice 2008). Hydrophobic bonding, defined as the thermodynamic tendency to bury nonpolar residues in the interior of a protein (away from the aqueous medium), was shown to be central for stable protein folding, and the related decrease of the surface exposure of nucleic acid bases by base stacking was established as the dominant interaction for stabilizing duplex DNA. However, many proteins bind to single and double-stranded DNA electrostatically, with little or no dependence on base or base pair composition or sequence (von Hippel 2007). These different forces contribute in varying degrees to proteins binding in a sequence-specific (tight) or non-sequence specific (loose) manner. For example, specific protein-DNA interactions are commonly mediated by an α-helix motif in the protein which inserts itself into the major groove of

the DNA, recognizing and interacting with a specific base sequence through H-bonds and salt bridges. In addition, the affinity and specificity of a particular protein-nucleic acid interaction can be increased through protein oligomerization or multi-protein complex formation (transcription initiation complexes, mRNA splicing complexes, etc.).

Proteins that interact with DNA can be divided into two groups: those that actively move along DNA ("active sliders") and those that do not. The first group contains proteins such as DNA and RNA polymerases and helicases. Many proteins in the second group slide or hop along DNA, either through attachment to active sliders or by diffusion ("passive sliders") (Williams and Maher 2011). Otto Berg, Peter von Hippel, and colleagues developed a theory (Berg, Winter et al. 1981) describing four possible "passive sliding" modes for DNA target searching (Figure 1.3) – more generally, modes of moving from one site to another on DNA:

1) sliding along DNA via continuous one-dimensional (1D) diffusion without dissociation;

2) hopping, where the protein effectively diffuses along a single molecule of DNA but does so via a series of dissociation and rebinding events;

3) jumping, which, in contrast to facilitated diffusion, amounts to ordinary 3D diffusion between DNA sites;

4) intersegmental transfer, where the protein swaps sites on the DNA via a looped intermediate.

To date with some minor differences all models of proteins moving on DNA are in line with these modes (Halford and Marko 2004). The first, second, and third modes present a continuum transition from 1D diffusion on DNA to 3D diffusion. The fourth mode, intersegmental transfer, requires at least two DNA-binding domains in a special arrangement.

Among the members of the second group are proteins such as DNA replication processivity factors (PCNA, b clamp), which are proteins responsible for detecting DNA damage (MutM (Blainey, van Oijent et al. 2006)), proteins involved in cutting DNA in specific places (EcoRV (Bonnet, Biebricher et al. 2008), BamHI (Blainey, Luo et al. 2009)), promoter and repressor DNA transcription factors (Tafvizi, Huang et al. 2008), adenoviral AVP–pVIc complex (Blainey, Luo et al. 2009), RNA polymerase (RNAP) (Harada, Funatsu et al. 1999), and DNA sliding clamps (PCNA (Kochaniak, Habuchi et al. 2009)) that aid critically in DNA replication and are also involved in DNA repair.

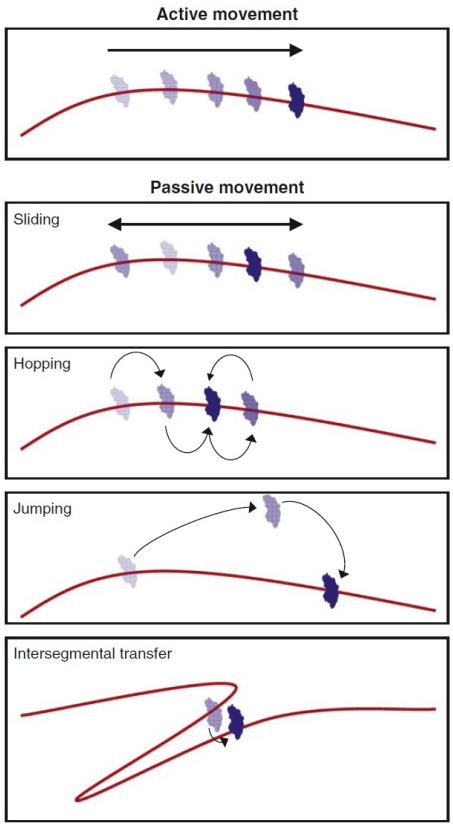


Figure 1.3. Modes of proteins searching for DNA targets (Williams and Maher 2011).

Dynamics of proteins on the DNA involve transient interactions that are difficult to observe and even more difficult to quantify. Crystal structures sometimes provide hints about these events, e.g. by the lack of electron density of DNA or protein regions, or by the observation of different conformations of amino acid side chains associated with the nucleic acids. Tomographic electron microscopy can provide structures of protein-nucleic acid complexes inside the cell (Brandt, Carlson et al. 2010), and NMR has the potential to do so (Sakakibara, Sasaki et al. 2009; Dominguez, Schubert et al. 2011). Analysis in solution by NMR is a powerful approach (Clore and Iwahara 2009), allowing the study of the kinetics of translocation (Doucleff and Clore 2008; Takayama, Sahu et al. 2010), as well as the structures of transient, nonspecific complexes. For instance, the structure of the Lac repressor bound to a nonspecific (low-affinity) DNA sequence suggests that binding is primarily driven by electrostatics, as most of the protein-DNA interactions do not involve the bases, but the phosphates and sugars of the DNA backbone (Kalodimos, Biris et al. 2004). NMR is uniquely suited to characterize folding-unfolding events that occur at disordered regions of proteins that become structured upon recognition of their target nucleic acids, and can be usefully complemented by small-angle X-ray scattering (Jimenez-Menendez, Fernandez-Millan et al. 2010; Mertens and Svergun 2010). Mass spectrometry is emerging as a potent tool for the study of dynamic or heterogeneous protein-nucleic acid complexes (Gordiyenko and Robinson 2008). Surface plasmon resonance and microcalorimetry are used to characterize structure-activity relationships.

Specific questions are also frequently addressed using single-molecule techniques, (Efcavitch and Thompson 2010) such as for instance RNA folding by a helicase (Karunatilaka, Solem et al. 2010), DNA transport (Ptacin, Nollmann et al. 2008), and DNA polymerization (Olasagasti, Lieberman et al. 2010). Moreover, using atomic force microscopy (AFM) (Binnig, Quate et al. 1986) and optical tweezers (OT) (Gordon 1973; Ashkin 1997), the interaction force between molecules becomes directly measurable and can be used to describe the interaction between different molecules at the single-molecule level (Sewald, Wilking et al. 2006). Direct force measurements on sequence-specific protein-DNA complexes have been performed at the single-molecule level with OT in the case of the restriction endonucleases BsoBI, XhoI, and EcoRI (Koch, Shundrovsky et al. 2002) and with AFM, in the case of transcriptional regulator proteins (Bartels, Baumgarth et al. 2003). In contrast to ensemble measurements, experiments handling single molecules are capable to detect fast intermediate transition states, provide details of the energy landscape, and describe structural changes, and, thus, 'individual' behaviors.

Special attention was paid to develop methods for measuring the parameters of binding, such as the binding affinity and stoichiometry of protein-nucleic acid complexes. One of the most commonly used methods is the electrophoretic mobility shift assay (EMSA), which allows visualizing the interaction between a protein and a radioactively labeled DNA or RNA molecule (Garner and Revzin 1981; Hellman and Fried 2007). This technique is based on the lower electrophoretic mobility of a protein-nucleic acid complex in comparison to free nucleic acid. The method is highly sensitive, allowing the assays to be performed with low protein and nucleic acid concentrations. When such high sensitivity is not needed, variants or the assay using fluorescence and chemiluminescence detection are also available (Berger, Duncan et al. 1993; Jing, Agnew et al. 2003). Since the EMSA method suffers from limitations, such as the fact that it is not a true equilibrium method, alternative methods have been developed such as affinity capillary electrophoresis (Xian, Harrington et al. 1996; Foulds and Etzkorn 1998; Guszczynski and Copeland 1998; Li and Martin 1998; Tajmir-Riahi 2006).

Many techniques are available for the detection and characterization of protein-nucleic acid complexes, showing advantages and disadvantages that differ from those of the EMSA. The most widely used alternative assays are nitrocellulose filter-binding (Hall and Kranz 1999) and footprinting (Galas and Schmitz 1978; Brenowitz, Senear et al. 1986). Filter-binding is simple to perform and the manipulations are rapid enough to allow kinetic studies as well as equilibrium measurements (Woodbury and von Hippel 1983). The process is based on the fact that proteins bind to nitrocellulose without losing their DNA binding capacity, while double-stranded DNA alone is not retained. Like the EMSA, filter-binding is a nonequilibrium technique. As a result, quantitative analyses require careful evaluation of filter-retention efficiency (Woodbury and von Hippel 1983; Oehler, Alex et al. 1999).

The DNase footprinting technique applied to protein-DNA complexes is a protection assay based on the ability of a DNA binding protein to protect DNA from cleavage by different deoxyribonucleases (DNase) or chemical cleavage agents. The method uses an enzyme to cut radioactively end-labeled DNA, followed by gel electrophoresis to detect the resulting cleavage pattern (Galas and Schmitz 1978; Helwa and Hoheisel 2010). This technique generally allows the identification of the DNA binding site of a protein. The use of different deoxyribonucleases that

specifically cleave nucleotides located in single-stranded or double-stranded regions also allow the identification of the secondary structures of the DNA bound by the protein. In addition, the footprint signal can be obtained under conditions of binding equilibrium for the protein of interest. This is an important advantage over nonequilibrium assays such as EMSA. Variants of the footprinting assay, optimized for quantitative detection of binding have been described (Brenowitz, Senear et al. 1986), and time-resolved methods have been developed that allow the analysis of binding kinetics as well as equilibria (Shcherbakova, Mitra et al. 2006).

Optical techniques such as circular dichroism (CD), Fourier transform infrared (FTIR), Raman, and UV-vis absorption spectroscopies are also widely used to study protein-nucleic acid interactions. CD and FTIR are of special interest since they allow monitoring changes in the secondary structures of peptides and proteins (Cary and Kneale 2008). Though these techniques are easy to use, they are generally limited in their sensitivity, so that only systems of moderate affinity can be analyzed. Moreover, they strongly depend on the amplitude of the signal changes occurring on binding, which is hard to predict and frequently low. The most used optical technique is fluorescence spectroscopy, which suffers less from these limitations and has thus been extensively used to study protein-nucleic acid interactions. This technique will be highlighted below.

During the last 40 years, technology for the detection of protein-nucleic acid interactions has enormously progressed. From initial qualitative in vitro assays on individual interaction partners, single molecule sensitivity, and in vivo quantification have been achieved. Currently, about 72 000 structures are deposited in the Protein Data Bank (http://www.wwpdb.org), but only about 4% are protein-nucleic acid complexes. It is difficult to know whether these figures mirror the prevalence of these complexes in the cell, or whether they arise from the difficulties in the identification and characterization of protein-nucleic acid complexes. It is consequently crucial to develop new technologies or improve existing ones for the analysis of protein-nucleic acid interaction in order to identify and characterize the binding partners.

1.2. Peptide-membrane interactions

The cell membrane is the basic structural part of the cell that encapsulates its contents and defines the intra- and extra- cellular space. It provides the integrity of the cell, preventing contents of the cell from leaking out. It regulates also the transport of molecules across the cell (ions, nutrients etc.) and maintains the cell potential. Furthermore, the cell membrane serves as a protective barrier, which prevents transport of undesired molecules and pathogens into the cell. Cell membranes are highly complex both in terms of composition and function, which vary from cell to cell (van Meer, Voelker et al. 2008). The classical representation of biological membranes, which was introduced by Singer and Nicolson in 1972, is the fluid mosaic model (Figure 1.4) (Singer and Nicolson 1972). In this description, a membrane is composed mainly of lipids and proteins that form a thin (about 5 nm width) bilayer film with membrane proteins either embedded in this structure or located at the surface of the membrane. Other components of the cell membrane may include cholesterol, sugars and other organic species. The membrane structure is highly flexible and allows lateral diffusion of both proteins and lipids. Due to the high complexity of natural biomembranes, simplified model membranes such as lipid vesicles and lipid bilayers are used to characterize the properties of lipids and their interaction with a broad variety of peptides.

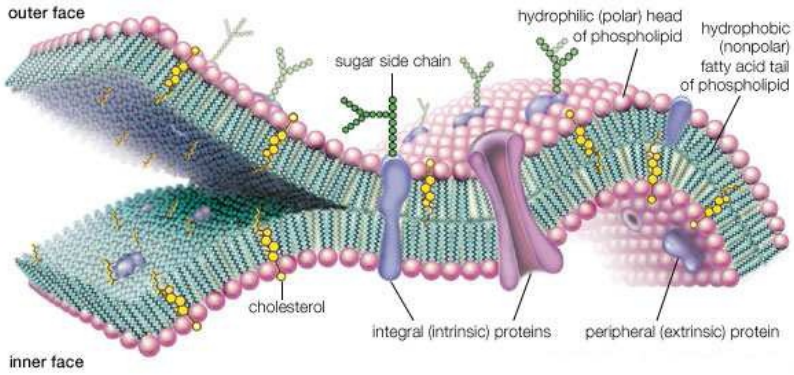


Figure 1.4. The fluid mosaic model. The membrane is composed of a bilayer structure, integral and peripheral proteins and several other organic molecules. The membrane proteins and the lipids are free to diffuse laterally in the bilayer. The figure was adapted from Encyclopedia Britannica web page.

Peptide-membrane interactions are an active research topic due to their wide implications in therapeutics, diagnostics, membrane permeation and the basic need for understanding the structure and function of membrane proteins. Peptide-membrane interactions play an important role in a number of biological processes, such as antimicrobial defense mechanisms, viral translocation, membrane fusion and functions of membrane proteins. For example, antimicrobial peptides are a family of peptides with a particular propensity to recognize and disintegrate bacterial pathogens. A number of these peptides have been identified as key components of the natural immune system (Yeaman and Yount 2003). A related family of peptides is the so-called cell-penetrating peptides (CPPs) capable of efficient translocation through the cell membranes, either by themselves or together with a molecular cargo (Langel 2007). These peptides are being explored as potential

programmable drug delivery vectors. As a part of larger proteins, ion-conducting channel peptides form well-organized transmembrane bundles capable of selective transport of ions. Other peptides are believed to play a key role in various complex cellular processes, such as membrane fusion. It is clear that a better understanding of peptide-membrane interactions on a molecular level is not only important in the elucidation of various biological processes, but also could be instrumental in designing peptides with tailored functionalities, for example, for antibiotic and drug delivery applications.

Peptides are broadly classified on the basis of their function, secondary structure, charge and/or predominance of certain amino acids. During the past three decades, a vast number of antimicrobial peptides (AMPs) (Hultmark, Steiner et al. 1980; Zasloff 1987) and other related cytolytic peptides (Kreger, Kim et al. 1971) have been discovered and extensively studied in order to understand the general principles of peptide-lipid interactions, as well as the relation of these interactions to the biological function of these peptides. More recently, several cell-penetrating peptides have been described and explored as potential programmable drug delivery vectors for transport of large macromolecules, such as proteins, peptides (Frankel and Pabo 1988; Soomets, Lindgren et al. 2000; Hallbrink, Floren et al. 2001), oligonucleotides (Morris, Vidal et al. 1997), polysaccharides (Henriques, Costa et al. 2005) and other large particles, like liposomes (Torchilin, Rammohan et al. 2001) and nanoparticles (Lewin, Carlesso et al. 2000), across cellular membranes both *in vitro* and *in vivo*. Surprisingly, many of these antimicrobial, cytolytic, and cell penetrating peptides fall into the same structural class (Henriques, Melo et al. 2006). These peptides vary considerably in chain length, hydrophobicity and overall distribution of charge, but they form an amphipathic α-helix, when bound to a membrane surface (Bechinger 1997; Shai 1999; Kourie and Shorthouse 2000). What has befuddled researchers for a long time is the absence of correlation between their sequence and their function or mechanism. The only element that appears to distinguish antimicrobial from cytolytic peptides is the fact that antimicrobials are usually cationic. This provides a simple explanation for their specificity because cationic peptides should bind better to the anionic membranes of most bacteria than to the neutral membranes of eukaryotic cells (Zasloff 2002).

Numerous parameters are of importance for the binding of peptides at the membrane, including peptide length, charge (distribution), hydrophobicity (distribution), secondary structure, and topology (Popot and Engelman 2000; von Heijne 2006; Stromstedt, Ringstad et al. 2010). These parameters may also tune other aspects of peptide-membrane interactions important for membrane lysis, e.g., the degree of peptide penetration into the membrane interior. Peptides that associate with membranes interact with the lipid bilayer by different mechanisms. Most often, the peptide has some amphiphilic character and both hydrophobic effects and electrostatic forces allocate the peptide in the hydrophilic-hydrophobic interfacial layer of the membrane. Upon association with the membrane, the peptides often change their conformation (White and Wimley 1998). In other cases, even for strongly hydrophilic and water-soluble peptides, the binding to the bilayer can be facilitated by hydrophobic anchors, such as acylation, prenylation or myristoylation. The various modes of the interaction of peptides with lipid bilayers are illustrated in Figure 1.5.

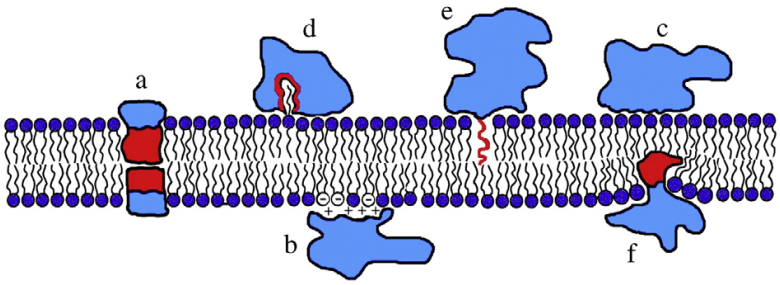


Figure 1.5. Schematic illustration of the different strategies used by peptides to bind to, anchor to, imbed in, and penetrate lipid membranes. (a) trans-membrane-spanning amphiphilic peptide dimer; (b) electrostatic binding; (c) non-specific binding by weak physical forces; (d) anchoring via a lipid extended conformation; (e) anchoring by an acyl-chain anchor attached to the protein; (f) amphiphilic peptide partially penetrating the bilayer (Khandelia, Ipsen et al. 2008).

Mechanical properties of the lipid bilayers influence the ability of peptides to associate with the membranes (Lee, Chen et al. 2004; Allende, Simon et al. 2005). For example, inclusion of cholesterol in the lipid bilayers tends to decrease the ability of peptides to associate with the membranes (Benachir, Monette et al. 1997; Wieprecht, Beyermann et al. 1999). In order to accommodate the bound peptide, the lipids have to rearrange in the proximity of the adsorbed peptide, resulting in an area expansion as well as a membrane thinning (Ludtke, He et al. 1995; Chen, Lee et al. 2003). The mass imbalance due to the adsorption of peptides to the outer leaflet induces a positive curvature strain on the bilayer. The stress is believed to be relieved by the formation of pores in the membranes which allows for a transport of lipids and peptides between the leaflets (Matsuzaki 1999). Bilayer properties such as spontaneous curvature, bending rigidity and area compressibility modulus affect the permeabilizing ability of the peptides (Benachir, Monette et al. 1997; Allende, Simon et al. 2005; Lee, Hung et al. 2005). Likely, different mechanisms apply to different peptides. Moreover, the disrupting mechanism appears to be dependent on the bilayer composition (Ladokhin and White 2001; van den Bogaart, Mika et al. 2007). Theories trying to unify different experimental results havebeen suggested (Matsuzaki 1999; Brogden 2005; Huang 2006). At low peptide/lipid ratios, peptides are bound parallel to the lipid bilayer. As the peptide/lipid ratio increases, peptides begin to orientate perpendicular to the membrane. At high peptide/lipid ratios, peptide molecules are orientated perpendicularly and insert into the bilayer, forming transmembrane pores which may lead to destabilization of the bilayer (Brogden 2005). Using model lipid systems, several mechanisms have been proposed for membrane disruption based on experimental data (Figure 1.6) (Shai 2002; Brogden 2005; Sanderson 2005; Lohner, Sevcsik et al. 2008; Almeida and Pokorny 2009). The classical mechanisms are based on formation of stable or short-lived membrane pores of the barrel-stave (Ehrenstein and Lecar 1977) or toroidal (Ludtke, He et al. 1996; Matsuzaki, Murase et al. 1996), or bacterial breakdown by detergent-like action (carpet type) (Bechinger, Kinder et al. 1999; Bechinger 2005) (for review see (Bechinger 1999)). Barrel-stave pores have only been experimentally demonstrated for a few peptides, such as alamethicin (Qian, Wang et al. 2008). Toroidal pores, on the other hand, can be formed by a greater variety of peptides. Prior to formation of barrel-stave and toroidal pores, the peptide adsorbs parallel to the membrane surface (Huang 2006; Huang 2009; Melo, Ferre et al. 2009). The main differences

between these models lie in the lipid structure around the pores and the pore stability. The holes, when the peptides are in high concentrations, may result into a complete collapse of the membrane.

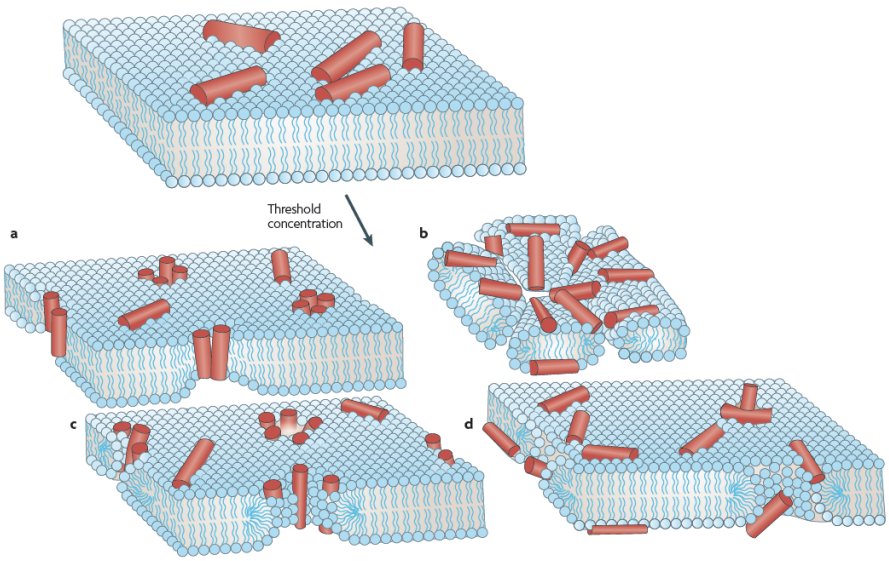


Figure 1.6. Proposed mechanisms of antimicrobial peptide-mediated membrane disruption (Melo, Ferre et al. 2009). a) Barrel-stave pore. Peptides insert perpendicularly in the bilayer, associate and form a pore. b) Carpet mechanism. Peptides adsorb parallel to the bilayer and, after reaching sufficient coverage, produce a detergent-like effect that disintegrates the membrane. c) Toroidal pore. Peptides insert perpendicularly in the bilayer and induce a local membrane curvature. d) Disordered toroidal pore. A recent modification to the toroidal pore proposes that less-rigid peptide conformations and orientations are formed.

The most studied peptides from the CPP family are the Tat peptide (transactivator of HIV-1 virus) (Frankel and Pabo 1988) and penetratin (Drosophila antennapedia transcription protein - pAntp) (Joliot, Pernelle et al. 1991), which were the first peptides identified to cross the membrane. Their ability to translocate the membrane is attributed to their basic amino acid sequences. The minimal peptide sequence necessary for translocation has been elucidated for Tat (Vives, Brodin et al. 1997) and pAntp (Derossi, Joliot et al. 1994). pAntp has attracted widespread interest because of its ability to directly target attached oligopeptides and oligonucleotides to the cytoplasm and nuclear compartments of cells (Prochiantz 1996). Many studies were performed to unravel the mechanism of translocation of penetratin into the membrane (Alves, Goasdoue et al. 2008; Lamaziere, Wolf et al. 2008; Alves, Correia et al. 2009). It was suggested that penetratin translocates through the lipid bilayer by transiently forming inverted micelles (Prochiantz 1996; Derossi, Chassaing et al. 1998). Endocytosis, macropinocytosis (Conner and Schmid 2003) and electroporation-like permeabilization mechanism (Binder and Lindblom 2003) may also participate in the uptake of penetratin.

The first antimicrobial and cell lytic peptides were isolated about 40 years ago. The number of antimicrobial peptides identified in nature is continuously increasing and compiled in the Antimicrobial Sequences Database (http://www.bbcm.univ.trieste.it/~tossi/antimic.html). By now, nearly 1000 peptides have been identified and many more have been created by design using synthetic combinatorial approaches. Antimicrobial peptides, such as alamethicin, isolated from

Trichoderma viride fungus (Leitgeb, Szekeres et al. 2007), the bee venom peptide melittin (Dempsey 1990; Raghuraman and Chattopadhyay 2007), the magainins of the African frog Xenopus laevis (Matsuzaki 1999), and many others, are among the most intensively studied peptides.

Magainin is a 23-amino acid cationic peptide which exhibits antibacterial, antifungal, antiviral and anticancer activity. The biological activity of magainin correlates with its ability to induce leakage of lipid vesicles (Huang 2006), confirming its bilayer-perturbation mechanism of action. Magainin folds into an amphipathic helix at the membrane interface (Bechinger, Zasloff et al. 1992; Marassi and Opella 2000; Porcelli, Buck-Koehntop et al. 2006). Its small size, cationic charge and helical amphipathic structure ensure its quick diffusion to the membrane interface, electrostatic binding to the lipid head groups and subsequent perturbation of membrane structure by hydrophobic association and/or a detergent-like action once a sufficient number of peptides aggregate locally (Khandelia, Ipsen et al. 2008).

Melittin is a strongly basic 26 amino-acid peptide which shows various biological, pharmacological and toxicological activities, including strong surface activity on cell lipid membranes, hemolyzing activity, antibacterial, antifungal and antitumor properties (Chen and Lariviere 2010). In aqueous solution at submillimolar concentration and low ionic strength, melittin exists as a random coil monomer (Figure 1.7A). The crystal structure shows that at higher ionic strength, especially in the presence of divalent anions, melittin forms a water-soluble tetramer with each subunit comprising two α-helical segments connected by a hinge at residues 11 and 12 with a kink of ~120° (Figure 1.7B) (Terwilliger and Eisenberg 1982; Terwilliger, Weissman et al. 1982). Upon interaction with lipid membranes melittin folds into a helix which makes the structure amphipathic with the hydrophobic face inside the bend and the hydrophilic face on the outside (Figure 1.7C). While the C-terminal section is approximately parallel to the lipid bilayer, the N-terminal section is inserted halfway into the membrane, as confirmed by fluorescence quenching, NMR, ATR-IR, and computational studies (Weaver, Kemple et al. 1992; Berneche, Nina et al. 1998). There is significant number of studies examining the orientation of melittin within the membrane but it is still a matter of controversy. Melittin can be oriented either perpendicular or parallel to the membrane surface, depending on pH, temperature, phospholipid composition, and peptide concentration (Raghuraman and Chattopadhyay 2007).

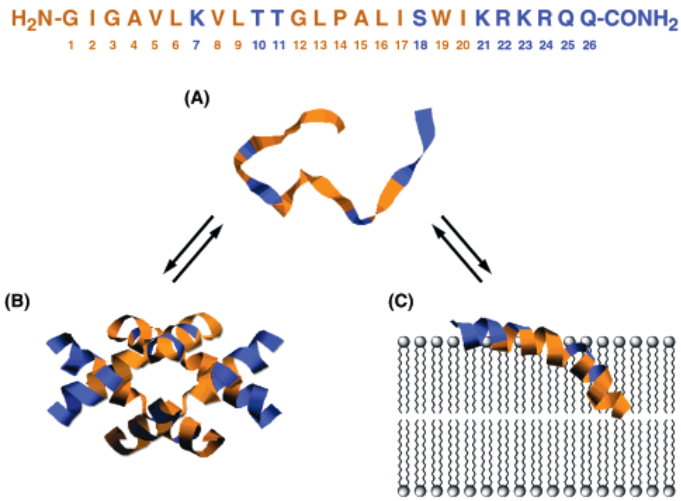


Figure 1.7. Amino acid sequence of melittin with hydrophobic (orange) and polar (blue) residues: (A) random coil conformation in equilibrium with the (B) tetramer and (C) lipid bound monomer (structures of the melittin tetramer and monomer are according to PDB-entry 2MLT (Terwilliger and Eisenberg 1982)) (Niemz and Tirrell 2001).

Many biophysical approaches have been used to investigate peptide-membrane interactions. The attachment, insertion and orientation of the peptide, as well as the orientation of the lipids and the thickness and integrity of the lipid bilayer can be measured by X-ray crystallography, NMR and CD spectroscopy, Fourier transform infrared spectroscopy (FTIR). All of these methods can provide information on the orientation of the peptides with respect to the bilayer normal (Salditt 2003), as well as on the secondary structure of the peptide. The structure, dynamics and conformation of membrane proteins can be investigated by electron paramagnetic resonance (EPR) (Tamm, Lai et al. 2007; Vamvouka, Cieslak et al. 2008) and solid-state NMR spectroscopy (Bechinger 1999; Dave, Billington et al. 2005; Hunter, Jing et al. 2005). NMR spectroscopy has proved particularly useful when information is already available for the structure of a peptide in the absence of membranes, as coupling constants and chemical shift anisotropies may be monitored to assess changes in conformation following binding (Yamaguchi, Huster et al. 2001). Solid-state NMR spectroscopy measures the secondary structure, orientation and penetration of peptides into lipid bilayers (Bechinger 1999). For example, it was used to establish the transmembrane alignment of the pore-forming antimicrobial peptides alamethicin (Bak, Bywater et al. 2001; Bechinger, Skladnev et al. 2001; Salnikov, De Zotti et al. 2009) and melittin (Naito, Nagao et al. 2000).

Circular dichroism and infrared spectroscopy allow monitoring protein conformation, picosecond dynamics and topologies (Wu, Huang et al. 1990; Hristova, Dempsey et al. 2001; Dave, Billington et al. 2005; Hunter, Jing et al. 2005). Oriented circular dichroism was applied to demonstrate that mellitin can be oriented parallel or perpendicular to the membrane to form a transmembrane pore (Yang, Harroun et al. 2001). To quantitatively study the thermodynamics of peptide membrane association, isothermal titration calorimetry has been used (Seelig 2004; Dave, Billington et al. 2005; Hunter, Jing et al. 2005).

The methods described above generally consist of ensemble measurements made at very high peptide and lipid concentrations. Single-molecule fluorescence microscopy is an exciting field that

allows investigation of binding and diffusion events of individual molecules (Michalet, Kapanidis et al. 2003; Gell, Brockwell et al. 2006; Wayment and Harris 2009). Total-internal-reflection fluorescence (TIRF) microscopy has been used to image proteins and peptides in live cells, supported lipid bilayers, or immobilized on surfaces at the single-molecule level or in ensemble measurements (Mashanov and Molloy 2007; Fox, Wayment et al. 2009; Wayment and Harris 2009).

Scanning probe microscopy is an excellent method for visualizing the structures formed by peptides and proteins in membranes (Myhra 2004). A particular advantage with these approaches is the ability to perform experiments directly on aqueous membrane preparations. This enables the observation of 2D arrays of membrane proteins, and the types of lipid phases formed by peptide-lipid mixtures. Atomic force microscopy (AFM) provides high resolution and allows the measurement of forces between single biomolecules, potentially allowing a direct measurement of peptide–lipid interactions (Engel and Muller 2000; Frederix, Akiyama et al. 2003; El Kirat, Morandat et al. 2010; Alessandrini and Facci 2011). AFM, as well as electron microscopy (EM) and cryo-EM, allows the direct observation of the pores formed by a number of channel proteins.

To determine the penetration depth of a membrane-active peptide, spin labels or heavy atoms, such as dibromo or nitroxide derivatives, are introduced at varying depths of model membranes (Chattopadhyay and Raghuraman 2004; Matos, Franquelim et al. 2010). The pattern of tryptophan-fluorescence quenching as a function of the depth of the quencher can be used to generate angstrom-level resolution of the membrane-penetration depth at a specified tryptophan residue (Ladokhin 1999; Chattopadhyay and Raghuraman 2004).

However, the exact manner in which the peptides interact with membranes and molecular details of this process are still an area of active research and a matter of extensive debate and controversy. Different peptides utilize different interaction mechanisms or combinations of mechanisms that are not limited to one class of peptides or the other. Furthermore, the mechanisms of interaction can change depending on the parameters of the system, such as pH, temperature and concentration of peptide. It is clear that a better understanding of peptide-membrane interactions on a molecular level not only is important in the elucidation of various biological processes, but also could be instrumental in designing peptides with tailored functionalities, for example, for antibiotic and drug delivery applications.

1.3. Fluorescence as a method for studying peptide interactions

Fluorescence spectroscopy-based techniques have been extensively applied since the late 1960s to study various aspects of peptide-peptide, peptide-membrane and peptide-nucleic acids interactions. These techniques encompass measurements of fluorescence excitation and emission spectra, fluorescence time decays (lifetimes), fluorescence polarization (or anisotropy) and fluorescence microscopy using a large variety of fluorescent probes. Fluorescence as a tool for biological studies has three major advantages over other methods: high sensitivity, high speed, and non-invasivity.

High sensitivity. Sensitivity is the most important issue since fluorescence signals down to the single molecule level can be measured. Whereas absorbance measurements can reliably be used at concentrations down to concentrations of several tenths of a micromolar, fluorescence techniques can still be used at pico- and even femtomolar concentrations.

High speed of response. Using fluorescence, it is possible to monitor very rapid processes, since the response is limited only by the fluorescence lifetime, which is in the range 10^{-8}–10^{-10} s.

Non-destructive character. Due to its non-invasive character, the studied samples are not affected or damaged by the fluorescence process and no byproduct is generated.

As a consequence, fluorescence has proven to be a versatile tool for a large range of applications, as for instance:

- Monitoring binding of ligands to biochemical species, including *in vivo*
- measurement of distances within macromolecules and biological assemblies
- study of the dynamics of proteins folding
- measurement of ion concentrations inside living cells
- study of membrane structure and function
- cell imaging

Fluorescence spectroscopy is a very sensitive and powerful technique that can be applied in many fields such as chemistry, physics, biochemistry, biology, physiology, photochemistry, medicine, environmental science... Continuous improvement of instruments and probes and their applications are the result of the increasing popularity of fluorescence, observed for the first time over 150 years.

1.4. Principles of fluorescence

Fluorescence belongs to the general photophysical phenomenon called luminescence, which can be defined as the emission of light from any substance that occurs from an electronically excited state. Depending on the nature of the excited state, two types of luminescence can be considered: fluorescence and phosphorescence. If the emission occurs from the singlet excited state, the process is called *fluorescence*. The emission rates of fluorescence are typically about 10^8-10^9 s^{-1}. If the emission occurs from the triplet excited state, the process is called *phosphorescence*. In this case, transitions to the ground state are forbidden and the emission rates are slow, about 10^6-10 s^{-1}. Fluorescence occurs in a limited number of molecules (generally polyaromatic hydrocarbons or heterocycles) called fluorophores or fluorescent dyes. The process responsible for fluorescence is illustrated by the Jablonski diagram (Figure 1.8). Three main steps are important in the fluorescence process: excitation, non-radiative relaxation, fluorescence emission (Lakowicz 2006).

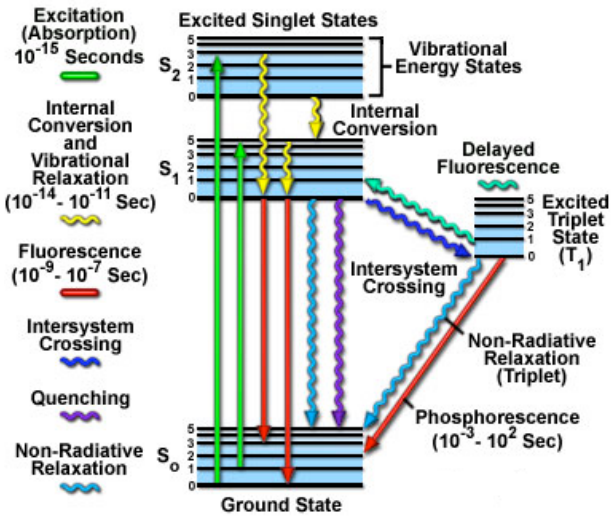


Figure 1.8. Jablonski diagram.
Taken from: www.olympusmicro.com/primer/java/jablonski/jabintro/index.html

Absorption of a photon by a molecule causes its excitation, and the electron moves from the occupied orbital of the ground state to an unoccupied orbital of the excited state. Each of these electronic states contains a number of vibrational levels. After the molecule has been excited, it rapidly relaxes from the higher vibrational states to the lowest vibrational state of the first excited electronic state. The rate for this relaxation is on the order of picoseconds. After reaching the lowest vibrational state of the first excited electronic state, the excited state can decay to the ground state by a number of mechanisms. The system can lose its excited energy by internal conversion (heat), quenching (external conversion), emission of a photon (fluorescence), or intersystem crossing (phosphorescence). Intersystem crossing produces a triplet state, where the spins of the excited and ground state electrons are no longer paired. It is important to mention that according to the selection rules of quantum mechanics, the crossing between two states with different multiplicities is forbidden. However, due to spin-orbital coupling between the orbital magnetic moment and the spin

magnetic moment, such crossing can take place. Other processes such as collisional quenching and fluorescence resonance energy transfer (FRET) may also depopulate S_1 and thus, together with internal conversion and intersystem crossing, compete with fluorescence.

Fluorescence emission is a radiative process in which a photon of energy $h\nu_{em}$ is emitted, with the fluorophore going back to its ground state S_0. As a consequence of the energy loss due to vibrational relaxation during the excited-state lifetime, the energy of the photon of fluorescence is lower, and therefore of longer wavelengths, than the excitation photon. The difference in energy or wavelength represented by ($h\nu_{ex} - h\nu_{em}$) is called the Stokes shift. This shift is an important parameter for the sensitivity of fluorescence techniques because it allows detecting emission photons with a low background, as they are easily discriminated from the excitation photons. In contrast, absorption spectrometry requires the measurement of transmitted light relative to high incident light levels at the same wavelength. An important characteristic of fluorescence is the fluorescence quantum yield, which is the ratio of the number of fluorescence photons emitted to the number of photons absorbed. Unless an irreversible reaction such as photobleaching occurs, the same fluorophore can repeatedly emit photons under constant illumination. The fact that a single fluorophore can generate a large amount of photons is of main importance for the high sensitivity of fluorescence detection techniques. For polyatomic molecules in solution, the discrete electronic transitions represented by $h\nu_{ex}$ and $h\nu_{em}$ are replaced by rather broad energy spectra called fluorescence excitation spectrum and fluorescence emission spectrum, respectively. The bandwidths of these spectra are important parameters for applications in which two or more different fluorophores are simultaneously detected. With only a few exceptions, the fluorescence excitation spectrum of a single fluorophore species in dilute solution is identical to its absorption spectrum (mirror effect). Under the same conditions, the fluorescence emission spectrum is nearly independent of the excitation wavelength (Kasha rule), due to the partial dissipation of the excitation energy during the excited-state lifetime (Figure 1.8).

1.4.1. Excited-state reactions

An excited-state reaction can be defined as a molecular process which occurs subsequent to electronic excitation and changes the structure of the excited-state fluorophore (Lakowicz 1999). Such reactions occur because light absorption changes the electron distribution within a fluorophore, which in turn changes its chemical and physical properties. An example of excited-state reaction is that of phenol, which in neutral solution can lose its phenolic proton in the excited state. Deprotonation occurs more readily in the excited state because the electrons on the phenolic hydroxyl groups are shifted to the phenol ring, making this hydroxyl group more acidic.

Another example of excited-state reaction is the mentioned above photoinduced charge transfer (PCT) that found wide applications in the construction of polarity sensitive fluorophores including labels for proteins. Among the well-known polarity sensitive probes are ANS (1-anilino-8-naphthalene sulfonate), TNS (p-toluidinyl-6-naphthalene sulfonate), PRODAN [6-propionyl-2-(dimethylaminonaphthalene)] and its derivatives. More details about them and their applications will be discussed in the following chapters.

To excited-state reaction belongs also excimer and exciplex formation that take places for example in pyrene, 2-phenylindole and anthracene, dimethylaniline. Intramolecular charge transfer (ICT) and twisted intramolecular charge transfer (TICT) accompanied with internal rotation are other examples of excited-state reactions that found applications in the design of viscosity sensitive fluorescent probes (molecular rotors).

Finally, an interesting example of excited-state reaction is the excited state intramolecular proton transfer (ESIPT) reaction. This ESIPT reaction is the basis of the probes developed in the present study. Its characteristics and applications will be described in the following chapters.

1.4.2. Solvent relaxation

The fluorescence process described above does not take into consideration the influence of the solvent that can strongly modify the fluorescence of the fluorophore. Solvent effects shift the emission to still lower energies owing to the stabilization of the excited state by the solvent molecules. Upon absorbing a photon of the appropriate energy ($h\nu_A$), the system is rapidly promoted to an excited singlet state (S^1). During this event, the system adopts a new electronic configuration with a dipole moment that differs significantly from that of the ground state. In many cases, the dipole shows an increased magnitude due to a photoinduced charge transfer (PCT). The electronic excitation occurs on a time scale that is much faster than that of the motions of atom nuclei (Frank–Condon principle). In the picosecond time-scale, the molecules of the solvent sphere reorient their dipoles to accommodate the new larger dipole of the fluorophore, which results in a higher ordered arrangement. This step, termed solvent relaxation, ultimately lowers the energy of the excited singlet state and narrows the energetic gap with the ground state (Figure 1.9). When the system finally returns to the ground state through a fluorescence event, the emitted photon is of a much longer wavelength (i.e. lower energy, $h\nu_F$) than that which was originally absorbed during excitation. This effect becomes larger with increasing solvent polarity, resulting in emission at lower energies or longer wavelengths. In general, only fluorophores that have high dipole moment in the ground or excited state display a large sensitivity to solvent polarity. Apolar molecules, such as unsubstituted aromatic hydrocarbons, or symmetric molecules, such as fluorescein, rhodamine and cyanine dyes are much less environmentally sensitive.

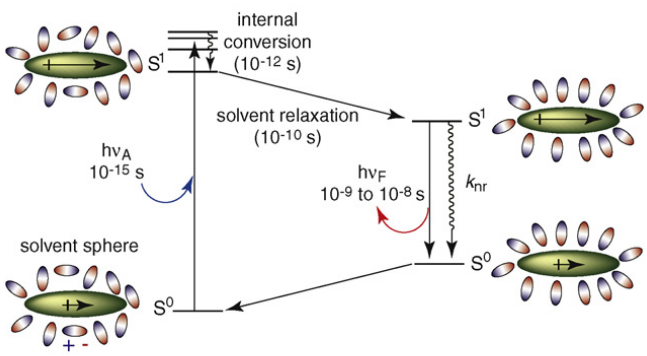


Figure 1.9. Solvent relaxation around a probe with a small dipole moment in its ground state and a large dipole moment in its excited state (Loving, Sainlos et al. 2010).

In the description of general solvent effects, the fluorophore is considered to be a dipole in a continuous medium of uniform dielectric constant. The interactions between the solvent and fluorophore affect the energy difference between the ground and the excited state. To a first approximation, this energy difference is a property of the refractive index (n) and dielectric constant (ε) of the solvent and is described by the Lippert-Mataga equation (Mataga 1970; Lippert 1975):

$$\bar{\nu}_a - \bar{\nu}_f = \frac{2}{hc}(\mu_e - \mu_g)^2 a^{-3} \Delta f + const \qquad (1)$$

where h is the Planck's constant, c is the velocity of light, $\mu_e - \mu_g$ is the change in dipole moment, a is the radius of the cavity in which the solute resides, *const* is a constant taking into account the non-radiative relaxation and Δf is the orientation polarizability defined as:

$$\Delta f = f(\varepsilon) - f(n^2) = \frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 + 1}{2n^2 + 1} \qquad (2)$$

Thus, according to this equation, the Stokes shift of a dye is proportional to the square of its transition moment and the orientation polarizability of solvent. It is frequently used to measure the transition dipole moment of new fluorophores based on spectroscopic measurements in solvents of different polarity. However, it should be noted that specific solvent-fluorophore interactions, such as H-bonding, produce deviations from the theory.

1.5. Natural and synthetic fluorescent dyes for protein labeling

Fluorescent dyes are widely applied in biology, biotechnology and medicine. Many organic dyes have been used to label proteins in order to investigate their dynamics, conformational changes and interactions. The selection of an appropriate dye and the corresponding fluorescence technique depends mainly on the biological system and the problem of study. Here, we will highlight only the most important examples of chromophores and their applications. All known fluorophores could be divided into two main classes based on their origin:

1) natural (intrinsic) fluorophores, which are synthesized by nature in living organisms;

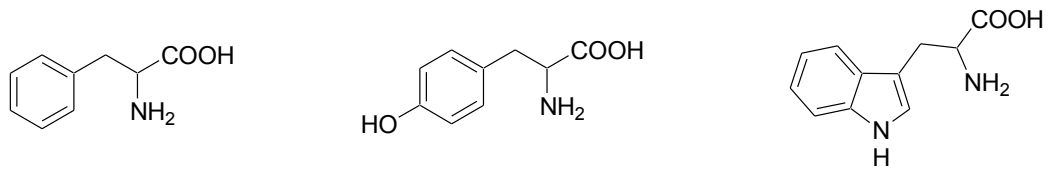
2) synthetic fluorophores, obtained artificially by synthesis.

The second could be further splitted into "classical" (environment-insensitive) and environment-sensitive dyes.

1.5.1. Natural fluorophores

Phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) aromatic amino acids serve as a natural fluorescent probes for proteins (Figure 1.10). Tryptophan has been widely used as the dominant fluorophore responsible for UV absorbance and emission in proteins (Demchenko 1986). In native proteins, the emission of Tyr is often quenched due to its interaction with the peptide chain or to energy transfer to Trp. Denaturation of proteins frequently results in an increased Tyr emission. Emission from Phe is observed only when the protein lacks both Tyr and Trp residues, which is a rare occurrence.

The main advantage of Trp residue is that it is constitutively present in native proteins. The emission of Trp is highly sensitive to its local environment. Due to this, it is often used to study protein dynamics and ligand binding (Demchenko 1986).



	Phenylalanine (Phe, F)	Tyrosine (Tyr, Y)	Tryptophan (Trp, W)
λ_{abs}	260 nm	275 nm	280 nm
λ_{em}	282 nm	304 nm	350 nm
ε	195 $M^{-1}cm^{-1}$	1400 $M^{-1}cm^{-1}$	5700 $M^{-1}cm^{-1}$
τ	7 ns	3-4 ns	3 ns

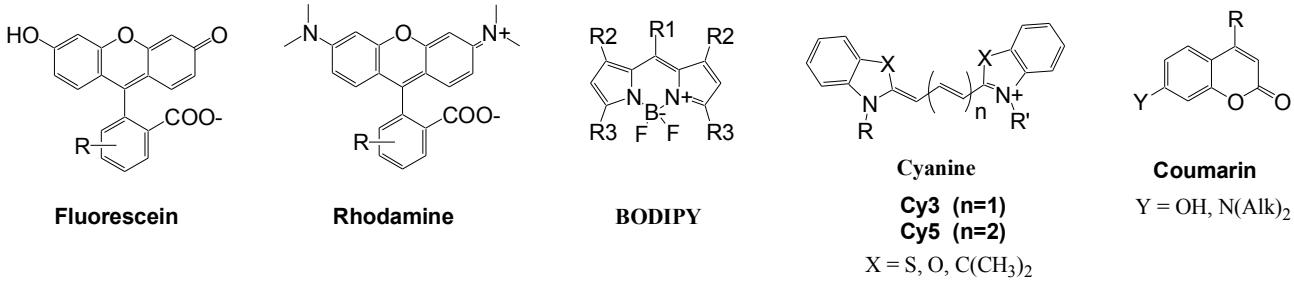
Figure 1.10. Natural fluorescent amino acids. λ_{abs}, λ_{em}, ε and τ are typical absorption and emission maxima, extinction coefficients, and lifetimes, respectively.

The main limitations of the aromatic amino acids, including Trp, as fluorescent dyes are their UV-range absorption, small extinction coefficient and relatively poor fluorescence quantum yield.

In addition, the complex photophysics of Trp makes interpretation of fluorescence data difficult and can often impair interpretation of fluorescence data (Lakowicz 1999). However, proteins usually possess more than one Trp residue, which complicates the assignment of a change in the fluorescence signal to a specific site. To overcome these problems, two approaches can be applied. The first one consists in removing Trp residues by site-directed mutagenesis, to reduce the complexity of the signal. The second approach consists in the biosynthetic substitution of Trp by its analogs. Among the noncoded Trp analogues studied so far, the most representative are 7-azatryptophan (7AW), 5-hydroxytryptophan (5HW) and 4-fluorotryptophans (4FW). These analogues display absorption and fluorescence properties that are red-shifted in respect with the Trp ones. They have been used to study ligand recognition and protein conformational changes (Twine and Szabo 2003). Nevertheless, their absorption and emission are still in the UV range, and site-selective substitution of natural amino acids by these Trp analogues is limited to peptides.

1.5.2. "Classical" synthetic dyes

The most well-known and probably the most common fluorophores are fluorescein and rhodamine dyes (Figure 1.11). Fluorescein was first synthesized by Baeyer in 1871 and remains one of the most popular dyes. Rhodamine is an amino-containing analog of fluorescein. It bears a positive charge delocalized between two amino groups through a conjugated aromatic system. Rhodamines are more photostable and less pH-sensitive than fluorescein.



Fluorescein **Rhodamine** **BODIPY** **Cyanine** **Cy3 (n=1)** **Cy5 (n=2)** X = S, O, $C(CH_3)_2$ **Coumarin** Y = OH, $N(Alk)_2$

Figure 1.11. Most common synthetic fluorescent dyes.

Another very common family is boron-dipyrrole (BODIPY) dyes, which were first discovered in 1968 by Treibs and Kreuzer (Figure 1.11). They exhibit optical properties that are often superior to fluorescein, tetramethylrhodamine, and other longer-wavelength dyes. They are strongly absorbing small molecules that emit a relatively narrow fluorescence band with high quantum yields. BODIPY dyes are relatively insensitive to the polarity and pH of their environment and have good photostability. Meantime, they have some disadvantages like poor water solubility and small Stokes shift. Nowadays, the chemistry and application of BODIPY dyes are increasingly developing due to their advantageous photophysical properties (Loudet and Burgess 2007; Ulrich, Ziessel et al. 2008).

Cyanine dyes (Mishra, Behera et al. 2000) are based on two aromatic or heterocyclic rings linked by a polymethine chain with conjugated carbon-carbon double bonds (Figure 1.11). This

family of dyes is characterized by high extinction coefficients and red-shifted absorption maxima. Both parameters depend strongly on the length of the polymethine chain. However, they have some drawbacks like short fluorescence lifetimes and low fluorescence quantum yields.

Coumarin dyes (Figure 1.11) is a relative old class of dyes, which found a number of biological applications mainly due to its very small size and good water solubility (Eggeling, Brand et al. 1997). However, its fluorescence properties, such as absorption in UV-region and low photostability are a clear disadvantage compared to all other dyes mentioned above.

Other families of dyes provided by commercial companies should be also mentioned such as Alexa, ATTO, DyLight, HiLyte Fluors, etc. These are recently developed fluorophores, characterized by significantly improved spectroscopic properties.

1.5.3. Environment-sensitive dyes

Environment-sensitive dyes are a special class of molecules displaying spectroscopic properties (e.g. fluorescence lifetimes, emission wavelengths, and quantum yields) that depend on the physicochemical properties of their environment (Lakowicz 2006). Such dynamic behaviour makes these species particularly well suited for investigating biomolecular interactions because it provides information at a highly localized region of the protein molecule (single amino acid level) (Loving, Sainlos et al. 2010). Their response to the environment is driven by excited state reactions in their fluorophore (conformational change, charge, electron and proton transfer, etc) and their non-covalent interactions with surrounding. These interactions could be classified as universal (van der Waals, dipole-dipole, dipole-external electric field, etc), and specific (H-bonding). Here, we do not consider pH- and ion-sensitive dyes as environment-sensitive, since the response of these dyes is associated with changes in their chemical structure: protonation/deprodonation or formation of a complex with an ion.

A widely used class of environment-sensitive fluorophores is solvatochromic dyes exhibiting shifts in their emission spectra as a function of polarity and hydration of their environment. These dyes exhibit strong changes in their dipole moments upon electronic excitation. Universal dipole-dipole and specific H-bonding interactions of these dyes with their surrounding change the energy of their electronic transition, and thus shift the maxima of their excitation and emission spectra (Mataga and Kubota 1970; Lippert 1975). A typical example of such dyes is Prodan (2-propionyl-6-dimethylaminonaphthalene, Figure 1.12) introduced more than 30 years ago (Weber and Farris 1979). In this fluorophore, the dipole moment increases dramatically upon electronic excitation due to an intramolecular charge transfer (ICT) from the electron donor, dialkylamino group, to the electron acceptor, carbonyl group (Figure 1.6). As a result, these dyes exhibit a red-shift of their emission spectrum in response to an increase in solvent polarity and relaxation rates of their surroundings (Parasassi, Di Stefano et al. 1994; Lakowicz 1999; Sykora, Kapusta et al. 2002). Moreover, an additional strong red-shift of the Prodan emission is connected with H-bonding to H-bond donor molecules in the surrounding (Catalan, Perez et al. 1991; Samanta and Fessenden 2000; Cerezo, Rocafort et al. 2001). This specific effect is typical for environment-sensitive dyes containing H-bond acceptor groups (such as carbonyl) and of particular interest to study

environment hydration (water is a strong H-bond donor). Most of these solvatochromic fluorophores are characterized by a low quantum yield in aqueous solution, due to electron transfer to water, but become highly fluorescent with strongly blue shifted emissions in non-polar solvents, or when bound to hydrophobic sites in proteins or membranes, due to efficient screening of these molecules from bulk water (Slavik 1982). Prodan has been used in many biological systems, notably in lipid bilayers and proteins presenting a binding site for this fluorophore (Hiratsuka 1998; Kaur and Horowitz 2004; Chakrabarti, Kelkar et al. 2006), (Haskard and Li-Chan 1998; Moreno, Cortijo et al. 1999; Kaur and Horowitz 2004; Sykora, Jurkiewicz et al. 2005; Bagatolli 2006). Derivatives of Prodan have also been attached covalently to proteins *via* reaction of thiols with Acrylodan (6-acryloyl-2-dimethylamino-naphthalen) (Baudier, Glasser et al. 1986; Lehrer and Ishii 1988; Harikumar, Pinon et al. 2002) or Badan (6-bromoacetyl-2-dimethylaminonaphthalene) (Hiratsuka 1999). However, the key disadvantage of this dye is its absorption in ultraviolet (360 nm), which limit its applications in cellular studies.

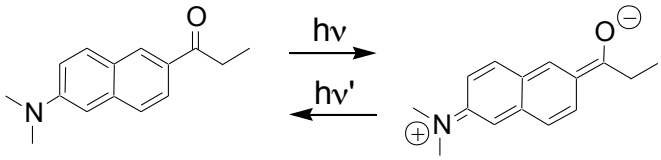


Figure 1.12. Typical solvatochromic fluorescent dye (Prodan) and its excited state intramolecular charge transfer.

In order to shift the absorbance of Prodan to the red, its benzo-analogue, 2-propionyl-6-dimethylaminoanthracene (Anthradan) was synthesized (Lu, Lord et al. 2006) (Figure 1.13). The main advantage of this red-shifted dye is the possibility to avoid interference from the autofluorescence of many biological components and allow a more favorable excitation wavelength for fluorescence microscopy applications. However, this dye showed decreased extinction coefficient and quantum yield as compared to parent Prodan.

Another example of Prodan improvement is substitution of its naphthalene core with a fluorene one, which presents longer electronic conjugation together with rigid conjugated structure. The synthesized fluorene derivative exhibits a 2-fold larger solvatochromism and absorption coefficient, a manifold larger two-photon absorption cross section, red-shifted absorption and emission, and higher photostability compared to Prodan (Kucherak, Didier et al. 2010).

Another typical example of environment-sensitive dyes is phtalimide-derivatives (Figure 1.13). The 4-(*N*,*N*-dimethylamino)phthalimide (4-DMAP) has been recognized as a fluorophore being highly responsive to the polarity and viscosity changes often found in biological media. 4-DMAP exhibits a 70-fold change in its quantum yield and a 110 nm fluorescence emission shift when transferred from diethyl ether to water (Soujanya, Fessenden et al. 1996). Moreover, 4-DMAP has a size comparable to that of tyrosine and tryptophan. However, it also shows an absorption in the UV region (380-390 nm) and a very low extinction coefficient. To improve its properties, the 4-DMAP fluorophore was extended, giving a new environment-sensitive dye, the 6-N,N-dimethylamino-2,3-naphthalimide (6-DMN) (Vazquez, Blanco et al. 2005), (Figure 1.13). This fluorophore displays interesting fluorescence properties with emission in the 460-590 nm range and

combined (fluorescence intensity and position of the maximum) response to changes in the environment polarity, though its absorption properties were not considerably improved. To overcome this drawback, 4-N,N-Dimethylamino-1,8-naphthalimide (4-DMN), an analogue of 4-DMAP, was synthesized. This new probe constitutes a significant improvement, since it shows an absorption maximum shifted to 440 nm (Loving and Imperiali 2008). It should be noted that 4-DMAP and all its analogues are low fluorescent in water, but exhibit high fluorescence in aprotic media. These properties are important for intensiometric detection of molecular interactions (Loving and Imperiali 2008). However, their low fluorescence in water makes them inefficient when the label is water exposed at all steps of the interaction process.

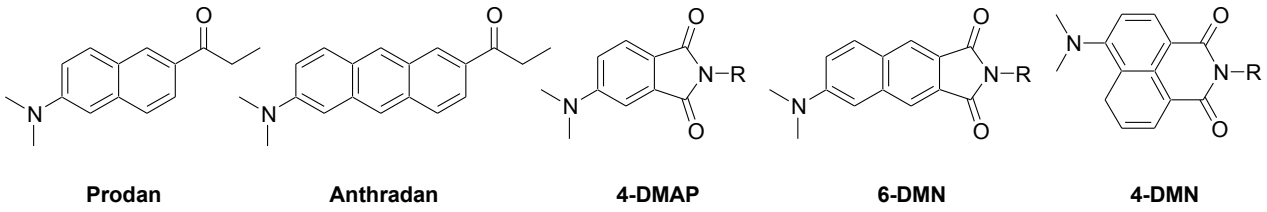


Figure 1.13. Prodan, its analogue Anthradan, phtalimide and naphthymide derivatives.

In the search for advanced environment-sensitive dyes, a bichromophoric dye Fluoroprobe (Figure 1.14) was developed (Mes, De Jong et al. 1984). This dye exhibits a charge transfer through space, which generates an exceptional transition dipole moment (27 D) and thus, exceptional solvent sensitivity. For the moment, Fluoroprobe remains the most environment-sensitive dye. However, this dye found no applications in biology, due to the extremely strong quenching of its fluorescence in protic media, as well as its UV absorption maximum (308 nm) and very low extinction coefficient.

Dansyl label (DNS) was one of the first solvatochromic labels used for protein labeling (Figure 1.14), which can be excited at 350 nm, where proteins do not absorb. The emission color and emission intensity of the dansyl moiety are also highly sensitive to solvent polarity, and emission maximum appears typically near 520 nm (Holmes-Farley and Whitesides 1986; Goncalves 2009). Moreover, the small size of DNS label makes it appropriate for biological applications.

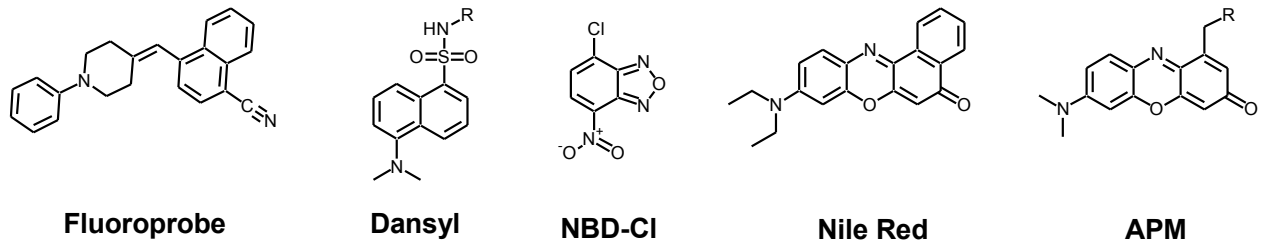


Figure 1.14. Advanced fluorescent solvatochromic dyes.

Another example of a label highly sensitive to solvent polarity is 7-nitrobenz-2-oxa-1,3-diazol-4-yl chloride (NBD-Cl) (Figure 1.14), which absorbs at 470 nm and emits at 530 nm

(Chattopadhyay 1990). NBD, due to its reactive chlorine group can be readily conjugated with the amino groups of proteins allowing probing polarity at the level of the labeling site (Lakowicz 1999).

Notably, Nile Red (Diaz, Melis et al. 2008) dyes and its analogues (Figure 1.14) represent environment-sensitive dyes which show significantly red shifted absorption and emission. The absorption maximum of Nile Red (Greenspan, Mayer et al. 1985) and its recently introduced analogue APM (Cohen, Pralle et al. 2005) are located at 530 and 560 nm, respectively. Moreover, the extinction coefficient and the fluorescence quantum yield of these dyes are relatively good. However, these favorable spectroscopic properties are compensated by a sever disadvantage: poor fluorescence solvatochromism and tendency to aggregate due to their high hydrophobicity. APM has a short linker between the probe and the protein, ensuring that it can closely follow the motions of the side chain to which it is attached. The spectral characteristics of APM provided information on the conformation changes of the water-exposed domain of the β2 adrenergic receptor during its interactions with ligands (Cohen, Pralle et al. 2005).

Noticeably, labeling of the amino acid side-chains with described above solvatochromic dyes increases the distance between the fluorophore and the protein backbone, which may reduce the sensitivity of the label to the local changes in the environment of the labeling site. To overcome this problem of non-specificity, fluorescent amino acids were synthesized.

1.5.4. Environment-sensitive amino acids

The Prodan-based amino acid, 6-dimethylaminonaphtoyl alanine (Aladan, Figure 1.15), was synthesized by two independent research groups (Cohen, McAnaney et al. 2002; Nitz, Mezo et al. 2002). In both cases, the substituted amino acid was prepared via an enantioselective synthesis as shown on Figure 1.9 and site-specifically incorporated into peptides. The probe was used to sense binding of the phosphoserine peptide to 14-3-3 protein involved in cell cycle control (Nitz, Mezo et al. 2002) and to estimate the local dielectric constant of the B1 domain of the streptococcal protein G at different sites (Cohen, McAnaney et al. 2002). Later, this amino acid was successfully applied in the study of δ-opioid receptor antagonist binding (Chen, Chung et al. 2005).

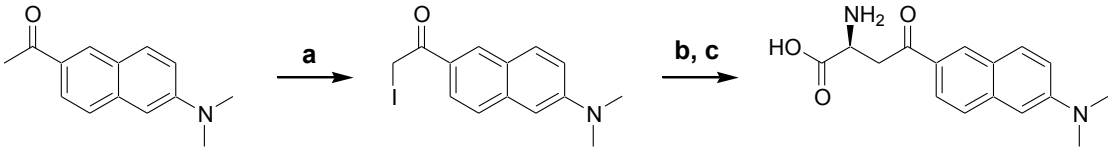


Figure 1.15. Scheme of Aladan (6-dimethylaminonaphtoyl alanine) synthesis (a) $LiN(SiMe_3)_2$/THF, -78°C, then I_2; (b) $Ph_2C{=}NCH_2CO_2$-*t*-Bu, catalyst, CsOH, DCM, -70°C; (c) TFA, EDT.

Based on 4-(*N*,*N*-dimethylamino)phthalimide (4-DMAP), another environment-sensitive amino acid (4-N,N-dimethylaminophthalimidoalanine – 4-DAPA) (Figure 1.16) was synthesized and incorporated into a peptide using standard Fmoc-based solid-phase peptide synthesis (SPPS). The label was able to sense the phosphorylation-dependent binding of the synthesized peptide to 14-3-3 protein (Vazquez, Rothman et al. 2004). Recently, the 6-DMNA amino acid (6-N,N-

dimethylamino-2,3-naphthalimidoalanine) (Figure 1.16) has been incorporated into peptide motifs that are recognized by SH2 phosphotyrosine binding domains, (Vazquez, Blanco et al. 2005) PDZ domains, opioid receptors, and class II MHC proteins (Venkatraman, Nguyen et al. 2007) and has proven very effective for monitoring binding. However, due to the intrinsic strain of the five-membered phthalimide ring of 4-DMAP and 6-DMN, these fluorophores are susceptible to nucleophilic attack, leading to the formation of ring-opened byproducts.

To overcome this problem, a new solvatochromic amino acid, 4-N,N-dimethylamino-1,8-naphthalimidoalanine (4-DMNA) was recently developed (Loving and Imperiali 2008). This amino acid shows the same solvatochromic properties than the dimethylaminophthalimide dyes. Moreover, it displays some advantages like an improved chemical stability due to incorporation of the six-membered imide ring of 4-N,N-dimethylamino-1,8-naphthalimide (4-DMN). Like 4-DAPA and 6-DMNA, the comparable size of 4-DMNA to the natural aromatic amino acids allows it to be incorporated into peptides or proteins without introducing a significant perturbation of the native peptide structure. To investigate study protein-protein interactions, 4-DMNA was introduced into a peptide that is recognized by calmodulin. The binding interaction between these two partners yields an increase in fluorescence emission greater than 900-fold.

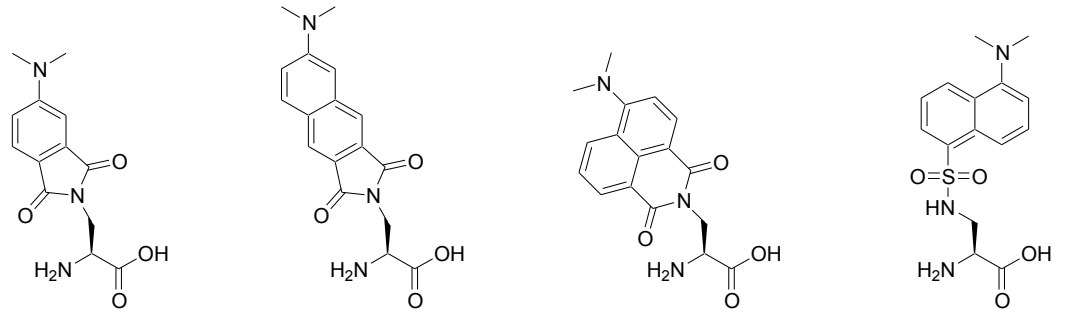


Figure 1.16. Structure of the solvatochromic amino acids 4-DAPA, 6-DMNA, 4-DMNA and dansylalanine.

Summerer et al. reported a strategy for selective and efficient biosynthetic incorporation of a low molecular weight fluorophore into proteins at defined sites (Summerer, Chen et al. 2006). Dansyl chloride was used in the synthesis of the fluorescent amino acid 2-amino-3-[5-(dimethylamino)naphthalene-1-sulfonamide] propanoic acid (dansylalanine) which displays small size, relatively large Stokes shift, and a high degree of sensitivity of its emission wavelength and quantum yield to environmental effects such as ligand binding, complex formation, conformational changes, or protein unfolding. Dansylalanine was incorporated into human superoxide dismutase through the aminoacyl-tRNA/nonsense codon technique and used to monitor unfolding of the protein in the presence of guanidinium chloride. This amino acid is potentially useful for in vitro studies of protein structure and function, and molecular interactions, due to its relatively small size and its sensitivity to the local environment (Goncalves 2009).

These examples clearly demonstrate that environment-sensitive probes are of great interest for protein investigation. Nevertheless, there is still a lack of probes with optimal parameters. Most

of the dyes described in this chapter (except APM) absorb in the UV range that is often not suitable for biological applications. Excitation at longer wavelengths would decrease significantly the photo-damage of the biological samples and the photo-degradation of the dye itself. Moreover, the low fluorescence quantum yield in aqueous media of most of the described dyes limits their application for the investigation of small peptides for which the label is exposed to water. Thus, there is a strong need for new fluorophores or derivatives of known fluorophores with improved water solubility, where the excitation and emission maxima are red-shifted and the fluorescence quantum yields are high. Finally, all dyes described above exhibit only a single emission band and respond to an environment change merely by a shift of their maximum emission wavelength. As a consequence, these dyes provide only limited information on local conformational changes. As it will be shown in the present work, a new class of environment-sensitive labels exhibiting dual fluorescence presents important advantages for protein labeling.

In conclusion, fluorescent labels are powerful tools for investigating biomolecular events. Their utility is amplified when their spectral properties are sensitive to the environment. Environment-sensitive probes allow real-time monitoring of binding events and accompanying conformational changes of biomolecules of interest by using a single labeling approach. For optimal sensitivity, dye incorporation in close proximity of the interaction site is commonly required.

1.6. 3-Hydroxychromones (3HCs): ESIPT-based environment-sensitive dyes

1.6.1. Excited State Intramolecular Proton Transfer (ESIPT)

Excited State Intramolecular Proton Transfer (ESIPT) reaction was observed and characterized for the first time in salicylic acid ester (Figure 1.17) (Weller 1956). The methylsalicylate ion displays an unusual large Stokes shift in comparison to its methylated analog methyl 2-methoxybenzoate, which was attributed to an excited state proton transfer from the hydroxyl to the carbonyl group. This process decreases the energy of the excited state species, giving a red-shifted emission band. Later, similar spectroscopic properties were observed for different carbonyl and heterocyclic compounds.

Figure 1.17. Methyl 2-methoxybenzoate (λ_{fl}= 360 nm) and methylsalicylate (λ_{fl}= 450 nm).

The main governing force for the ESIPT reaction is related to the changes in the acidicity of the proton donor hydroxyl-group and the basicity of the proton-acceptor carbonyl in the excited state (Formosinho and Arnaut 1993). Due to the charge redistribution, the carbonyl group becomes strongly basic, while the hydroxyl group – strongly acidic, which results in the proton transfer to give the tautomer form (T*). Both the rate and the equilibrium of such a process are controlled by the proton transfer energy barrier and the energy difference between the normal (N*) and tautomer (T*) forms of the excited molecule (Figure 1.18). In the ground state, the normal (N) form is strongly favored. After excitation, the tautomer (T*) form may present lower energy and thus, the ESIPT reaction becomes favorable. The ESIPT reaction can be characterized by its rate, which depends on the energy barrier (ΔE_{act}). If ΔE_{act} is too high, the ESIPT will not occur. In contrast, if ΔE_{act} is low enough, the equilibrium between the N* and T* forms can be reached during the excited state lifetime, and will be governed by the energy level difference between the two excited states (E_{N*}-E_{T*}).

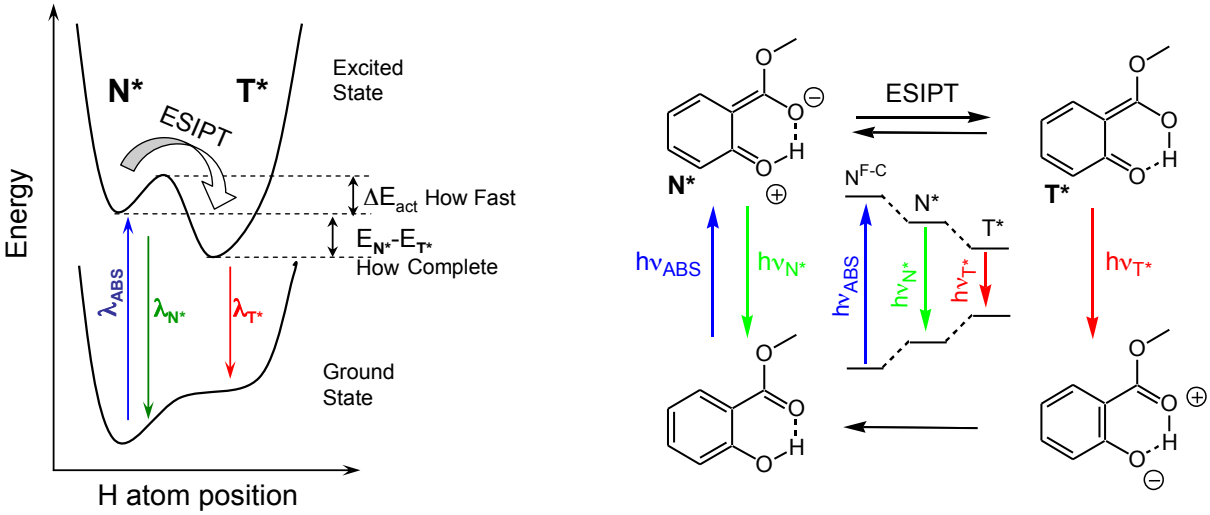


Figure 1.18. Energy diagram of the ESIPT reaction (left). Example of ESIPT in methyl salicylate (right).

The sequence of events that results in two emission bands is shown on Figure 1.18. Absorption of a quantum of light $h\nu_{ABS}$ generates the normal Franck–Condon state N^{F-C}, which then relaxes to the N* state followed by relaxation of the solvent shell. Then, the N* state can emit a quantum of light $h\nu_{N^*}$ or undergo ESIPT to generate the T* state. Since the ESIPT reaction occurs simultaneously or slower than the solvent relaxation rate (Swinney and Kelley 1993), the solvent environment of the emissive T* state can be considered relaxed. The emission of a quantum of light $h\nu_{T^*}$ results in populating the solvent-unrelaxed ground T state with subsequent relaxation and, proton back transfer to the ground N state, which closes the cycle.

The ESIPT reaction is probably one of the fastest chemical processes described to date. Its rate can vary in a rather large time-scale range (from 10^{-14} to 10^{-8} s). Only in the 90s, new instrumental methods appeared which allowed studying such processes with enough precision, namely the so-called "ultra-fast laser spectroscopy" methods, including the ultra-fast transient absorption (TA) and the time-correlated single photon counting (TCSPC) methods.

In most cases described so far, the ESIPT reaction occurs between the hydroxyl group and the oxygen atom of the carbonyl (Ormson and Brown 1994) or the sp^3 hybridized nitrogen atom (Ormson and Brown 1994). In salicylic acid, the hydrogen bond results in the formation of a six-member cycle. The short length of this bond (~2 Å) is optimal for proton transfer, explaining the low energy barrier of the process (<15.5 kJ/mol) and its very high rate ($k \sim 10^{13}$ s^{-1}) (Herek, Pedersen et al. 1992). In most cases, the salicylic acid derivatives exhibit only one fluorescence band that belongs to the tautomer form, due to the large energy difference between the two excited–state forms (~ 37 kJ/mol).

The ratio of intensities of the N* and T* bands (I_{N^*}/I_{T^*}) in the emission spectrum is an important parameter. The influence of electron-donating substituents on the ratio was investigated with 4- and 5-methoxysalicylic acid esters. In these derivatives, the methoxy group increases the electron density on the substituent in the *para*-position. As a result, the basicity of the carbonyl

group increases for the 4-methoxy derivative while the acidity of the hydroxyl group decreases in the case of the 5-methoxy derivative. In cyclohexane, the band of the tautomer form dominates in the fluorescence spectrum of 4-methoxy-methylsalicylate, while in the case of 5-methoxy-methylsalicylate, the band of the normal form dominates (Acuna, Toribio et al. 1985) (Figure 1.19). Thus, the ratio of the two emission bands could be modulated by the functional groups.

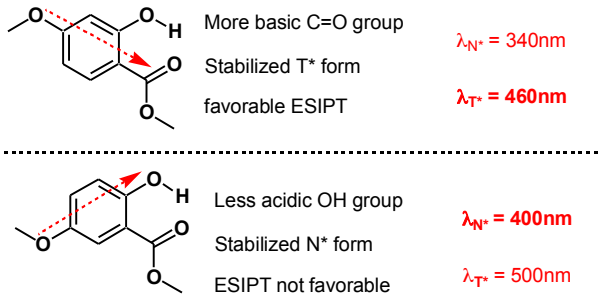


Figure 1.19. Effect of substituents on ESIPT in 4-methoxy and 5-methoxy methyl salicylate.

Among compounds that exhibit an ESIPT reaction of particular interest are 3-hydroxychromones (3HC), which form an intramolecular H-bond through a 5-membered cycle (Figure 1.20) providing a pathway for the ESIPT reaction (Sengupta and Kasha 1979).

1.6.2. 3-Hydroxychromone dyes

Chromones (chromen-4-ones) are natural dyes, which are widespread in the vegetal kingdom. Most natural chromones, containing phenyl (aryl) substituents in position 2 of the heterocycle, are commonly called flavones (Figure 1.20). The flavone derivatives bearing a hydroxyl group in position 3 of the heterocycle are called 3-hydroxyflavones (3HF). These compounds attracted the attention of spectroscopists after the discovery of their dual fluorescence. Later, Sengupta and Kasha interpreted this phenomenon by an excited state intramolecular proton transfer (ESIPT) reaction (Sengupta and Kasha 1979). Importantly, the pathway for ESIPT in 3-hydroxychromones is provided by an intramolecular H-bond through a 5-membered cycle (Figure 1.20), which is much weaker than in the 6-membered cycle presented by other ESIPT systems. Therefore, it can be easily disrupted by H-bonding interactions with the environment (McMorrow and Kasha 1984), so that 3HC dyes can exhibit two emission bands of comparable intensities: the short-wavelength band corresponding to the normal excited state (N*) and the long-wavelength band to the photo-tautomer product (T*) (Figure 1.21).

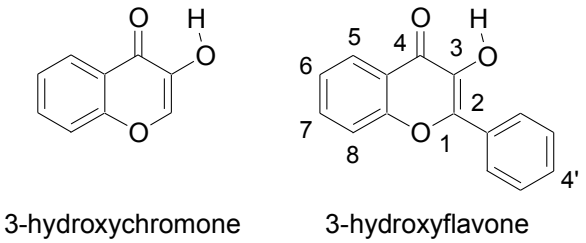


Figure 1.20. Chemical structure of 3-hydroxychromone and 3-hydroxyflavone.

While classical solvatochromic dyes respond by a shift in their emission maximum and fluorescence intensity in response to changes in their environment, such as polarity, 3HC dyes also respond through changes in the ratio of their two emission band intensities. Measuring the intensity ratio of the two emission bands is more precise and convenient than measuring band position or absolute intensities, since this ratio is independent from dye concentration and instrumental settings. Moreover, due to two-band emission, ESIPT dyes, as compared to one-band solvatochromic dyes, provide an additional channel of spectroscopic information for analysis of their environment.

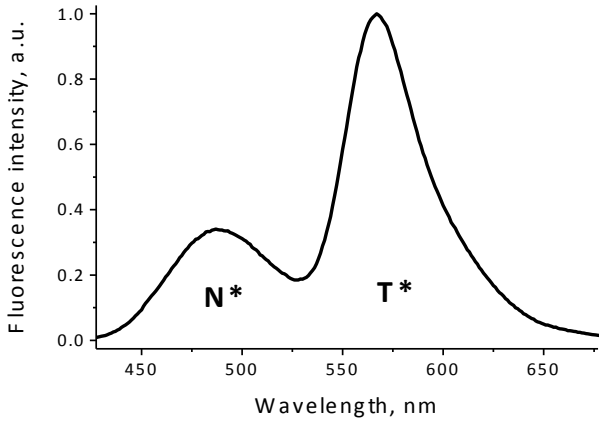
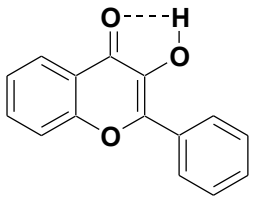


Figure 1.21. Fluorescence spectrum of 3-hydroxyflavone in methanol. Excitation wavelength was 350 nm.

1.6.3. Solvent effects on 3HC dyes

The non-substituted 3HF shows only a T* band emission in aprotic solvents, while a dual fluorescence is observed in protic environment (McMorrow and Kasha 1984) (Figure 1.22). The intensity ratio of the two emission bands (N*/T*) is higher in more polar alcohols making it possible to sense the environment polarity just by recording the fluorescence spectrum. Meanwhile, in aprotic media the N* band intensity is too low, thus rendering environment sensing difficult to perform through ratiometric measurements.



3-hydroxyflavone

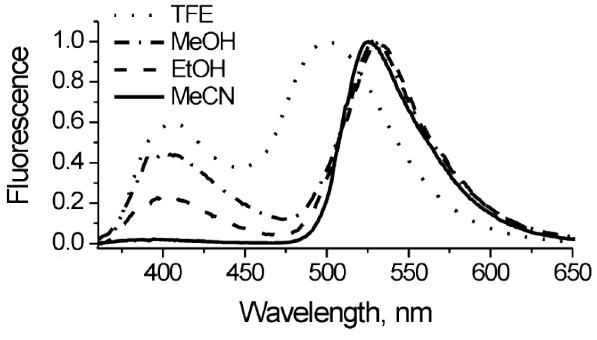


Figure 1.22. Structure of 3-hydroxyflavone (3HF) and fluorescence spectra of 3HF in protic and aprotic solvents.

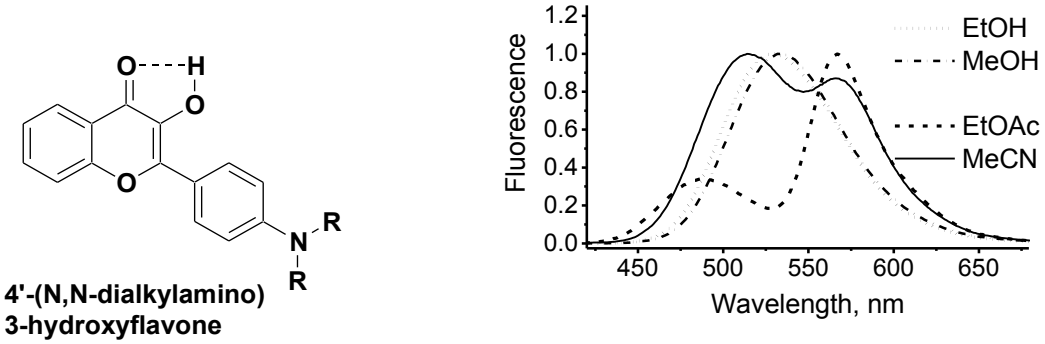


Figure 1.23. Structure of 4'-(N,N-dialkylamino)-3-hydroxyflavone and its fluorescence spectra in protic and aprotic solvents.

A significant progress was done after synthesis of dialkylamino derivatives of 3HF (Chou, Martinez et al. 1993; Ormson, Brown et al. 1994) (Figure 1.23). Due to the 4'-dialkylamino group, the N* excited state in these compounds exhibit a very large dipole moment, where the electronic charge is transferred from the dialkylamino group to the chromone moiety (Chou, Martinez et al. 1993; Swiney and Kelley 1993; Nemkovich, Pivovarenko et al. 2005). In contrast, the excited state tautomer T* exhibits much lower charge separation (Figure 1.24) and, thus, lower dipole moment (Chou, Pu et al. 2005; Yesylevskyy, Klymchenko et al. 2005). Therefore, the N* state, unlike the T* state, shows a significant solvatochromism. Indeed, on increase in the solvent polarity, the N* band shifts to the red, while the T* band remains almost unchanged (Figure 1.24). In addition, an increase in the solvent polarity results in an increase in the relative intensity of the N* emission band. Therefore, the intensity ratio of the N* and T* bands in emission, I_{N^*}/I_{T^*} , is an important indicator of solvent polarity (Ercelen, Klymchenko et al. 2002; Ercelen, Roshal et al. 2002; Klymchenko, Ozturk et al. 2002; Klymchenko, Pivovarenko et al. 2003; Klymchenko, Pivovarenko et al. 2003).

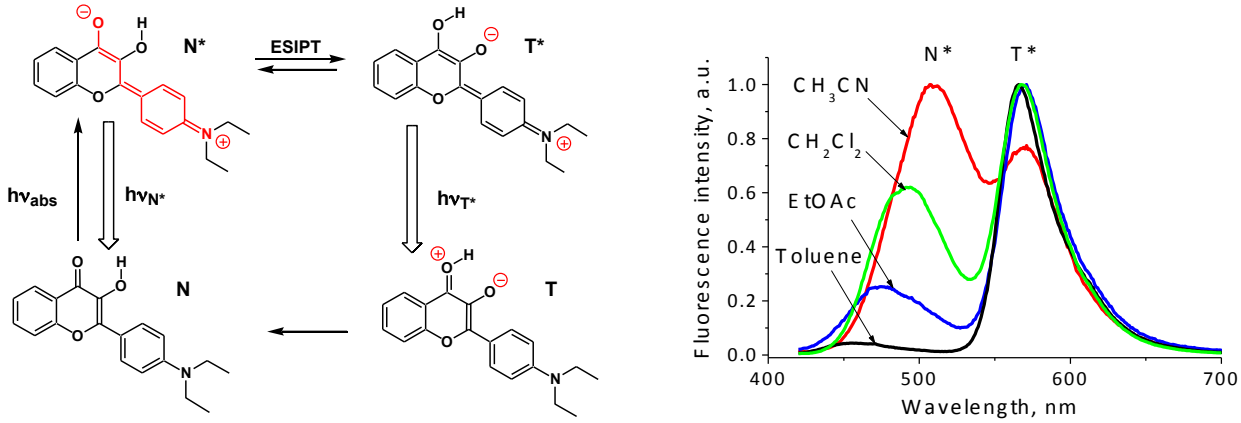


Figure 1.24. Photophysical cycle of 4'-(N,N-diethylamino)-3-hydroxyflavone dye (left). Fluorescence spectra of 4'-(N,N-diethylamino)-3-hydroxyflavone in aprotic solvents of different polarity (right).

3HF molecules allow the simultaneous determination of four parameters: ν_{abs}, ν_{N*}, ν_{T*} and I_{N*}/I_{T*}, which can differently characterize the physical properties of the microenvironment. An algorithm using these four spectroscopic parameters was proposed to simultaneously estimate three microenvironment characteristics: polarity, polarizability and H-bond donor ability (Klymchenko and Demchenko 2003). This algorithm was successfully applied in binary solvent mixtures, AOT reverse micelles and proteins.

Hydrogen Bonding. A particularly attractive feature of 3HFs is their high sensitivity to hydrogen bonding. The ESIPT reaction site is strictly localized between the 3-hydroxyl and 4-carbonyl groups, which form an H-bond that closes a low-stable five-member ring (Sengupta and Kasha 1979). Therefore, this reaction shows an extreme sensitivity to intermolecular H-bond perturbation (McMorrow and Kasha 1984). 3HFs show two emission bands in alcohols; meanwhile in aprotic solvents of similar polarity, the T* band strongly dominates. This effect could be better observed in dialkylamino derivatives of 3HCs, namely 4'-diethylamino-3-hydroxyflavone (FE) (Klymchenko and Demchenko 2003). The I_{N*}/I_{T*} ratio of FE was shown to be up to 10 times higher in protic solvents than in aprotic solvents of the same polarity. This high sensitivity was related to the H-bonding of solvent proton donors with the 4-carbonyl group of FE acting as an H-bond proton acceptor (Figure 1.25).

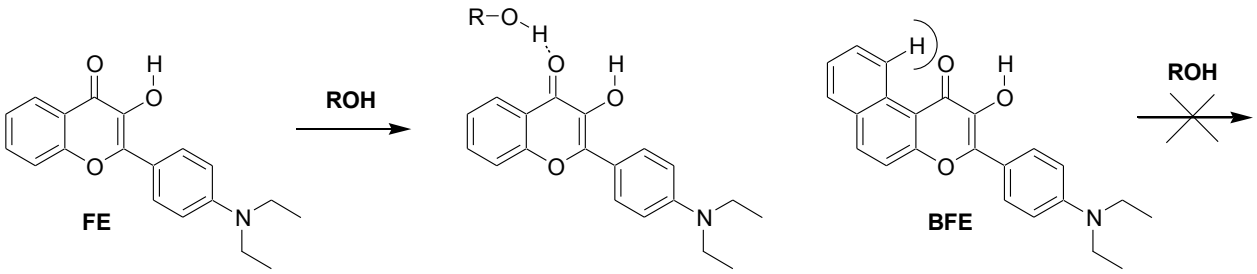


Figure 1.25. Formation of an H-bonding complex of FE with alcohol is blocked by the additional benzene ring in BFE.

Interestingly, a benzo-analogue of FE, 3-(4-diethylamino-phenyl)-2-hydroxy-benzo[f]chromen-1-one (BFE), with a 4-carbonyl group sterically protected from the intermolecular H-bonding (Figure 1.25) does not show any sensitivity of its dual emission to protic solvents. These results suggested that only H-bonding between protic solvents and the carbonyl group affects the dual emission of dialkylamino substituted 3HFs, while H-bonding of the solvents with the 3-hydroxy group are not detected in the fluorescence spectra (Klymchenko, Pivovarenko et al. 2003).

1.7. Approaches for protein labeling

It has to be considered that among the 20 natural amino acids, only Tyr and Trp display significant emission, and of those two, only the Trp fluorescence is sensitive to its immediate environment. However, Trp suffers from strong photobleaching, and its excitation and emission are frequently masked by the strong fluorescence background typical of biological media. Therefore, applications that require measuring fluorescence in complex mixtures will demand the introduction of extrinsic fluorophores with long-wavelength excitation and emission fluorophores.

There are several ways to label a protein:

a) direct labeling of native proteins;

b) incorporation of a labeled amino acid by peptide synthesis;

c) indirect labeling;

d) translational incorporation of nonnatural amino acids.

1.7.1. Direct labeling of native proteins

A large number of methods for site-selective chemical modification of proteins with extrinsic fluorophores has been developed (Haugland 2005). The most appropriate one strongly depends on the presence and accessibility of targeted functional groups. The most useful protein functionalities for specific chemical modification are the natural amino acid side chains of cysteine (Cys), lysine (Lys), tyrosine (Tyr), as well as the α-N-terminus of the peptide backbone (Thordarson, Le Droumaguet et al. 2006).

Different reagents have been developed for labeling amino groups of proteins. Active esters and isothiocyanates are probably the most common reactive groups for amino group modification. The reaction is easy to control and it occurs over a period of a few minutes. Higher pH favors the -NH_2 form of lysine groups, which improves the labeling efficiency. Aldehydes can react reversibly with primary amino groups resulting in a Schiff base (imide), which can in turn be reduced with $NaCNBH_3$ to form a secondary amine linker. As the ε-amine group of Lys and the α-N terminus have different pKa values, (about 10 and 7.8, respectively), their difference in reactivity could be used for the selective functionalization of the α-N-terminus. Thus, only the α-N terminus will react with aldehydes at pH 5. Sulfonyl chlorides are highly reactive for amino group labeling, though they are not selective. Therefore they are mainly used in cases where a single amino group is available.

Cysteine contains a mild nucleophilic thiol group in its side chain. Under appropriate conditions, Cys can be modified selectively, rapidly, and quantitatively. In general, cysteine residues are ideal because this amino acid occurs relatively infrequently in proteins (de Graaf, Kooijman et al. 2009) and possess excellent nucleophilic properties under most physiological conditions. Proteins with unique cysteines can be prepared readily by site directed mutagenesis, particularly in cases where proteins do not contain any Cys residue. Labeling of Cys can be achieved using maleimides and halocarbonyl compounds (usually iodine), along with amine-selective acylating agents such as the O-succinimidyl esters (Figure 1.26a). The reaction can be performed in a pH range (pH 7.0-8.0) where alkylation of cysteines is much more efficient than

alkylation of Lys amino groups, which are protonated in this pH range. Maleimide-functionalized groups react as Michael acceptors with Cys. This method works also at mild conditions (proteins can be modified at neutral pH) and greater specificity (side-reaction with amines occurs only at pH >8). Disulfide formation, especially with dyes and heterofunctional spacers terminated with the o-pyridyldisulfide group, is the most specific method for functionalization of Cys. It is also a fully reversible reaction through the use of standard reducing agents such as dithiothreitol (DTT).

The direct labeling approach, particularly thiol-labeling, is frequently used with environment-sensitive dyes. However, as it was already mentioned, the main disadvantage of this method is the specificity of labeling, purification of the conjugates and inability of this method to label proteins *in situ* inside the cell.

1.7.2. Tag-labeling

Specific fluorescence labeling of proteins inside the cell can be provided by specific protein tag sequences, which can be labeled with high specificity *in situ* by a specially designed fluorescent dye. These short tag sequences of amino acids can be expressed together with the protein of interest and then further recognized by a fluorescent ligand, in order to obtain labeled proteins (Griffin, Adams et al. 1998; Kapanidis, Ebright et al. 2001).

This approach has been notably used with As^{III} compounds, which can form covalent linkages with closely spaced paired cysteine-thiol groups in proteins. Based on this interaction, an interacting pair has been developed that consists of a small peptide tag with tetracysteine (TC) motif and a biarsenical probe (Griffin, Adams et al. 1998; Sadhu, Mizukami et al. 2011). In the TC part, two cysteine pairs are separated by a spacer of two different amino acids (CCXXCC), and this motif binds to organo-arsenical compounds under reducing conditions. This successful strategy was named “fluorescein arsenical hairpin binder” (FlAsH). The interaction results in an extremely high (~10 pM) affinity. Incorporation of the green fluorophore FlAsH or the red fluorophore ReAsH are the most known applications of this scheme (Lin and Wang 2008).

Though the tag-labeling provides a possibility for intracellular specific labeling, this approach is not very useful for application of environment-sensitive dyes. Indeed, due to the relatively large size of the tag, the fluorophore will be positioned very far from the site of interaction or the protein conformational change, and thus will not be sensitive to them. Though there were some attempts to apply environment-sensitive dyes (biarsenic derivative of Nile Red) for tetracysteine-tag protein labeling (Nakanishi, Maeda et al. 2004), this labeling approach is mainly used with “classical” dyes for tracking and FRET-based techniques.

1.7.3. Solid-state synthesis of labeled peptides

Relatively short peptides could be prepared synthetically. This method is faster than recombinant protein production in bacteria and has no limitation concerning the use of non-natural amino acids and post-synthesis modifications. Nowadays, solid-phase peptide synthesis (SPPS), pioneered by Merrifield (Merrifield 1963), allows synthesis of almost any peptide or protein of less than 100 amino acids. In SPPS, the peptide remains covalently attached to the solid bead until the

end of the synthesis. The general principle of SPPS consists in repeated coupling-deprotection cycles. The two most common types of SPPS chemistry, Fmoc- and Boc-based, differ by the use of amino acid N-protecting groups, resins and deprotection reagents. In our laboratory, we use Fmoc-based chemistry. It uses repeatedly piperidine as a Fmoc deprotection agent followed by full deprotection and removal from the resin by trifluoroacetic acid.

Incorporation of the labels during SPPS could be done in one of the following ways: (a) using synthetic fluorescent amino acids and (b) by direct labeling on resin. Synthetic fluorescent amino acids could be incorporated into any position of the peptide during the SPPS. Such a synthesis is useful for relatively small fluorophores incorporated instead of hydrophobic amino acids to limit the perturbation of the natural structure of the peptide. On-resin labeling of peptides is a complementary approach. The simplest version of on-resin labeling consists in coupling the fluorophore to the terminal amino group after a Fmoc-deprotection step.

1.7.4. Incorporation of nonnatural amino acids via synthesis and semisynthesis

To overcome the peptide length limitations of SPPS, efficient strategies have been developed to ligate synthetic peptides together to make larger proteins, including a chemoenzymatic strategy (Jackson, Burnier et al. 1994) and chemical ligation (Schnolzer and Kent 1992).

Chemical engineering of proteins provides a pool of various synthetic and semisynthetic methods. Dawson et al. (Dawson, Muir et al. 1994; Dawson and Kent 2000) introduced a simple and elegant method called native chemical ligation (NCL) for the synthesis of peptides by condensation of their unprotected segments. The coupling of synthetic peptide-thioesters with peptides carrying an N-terminal Cys leads to an amide-bond at the ligation site. This approach has proven to be useful for the synthesis of smaller proteins up to 120 amino acids in length; larger proteins cannot be obtained easily in one ligation step. Multistep NCL of different peptide-segments, however, can lead to larger proteins (Becker, Hunter et al. 2003). An extension of this NCL strategy is the expressed protein ligation (EPL) method (Muir, Sondhi et al. 1998).

Expressed protein ligation (Muir, Sondhi et al. 1998; Muir 2003) is an extension of the NCL method. A recombinant Ca-thioester reacts with a chemically synthesized or expressed peptide/protein possessing an N-terminal Cys under the conditions of NCL to form a native peptide bond. The synthetic peptide that is prepared constitutes the omitted portion of the native protein, but includes a fluorescent amino acid in place of another residue. The chemically synthesized section of the protein can be as small as possible whereas the expressed part is not limited in size. This can lead to very large specifically labeled proteins. The chromophore can be inserted either as an amino acid building block (e.g. Fmoc-protected amino acid) during SPPS, or afterwards by labeling a sidechain residue with an appropriate electrophilic derivative of the chromophore (represented as a red circle at Figure 1.26b). This ligation method combines the advantages of molecular engineering and chemical peptide synthesis in many cases and allows site-specific introduction of unnatural amino acids and chemical or biophysical tags into large proteins. The disadvantage of this method is the necessity of Cys residue or a homologue at the ligation site. The occurrence of this amino acid in globular proteins is very low and the insertion of additional Cys residues can alter the protein structure and function by the formation of disulfide bridges. Through synthetic and semisynthetic

approaches, many unnatural amino acids bearing solvatochromic fluorophores have been incorporated into peptides and proteins (Dufau and Mazarguil 2000; Cohen, McAnaney et al. 2002; Vazquez, Rothman et al. 2004; Vazquez, Blanco et al. 2005; Summerer, Chen et al. 2006; Loving and Imperiali 2008).

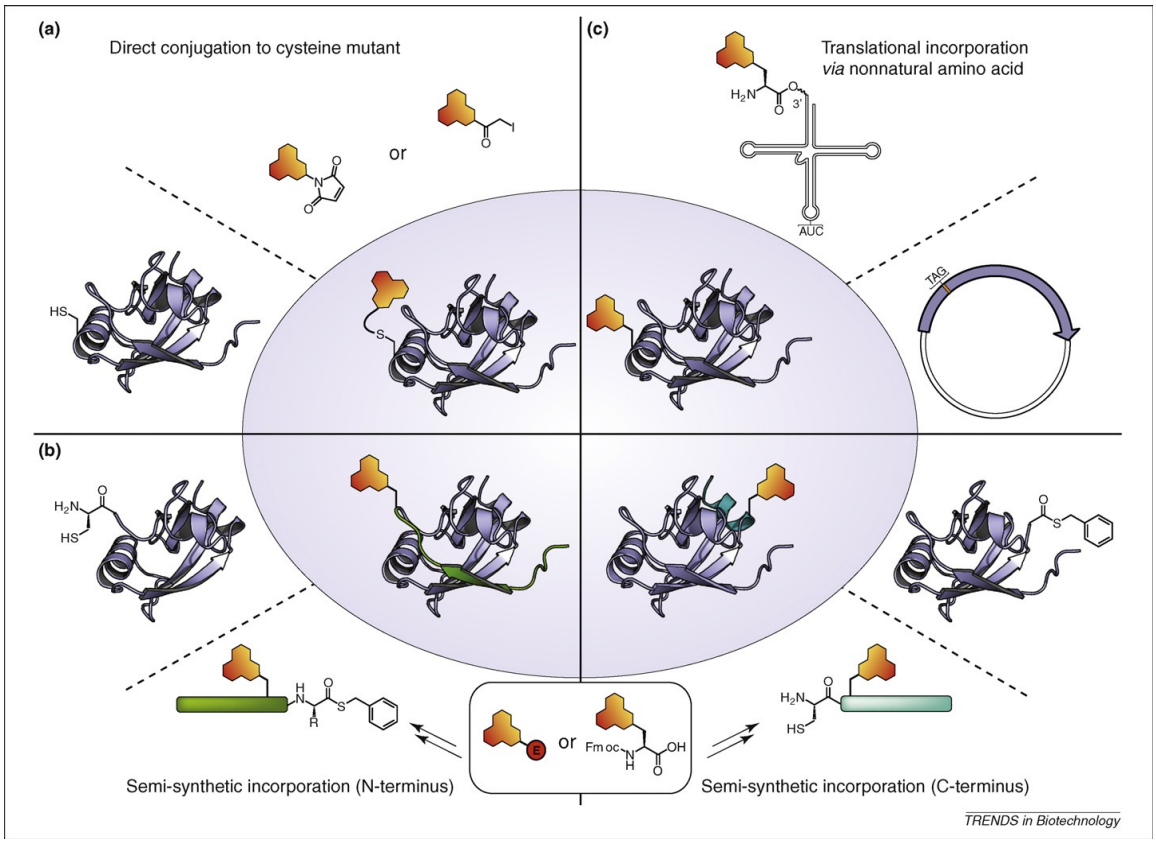


Figure 1.26. Methods for site-selective and site-specific incorporation of fluorophores into proteins (Loving, Sainlos et al. 2010). (a) Direct labeling. (b) Incorporation of nonnatural amino acids via synthesis and semisynthesis (expressed protein ligation). (c) Translational incorporation of nonnatural amino acids.

1.7.5. Indirect labeling by click chemistry reactions

Click chemistry is the easiest way to functionalize peptides and proteins. It requires the incorporation of a non-natural amino acid containing a group suitable for click chemistry (azides or alkynes) or grafting of such groups to protein by another way. Two types of click chemistry reactions are used in protein labeling: azide-alkyne [3 + 2] cycloaddition reaction and Staudinger ligation.

[3 + 2] cycloaddition reaction with acetylene or azide derivatives is characterized by extremely high selectivity. This reaction can be carried out at room temperature under aqueous conditions by the addition of catalytic amounts of Cu(I) salts to the reaction mixture (Agard, Baskin et al. 2006; Angell and Burgess 2007). Moreover, Cu-free methods and reagents were developed recently (Baskin, Prescher et al. 2007)**.** Staudinger ligation exploits the smooth reaction between an azide and a phosphane to form a phospha-aza-ylide. This ylide can be trapped by an acyl group with

formation of a stable amide bond. Since the azide moiety is absent in almost all naturally occurring compounds, click reactions become widely used in chemical biology for highly specific labeling (Kohn and Breinbauer 2004). Moreover, the labeling based on click chemistry is more suitable for application of environment-sensitive dyes, since the obtained linker group is relatively small, so that the fluorophore could be placed at any position close to the site of interest in a given protein. Moreover, recent studies show possibilities to use the click-approach to in situ fluorescence labeling inside living cells (Baskin, Prescher et al. 2007; Neef and Schultz 2009), which opens new possibilities for applications of environment sensitive dyes for intracellular monitoring of protein activity.

1.7.6. Translational incorporation of nonnatural amino acids

With the recent developments in synthetic biology, it has become possible to engineer bacteria and yeast in such a way that they can incorporate nonnatural amino acids with selective reactivities into proteins. With this technique, the possibilities for site-specific conjugation to functionally active proteins can be substantially increased. Several methods currently exist for the translational incorporation of nonnatural amino acids (de Graaf, Kooijman et al. 2009).

One of the more successful approaches utilizes the amber stop codon (TAG) to encode a new amino acid (Ellman, Mendel et al. 1991). The method involved the semisynthesis of an artificial amino acyl tRNA (AA-tRNA) molecule designed to recognize the amber codon through base pairing of the anticodon loop. The artificially acylated tRNA is introduced into a cell extract derived from *Escherichia coli* (or in other cases from yeast or rabbit reticulocyte lysates), which is rich in the molecular machinery necessary for protein synthesis. The gene for the protein of interest, which has been mutated to incorporate the amber codon site-specifically, is then translated with the nonnatural amino acid integrated into the protein at the desired position (Figure 1.26c). The general disadvantage of these methods is their difficulty of realization. Nevertheless, intensive research in this field is in progress due to its importance in cellular investigations.

1.8. Fluorescent nucleoside analogs

In nucleic acids, the intrinsic fluorescence of the common nucleotides is too low to be useful. Accurate measurements reveal exceedingly low fluorescence quantum yields for the natural nucleobases (0.03 % or less (Daniels and Hauswirt 1971; Serrano-Andres and Merchan 2009)) associated with excited state lifetimes in the order of picoseconds. The only exceptions to this rule are several naturally modified nucleosides, including 4-thiouridine, 7-methylguanosine and the Wye derivatives. With such a limited number of naturally occurring fluorescent base analogs, development of new extrinsic fluorescent nucleoside analogues that are isomorphic to their native counterparts with favorable photophysical properties are highly demanded (Wilson and Kool 2006; Tor 2007; Sinkeldam, Greco et al. 2010; Wilhelmsson 2010). Several factors must be considered to maintain favorable photophysical properties, while avoiding significant nucleic acid structural disturbance (Tor 2007). First, it is important to design analogs that are structurally similar to native nucleobases to allow hybridization and other recognition processes to occur. Second, the analogs should have red shifted absorption spectra relative to the native nucleosides. Minimizing the absorption overlap permits selective excitation. Third, the analogs must have sufficient quantum efficiency in order to be detectable. A long emission wavelength is also preferable, so the fluorescence can be in the visible range. Lastly, the emission of the fluorescent analogs should be sensitive to changes in its microenvironment. While many fluorescently modified nucleosides exist, very few actually follow the ideal design criteria.

Schematically, the fluorescent nucleoside analogs can be divided in five classes:

a) Fluorescent nucleoside analogues derived from natural bases;

b) Pteridines;

c) Nucleoside analogues with extended nucleobases;

d) Nucleic bases combined with a fluorophore;

e) Chromophoric base analogs.

1.8.1. Fluorescent nucleoside analogues derived from natural bases

This class of compounds includes derivatives of purine and pyrimidine. The modified fluorescent base closely resembles its natural analogue, with the same shape, size and ability to form hydrogen bonds. The advantages for these analogs are their close similarity to the natural bases and minimally perturbing nature.

2-Aminopurine (2-AP) is the most widely used fluorescent nucleoside, being a derivative of adenosine with improved photophysical properties (Figure 1.27). Its ability to form WC-like base pairs with dT/U, high quantum yield (Φ = 0.68 in water), minimal sensitivity to pH changes, and sensitivity to environmental polarity, make 2-AP an ideal emissive nucleoside analogue. However, when introduced in a single or double-stranded DNA, 2-AP's emission is significantly quenched, thereby requiring the use of large quantities of 2-AP-labeled oligonucleotides. (Ward, Reich et al. 1969)

The 8-vinyl-deoxyadenosine (8-vda) (Figure 1.27), a new analogue of dA was synthesized (Gaied, Glasser et al. 2005). Substitution of the hydrogen at the 8 position of adenine with a vinyl moiety results in remarkable photophysical properties (Gaied, Glasser et al. 2005; Kenfack, Burger et al. 2006). The emission of 8-vdA was shown to be responsive to changes in temperature and solvent, while insensitive to pH changes (between 5 and 10). This residue was found to be closer to natural nucleosides and exhibit higher quantum yield in oligonucleotides than 2-AP. Nevertheless, its emission is still too much quenched in oligonucleotides to enable working with low concentrations of labeled oligonucleotides (Gaied, Glasser et al. 2005).

Tor and co-workers (Tor, Del Valle et al. 2007) developed fused thiophene derivatives 1, 2 (Figure 1.27). The nucleoside analogue **1** shows spectroscopic characteristics similar to 2-AP (Tor, Del Valle et al. 2007) while the isomeric [3,4] analogue **2** show strong red shift in the emission spectrum (412 nm) (Srivatsan, Weizman et al. 2008). The most important property of these fused analogues is the sensitivity of their photophysical parameters to polarity.

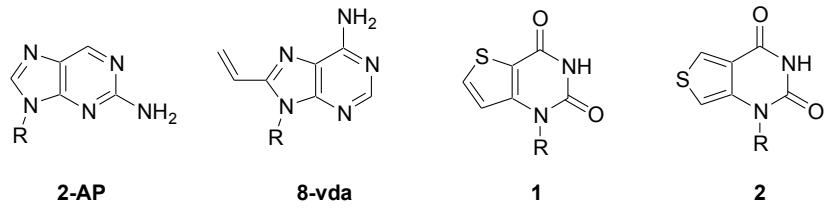


Figure 1.27. Selected examples of fluorescent probes derived from natural bases, where R = ribose or 2′-deoxyribose.

1.8.2. Pteridines

Pteridines are naturally occurring, heterocyclic structures similar to purines. DMAP and 3-MI are two examples of dA and dG nucleoside analogues, respectively (Figure 1.28). Their emission wavelengths are relatively red shifted (~430 nm) and their quantum efficiency is very high (Hawkins, Pfleiderer et al. 2001). However, the incorporation of these probes in a single or double strand causes a strong fluorescence quenching (up to 96%). These probes are interesting alternatives to 2-AP. Since these probes are highly quenched in oligonucleotides, they can be used to monitor events associated with a distortion of the labeled DNA, which results in a strong increase in fluorescence (Parsons and Hermann 2007).

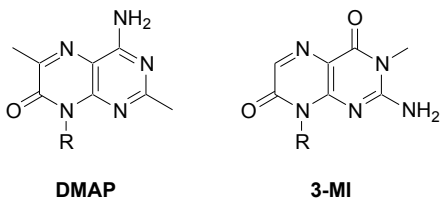


Figure 1.28. Selected examples of pteridines, where R = 2′-deoxyribose or ribose.

1.8.3. Nucleoside analogues with extended nucleobases

Besides the functional changes of the natural bases, it is also possible to access new fluorescent analogues by increasing the conjugation of purine and pyrimidine moieties. In general, the extension of conjugation enhances the fluorescence properties but the obtained large surface area can structurally perturb the oligonucleotides labeled with these compounds.

Leonard and co-workers first investigated etheno-A (εA) (Secrist, Barrio et al. 1972; Secrist, Barrio et al. 1972) and benzo-A (Scopes, Barrio et al. 1977) (Figure 1.29) in the early 1970s showing that adenine could be cyclized to produce nucleobases with red-shifted absorption bands (Kochetkov, Shibaev et al. 1971). While the fused structure of εA, similar to the naturally occurring fluorescent nucleoside wyosine (Rajbhandary, Chang et al. 1967) masks the hydrogen bonding face, it also improves its photophysical properties (Secrist, Barrio et al. 1972; Rist and Marino 2002).

The naphthalene substituted nucleoside BgQ is a relatively recent addition to this class of fluorescent nucleosides (Figure 1.29) (Godde, Toulme et al. 1998; Godde, Toulme et al. 2000). This analogue displays a red-shifted absorption band (360 nm) and a strong emission band in the visible range (Φ = 0.82). Due to its photophysical properties and large surface area, BgQ was employed for the study of double- and triple-stranded oligonucleotides.

Recently, benzopyridopyrimidine (BPP) derived from 2'-deoxycytidine has been described (Figure 1.29) (Okamoto, Tainaka et al. 2003; Okamoto, Tainaka et al. 2003). It shows maxima absorption and emission at 350, 390 nm, respectively. Due to its low quantum efficiency (0.04), the naphthopyridopyrimidine was synthesized. While showing similar absorption and emission wavelengths as BPP, it displays a substantially higher emission quantum efficiency (Φ = 0.26) (Okamoto, Tainaka et al. 2003).

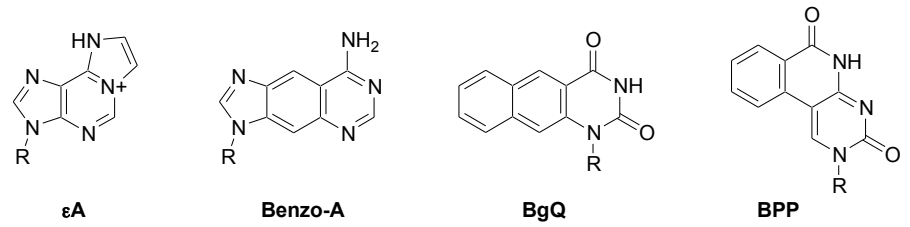
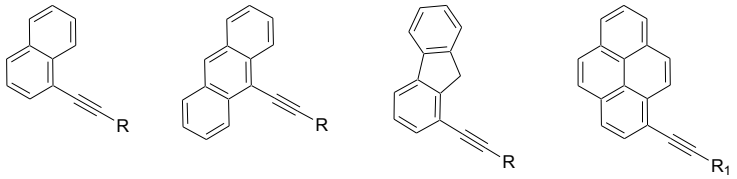


Figure 1.29. Selected examples of improved nucleoside analogues with extended conjugated rings, where R = 2′-deoxyribose or ribose.

1.8.4. Nucleic bases combined with a fluorophore

This class of compounds is characterized by fluorescent moieties that are linked or conjugated to the natural nucleobases, via either flexible or rigid linkers. This approach has been studied with polycyclic hydrocarbons (Figure 1.30). Thus, analogues of dU and dA, respectively, have been functionalized at position 5 and 8, with naphthalene, anthracene (Xiao, Ranasinghe et al. 2007), fluorene (Ryu, Seo et al. 2007) and pyrene (Saito, Miyauchi et al. 2004; Hwang, Seo et al. 2005; Okamoto, Tainaka et al. 2005; Saito, Hanawa et al. 2005) rings.



R = 5-dU; R_1 = 5-dU, 5-dC, 8-dA, 8-dG

Figure 1.30. Selected examples of nucleic bases combined with a fluorophore.

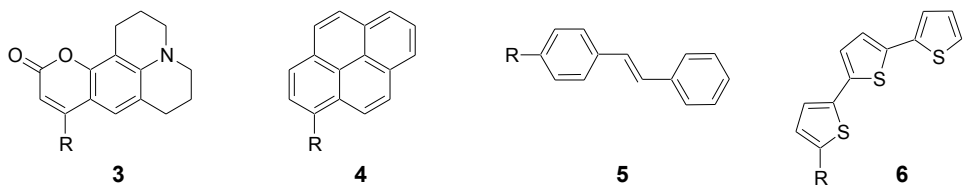
Berlin and co-workers linked pyrene via an ethynyl linkage to the 5-position of dU (Korshun, Prokhorenko et al. 1997). All pyrene nucleosides share similar photophysical properties with red-shifted absorption in comparison to the native nucleobases (342-392 nm) and a weak emission band (Φ = 0.002-0.027) near the visible range (395-474 nm). 5-(1-Ethynylpyrenyl)-dU has been utilized by numerous groups for a variety of applications (Netzel 2007; Wagenknecht 2008). The corresponding 8-substituted purines can be viewed as rather perturbing analogues (Wagenknecht 2008).

1.8.5. Chromophoric base analogues

The next family of compounds was developed by replacing the natural base with an unusual family of chromophoric base analogues, typically polycyclic aromatic hydrocarbons (PAH) (Figure 1.31). The structure of these analogues does not allow the formation of Watson-Crick bonds but the loss of hydrogen bond pairing may be compensated by other interactions, such as stacking.

Coleman and co-workers described the synthesis of coumarin **3** containing nucleoside (Figure 1.31) (Coleman and Madaras 1998; Andreatta, Sen et al. 2006; Coleman, Berg et al. 2007) as a fluorescence probe to study the dynamics of DNA. Coumarin is known for its high quantum efficiency and absorption maximum at 400 nm, which allows selective excitation in the presence of natural bases.

Numerous examples have been described in the literature for which the base is a polycyclic aromatic moiety such as pyrene **4**, stilbene **5**, terthiophene **6** (Figure 1.31) (Ren, Chaudhuri et al. 1996; Strassler, Davis et al. 1999).



3 **4** **5** **6**

Figure 1.31. Selected examples of chromophoric base analogues, where R = 2′-deoxyribose.

1.9. Approaches for oligonucleotide labeling

Two approaches have been developed to get fluorescent nucleic acids. The first is to intercalate within the DNA bases, fluorescent molecules such as acridine orange (AO) (Brauns, Murphy et al. 1998; Coleman, Berg et al. 2007), ethidium bromide and thiazole orange (Tse and Boger 2004). The fluorescence probe in this case is associated non covalently, its location is not defined and will be governed by its chemical nature and the interactions with the DNA bases. In most cases, the use of intercalated probes causes distortion of the DNA (Brauns, Murphy et al. 1998; Coleman, Berg et al. 2007) and does not provide all required information.

The second approach consists of covalent attachment of the fluorescent molecule to DNA:

a) by addition of a chromophore via a linker

In an oligonucleotide sequence, different positions (5' or 3') can be labeled by fluorescent dyes, such as derivatives of fluorescein, rhodamine. Generally, the length of the linker which is often a linear alkyl chain (Jiao and Burgess 2003) is between 6 and 12 bonds. This gives the fluorophore high mobility and creates uncertainty about its position and orientation which is important in some cases where the distance and structure play a major role. Moreover, the linker and attached fluorophore form a group which occupies a large volume that can interact with DNA or other macromolecules, negatively affecting the properties of DNA.

b) by incorporating fluorescent nucleoside analogues (which will be highlighted below).

Ideally, the incorporation of the fluorescent probe in the DNA must fulfill three important conditions: the probes should occupy a single well-defined position; they should not disrupt the structure and functions of nucleic acids and these probes must have appropriate spectroscopic properties (Coleman, Berg et al. 2007). Moreover, these fluorescent probes must also be compatible with the synthesis of oligonucleotides.

1.9.1. Incorporation of fluorescent nucleoside analogues into oligonucleotides

This labeling strategy shows some advantages over the others: fluorescent nucleoside analogues have small size which results in less disturbance of the structure and interactions of DNA with other biomolecules. These nucleoside analogues can be incorporated into DNA in place of a natural nucleotide in a strategic position allowing studies of many biological processes.

Whatever the desired labeling position on the oligonucleotide, it is important that it is specific. There are several approaches to incorporate modified nucleosides into oligonucleotides, such as solid-phase assisted synthesis, enzymatic incorporation, and postsynthetic modification (Goodchild 1990; Verma and Eckstein 1998; Chow, Mahto et al. 2008). The first two approaches are most relevant to the fluorescent nucleoside analogues discussed above. The most popular is the incorporation of modified nucleosides into oligonucleotides using solid-phase assisted synthesis, which enables modification of oligonucleotides at any position. The predominantly used techniques rely on the use of a 3′-phosphoramidite group for the introduction of the phosphate group in both DNA and RNA oligonucleotides (Gait 1984; Caruthers, Barone et al. 1987; Beaucage and Iyer

1992). The cycle used in the standard solid-phase phosphoramidite chemistry is shown in Figure 1.32. There are several key steps:

1) acidic removal of the 5′-DMTr (5′- dimethoxytrityl) protecting group of the solid-supported nucleoside

2) mild acidic activation of the phosphoramidite (e.g., using 1*H*-tetrazole)

3) coupling of the available 5′-hydroxyl to the activated phosphoramidite

4) oxidation of the trivalent phosphorus to the pentavalent phosphorus

5) capping of unreacted hydroxyl groups.

This cycle continues until termination and removal of the full-length oligonucleotide from the solid support, typically under basic/nucleophilic conditions (e.g., ammonium hydroxide) (Sinkeldam, Greco et al. 2010).

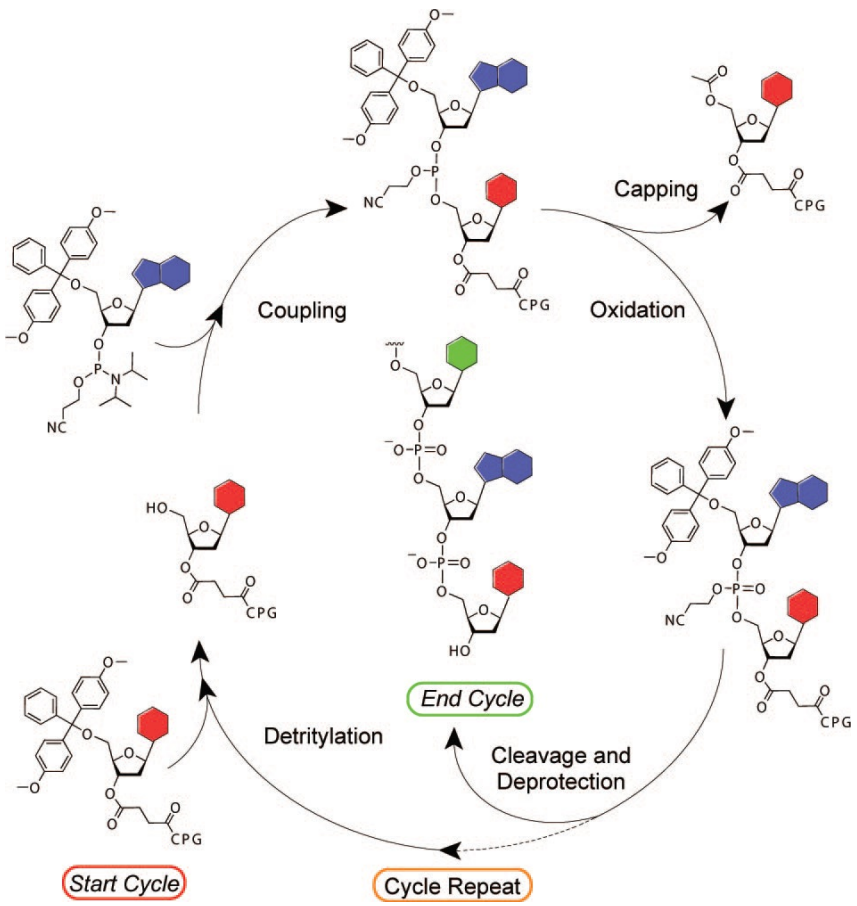


Figure 1.32. Common cycle for solid-phase-assisted phosphoramidite oligonucleotide synthesis (Sinkeldam, Greco et al. 2010).

As an alternative method, enzymatic incorporation of fluorescent nucleotides can, in some cases, be used. Enzymatic incorporation has to be evaluated on a case by case basis because DNA and RNA polymerases of different origin display diverse tolerance levels to unnatural nucleotides (Sandin, Stengel et al. 2009; Srivatsan and Tor 2009). Conversion of the modified nucleosides into the corresponding triphosphate is necessary for implementation of this approach (Da Costa, Fedor et al. 2007; Srivatsan and Tor 2007). The synthesis of many triphosphates does not require any

protection of the exocyclic groups and is performed in a one-pot/two-step procedure following established methods (Vaghefi 2005). Additionally, enzymatic incorporation of fluorescent nucleotides has been achieved through unnatural base pairs (Benner 2004; Kawai, Kimoto et al. 2005; Loakes, Gallego et al. 2009), use of terminal transferases (Barone, Chen et al. 2001), or a reactive amino group on the recently incorporated nucleoside able to undergo conjugation with an appropriately selected tag (Cox and Singer 2004). The main disadvantage of this strategy is that the fluorescent nucleotide analogue must be recognized by the enzyme and incorporated into the nucleic acid during the elongation of a defined position in order to avoid uncontrolled multiple labeling.

1.10. Human immunodeficiency virus type 1 (HIV-1)

HIV-1 infection targets the immune system with progressive impairment of cell mediated immune response resulting in Acquired ImmunoDeficiency Syndrome (AIDS) characterized by severe immune deficiency and the development of opportunistic infections. The first case of AIDS was registered in USA in 1981. This disease is an important public health problem as it has caused death of around 25 million people since 1981. In 1983, the causative agent of AIDS, Human Immunodeficiency Virus (HIV, type 1) was discovered by Luc Montagnier and his team in France. People living with HIV/AIDS around the world counts up to 33.3 millions with 22.5 millions in sub-saharan Africa (UNAIDS, 2010).

HIV is a lentivirus, a member of the retrovirus family. It infects lymphocytes CD4+ and other cells of the immune system, leading to their death. When the number of immune system cells declines below a critical level, the cell-mediated immunity is lost, and the body becomes progressively more susceptible to opportunistic infections. HIV is transmitted by sexual contact, through blood and from infected mother to child.

Two types of HIV have been identified until now: HIV-1 (Barresinoussi, Chermann et al. 1983) and HIV-2 (Clavel, Guyader et al. 1986). HIV-1 is the most prevalent in the worldwide pandemic. HIV-2 is present mainly in West Africa. Compared to HIV-1, HIV-2 is less transmissible and less pathogenic, with lower viral loads and giving an asymptomatic and a slower progression to AIDS.

Due to this variability, HIV-1 variants are classified into three major phylogenetic groups: group M (Main, represents 90% of infections), group O (Outlier, restricted to west-central Africa), and group N (Non-M, Non-O. Rare group discovered in 1998 in Cameroon). Group M, which is responsible for the majority of infections in the worldwide HIV-1 epidemic, can be further subdivided into ten recognized phylogenetic subtypes, or classes (A to K, excluding E, which is actually a CRF - circulating recombinant form).

1.10.1. Genetic organization

a. Viral RNA

The genome of HIV-1 consists of two copies of an homologous single-stranded RNA of 9.2 kb linked noncovalently near their 5' end. The RNA is capped and polyadenylated. It is produced after the integration of the provirus into the genome of infected cells. The transcription of the viral RNA starts at the 5' end of the R sequence located in the 5'LTR and terminates at the 3' end of the R sequence located in the 3'LTR (Figure 1.33).

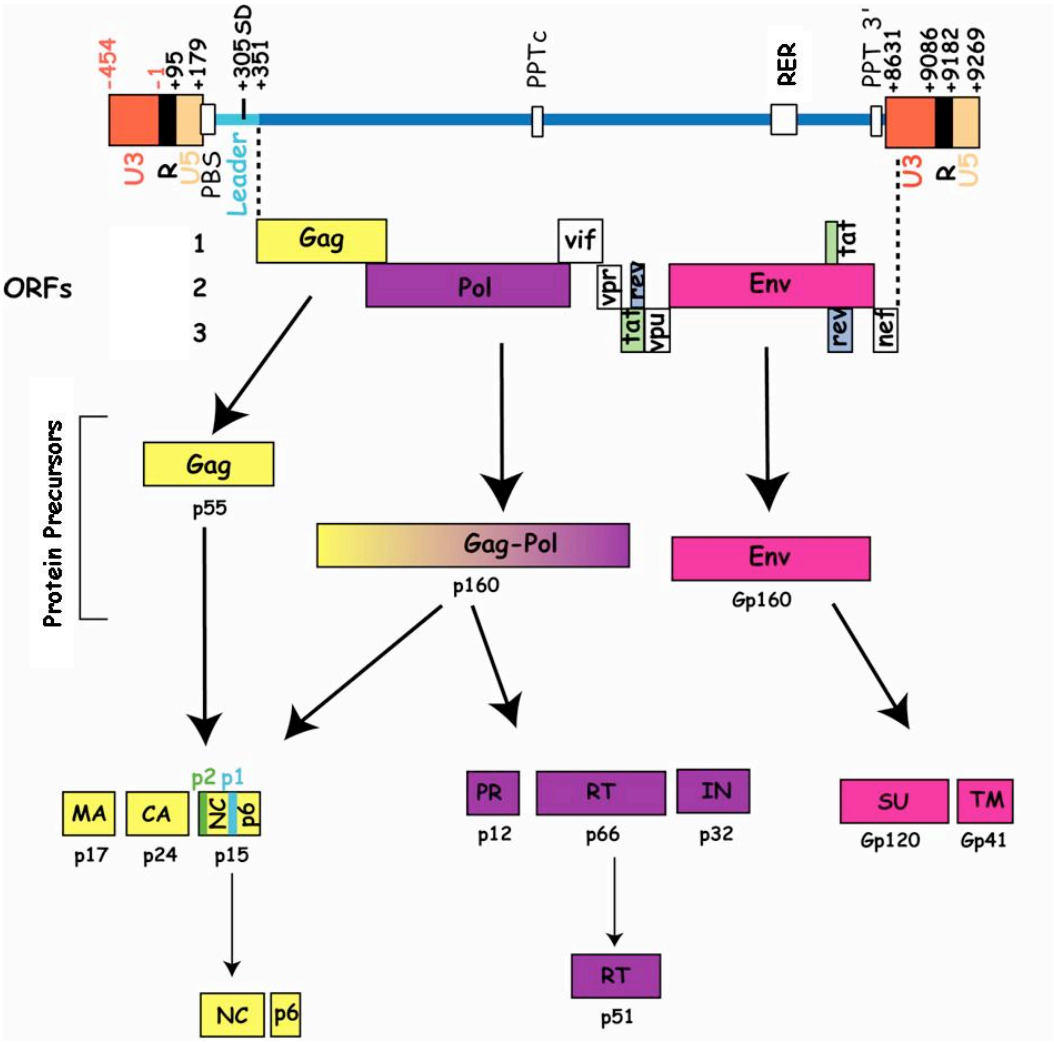


Figure 1.33. Genetic organization of HIV-1. Figure adapted from (Goldschmidt 2004).

5' region of the viral RNA (Figure 1.34)

Many stem-loops are found in the 5' region of the RNA genome (Berkhout 1996; Watts, Dang et al. 2009) and are involved in different steps of the replicative cycle.

Region R consists of two stem-loop sequences with different functions:

- stem-loop TAR (Trans-Activating Response element) composed of the 55-60 first nucleotides of the viral RNA. The TAR hairpin binds Tat and several cellular proteins (Gatignol, Kumar et al. 1989) and allows the transcriptional regulation of the viral genome (Karn 1999). TAR also plays a role in the initial steps of reverse transcription (Harrich, Ulich et al. 1996) and the assembly of the virus (Clever, Eckstein et al. 1999);

- stem-loop poly-A containsthe polyadenylation signal (AAUAAA) in its loop. At the 3' end of the viral RNA, this motif allows the cleavage of the viral RNA and the addition of the poly-A tail (Klasens, Thiesen et al. 1999).

Primer Binding Site (PBS) is a sequence of 18-nucleotides complementary to the 3' end of a specific tRNA: $tRNA^{Lys3}$, which is important to initiate the reverse transcription (Prats, Sarih et al. 1988).

The leader region is located between the PBS sequence and the starting codon. It contains four stem loops (SL1 – SL4) crucial for genome recognition and RNA dimerization (Figure 1.34) (Skripkin, Paillart et al. 1994; Paillart, Shehu-Xhilaga et al. 2004):

- SL1, also called DIS (Dimerization Initiation Site) belongs to the dimerization sequence of genomic RNA (Skripkin, Paillart et al. 1994).

- SL2, contains the splicing donor site (SD) (Ashe, Pearson et al. 1997) that is used to generate more than 30 different subgenomic spliced mRNAs (Purcell and Martin 1993; Oreilly, McNally et al. 1995).

- SL3 (or ψ) is an encapsidation signal (Hayashi, Shioda et al. 1992) that is critical for virus assembly (Clever, Eckstein et al. 1999; Helga-Maria, Hammarskjold et al. 1999).

- SL4 serves as an initiation codon for the Gag gene.

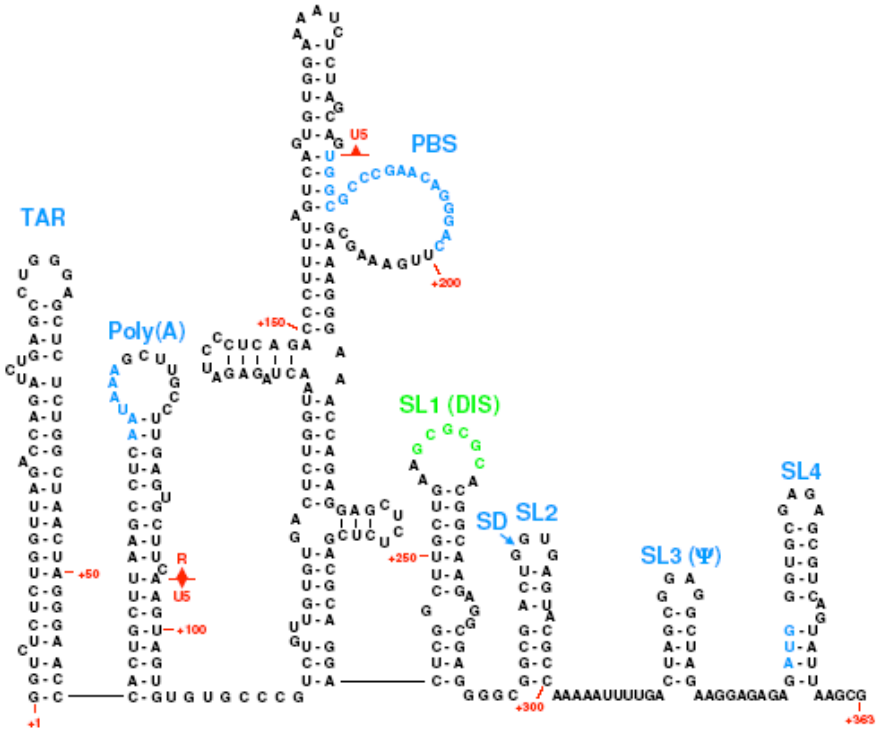


Figure 1.34. Sequence and structure of the 5' region of HIV-1 RNA (Russell, Liang et al. 2004).

Internal elements:

- RRE sequence (Rev Response Element), close to the 3' end of the viral genome (Frankel and Young 1998), contains 234 nucleotides structured in hairpins. Rev protein binds to the RRE sequence in the non- or mono-spliced viral RNA and regulates their export (Malim, Bohnlein et al. 1989; Zemmel, Kelley et al. 1996).

- PPT sequence (Polypurine tract). There are 2 PPT sequences (PPTc and 3' PPT), that are rich in purines and serve as primers for the synthesis of the (+) strand of the proviral DNA due to their resistance to the degradation by the RNAse H activity of HIV RT (Frankel and Young 1998).

3' region of the viral RNA

The U3 region contains signals required for the integration of proviral DNA (Reicin, Kalpana et al. 1995), and a polyadenylation signal in R.

Proviral DNA (Figure 1.33)

The genomic DNA is flanked by two LTRs (Long Terminal Repeats) that are non coding regions important for the integration of the viral DNA and for the transcription of viral RNA. HIV-1 genomic DNA encodes three polyproteins, Gag, Pol and Env that are clived either by cellular or viral protease.

- Gag gene ("Group Specific Antigen") encodes a myristyl precursor of 55 kDa – $Pr55^{Gag}$. After cleavage by the viral protease, $Pr55^{Gag}$ gives rise to the matrix protein p17 (MA), capsid protein p24 (CA), nucleocapsid protein p7 (NCp7) and p6 protein. Peptides p1 and p2 are also generated from Gag polyprotein but their function is not clear.

- Pol gene ("polymerase") encodes a large protein precursor $Pr160^{Gag\text{-}Pol}$ that generates three enzymatic proteins: protease (PRp12) that is cleaved by self-catalysis, reverse transcriptase (RTp66/p51), and integrase (INp32). These two proteins are cleaved by the viral protease. Pol precursor is obtained by a -1 ribosomal frameshift mechanism that occurs at a frequency of 5 to 10% of Gag translation events (Wapling 2007).

- Env gene ("envelope") encodes a polyprotein precursor $Pr160^{Env}$ whose maturation occurs in the endoplasmic reticulum. The Env polyprotein is cleaved by the cellular protease to produce a transmembrane protein (TMgp41) and a surface protein (SUgp120), two membrane glycoproteins involved in virus-host cell interactions.

HIV-1 also encodes regulatory (Tat and Rev) and accessory proteins (Nef, Vif, Vpu, Vpr) required for the viral pathology (Frankel and Young 1998).

1.10.2. Structure of the viral particle

The mature virion is a spherical particle of 110-130 nm in diameter (Figure 1.35) (Sierra, Kupfer et al. 2005). The virion contains genomic RNA which is diploid, consisting of two copies of homologous single-stranded RNA. Viral genome is highly condensed in association with the nucleocapsid protein (NC), the viral dependent DNA polymerase, Pol, andthe reverse transcriptase (RT) enclosed in the conical shaped capsid or "core" composed of the p24 Gag capsid (CA) protein. The core is surrounded by a viral shell made up of the matrix protein (MA), which in turn is surrounded by the envelope. This viral envelope originates from the host cell membrane acquired during budding. It contains the external surface viral glycoproteins gp120 (SU) and gp41 transmembrane (TM) proteins (300-500 copies per particle) while the inner surface is bound to the MA. Accessory proteins, such as Nef and Vif are thought to be associated with the core. The virus is build up from a combination of three polyprotein precursors (Gag, Env, Pol), whose cleavage during the maturation step generates the architecture and infectivity of the virus particle matures.

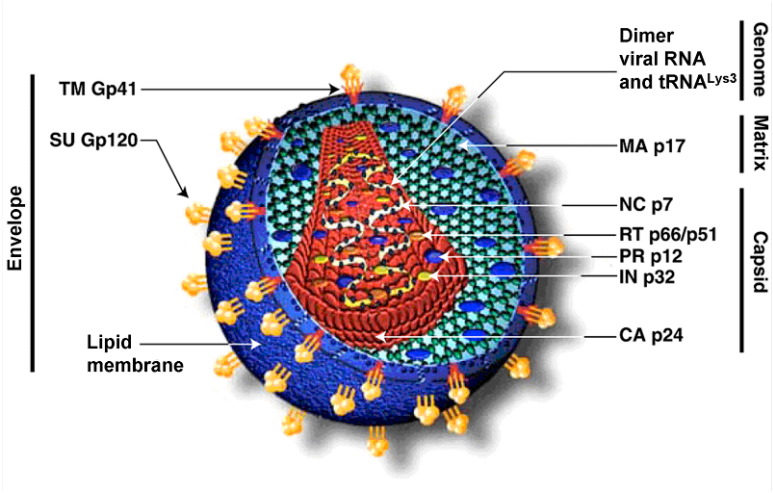


Figure 1.35. Schematic structure of the mature HIV-1 virion. Figure adapted from (Godet 2010).

1.10.3. HIV proteins

Depending on their function, the viral proteins could be classified in four groups:

- Envelope proteins - Transmembrane (TM), Surface (SU) proteins
- Structural proteins - Matrix (MA), Capsid (CA), Nucleocapsid (NC)
- Enzymatic proteins - Protease (PR), Reverse transcriptase (RT), Integrase (IN)
- Regulatory/accessory proteins - Vpu, Vif, Vpr, Nef, Tat, Rev

Envelope proteins

The HIV-1 envelope glycoproteins are the product of the Env gene. After a cleavage mediated by the cellular protease, the Env precursor (gp160) allows the production of the surface envelope glycoprotein (SUgp120) and the transmembrane glycoprotein (TMgp41). The envelope glycoproteins TMgp41 form a trimer in which each monomer is linked non-covalently to a glycoprotein envelope SUgp120 (Roux and Taylor 2007).

Surface protein (SUgp120) is a highly glycosylated 515 amino acids protein located at the surface of the viral particle by being non-covalently attached to TMgp41. Gp120 contains five variable regions (V1 to V5) of which four loops are exposed. These variable regions and notably the V3 region are critical. In fact, after a specific recognition between gp120 and the CD4 receptor at the surface of the target cell, a conformational change of gp120 allows the exposition of the V3 loop that could interact with co-receptors CCR5 or CXCR-4 (Huang, Tang et al. 2005).

Transmembrane protein (TMgp41) is a 345–amino acid glycosylated protein involved in membrane fusion during viral entry into the cell. In the virus, TMgp41 forms two antiparallel helices that combine to form a trimer. The protein TMpg41 contains three domains: an intracellular (endodomain), a transmembrane (TM) and an extracellular (ectodomain) domain. The ectodomain

is a symmetric trimer where each monomer consists of two antiparallel helices connected by a large loop (Figure 1.36). The N-terminus of the ectodomain is called the fusion domain and is important for the fusion between the virion membrane and the plasma membrane of the host cell (Prabakaran, Dimitrov et al. 2007). This step takes place after the interaction between gp120/CD4 and the coreceptor.

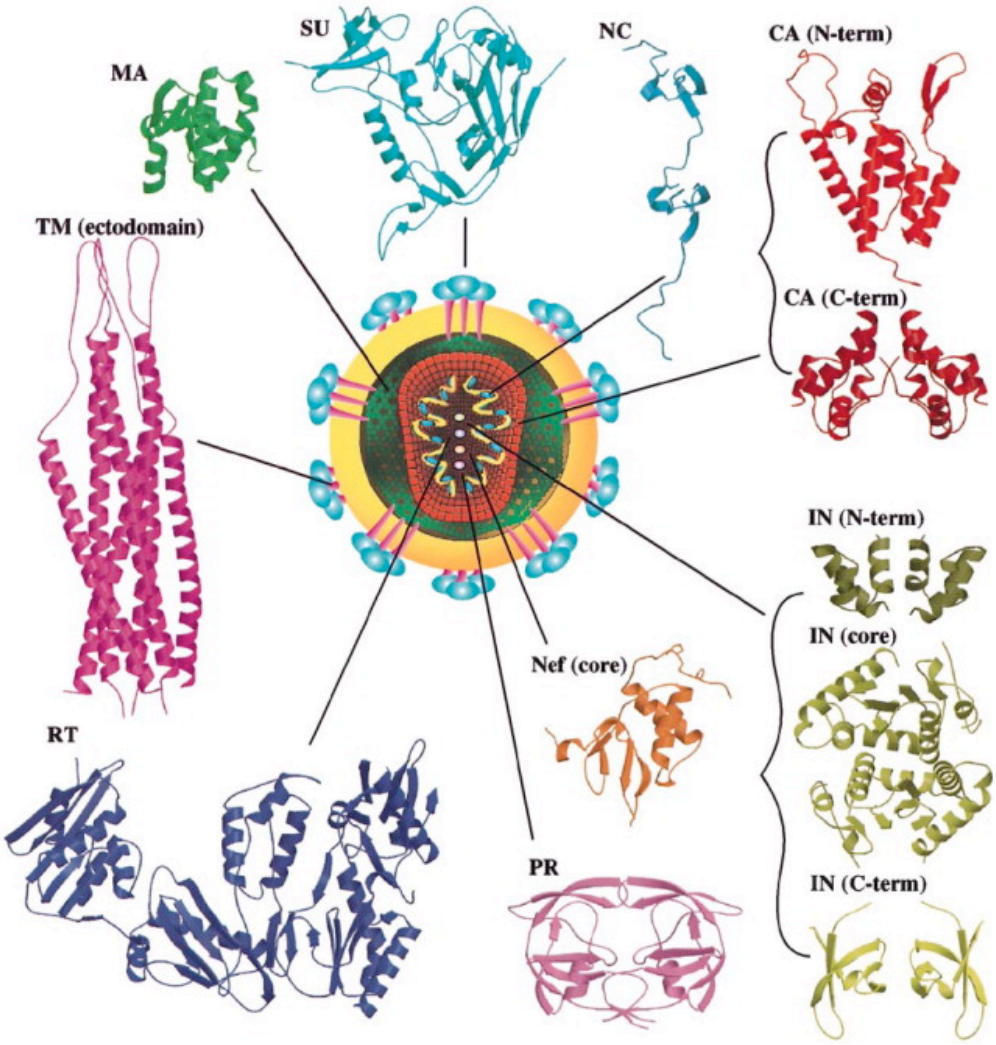


Figure 1.36. Structures of the proteins (or protein domains) of HIV-1 (Turner and Summers 1999). The TM ectodomain shown is that determined for the closely related simian immunodeficiency virus (SIV).

Structural proteins

The Gag gene encodes a polyprotein (55kDa) that self-assembles at the membrane cell to form the immature virion. During virion maturation, Gag is cleaved by protease into three structural proteins (matrix, capsid and nucleocapsid) and other products of lower molecular weight: p1 (SP2), p2 (SP1) and p6.

Matrix (MA) is the N-terminal component of the Gag polyprotein and is obtained after cleavage of the $Pr55^{Gag}$ (Figure 1.36) (Bukrinskaya 2007; Hearps and Jans 2007). MA is important

for targeting the Gag and Gag-Pol precursors to the plasma membrane prior to viral assembly. In the mature viral particle, the 132 amino acids MA protein lines the inner surface of the virion membrane forming an interface between the envelop and the capsid (Figure 1.35). Two distinct features of MA are involved in membrane targeting: an N-terminal myristate group and basic residues located within the first 50 amino acids. The trimeric form of MA is presumed to be biologically relevant since mutation of residues involved in trimerization (residues 42–77) abolishes viral assembly (Fiorentini, Marini et al. 2006) and since basic residues which important for membrane localization are arranged on the putative membrane-binding surface of the trimer. Moreover, the monomeric form of MA is involved in the uncoating step (Wu, Alexandratos et al. 2004), and the transport of the reverse transcription complex in the cytoplasm through its interaction with actin.

Capsid (CA) is the second component of the Gag polyprotein. The CA protein forms the viral capsid, which has an icosahedral fullerene form in the mature virus. The capsid is constituted by the assembly about 2000 molecules of the CA, also called p24 protein (232 amino acids), which is produced from the maturation of the $Pr55^{Gag}$. The complete structure of CA was determined by electron microscopy and NMR (Ganser-Pornillos, Cheng et al. 2007; Ganser-Pornillos, Yeager et al. 2008; Byeon, Meng et al. 2009; Pornillos, Ganser-Pornillos et al. 2009). CA is composed of two independent domains connected by a flexible linker, namely the N-terminal domain (NTD, aa 1–151) that consists of 7 helices and the C-terminal domain (CTD, aa 152–231) which has four α-helixes in its structure (Figure 1.37). The NTD is involved in condensation and morphogenesis of the capsid. It plays a major role in infectivity, by participating in viral uncoating through its association with a putative cellular chaperone, cyclophilin A (CypA) (Gamble, Vajdos et al. 1996; Gamble, Yoo et al. 1997). It is also important in the morphogenesis of the capsid. The CTD is primarily involved in assembly, but also in CA dimerization and Gag oligomerization. Thus, CA plays a fundamental role in structuring the viral particle.

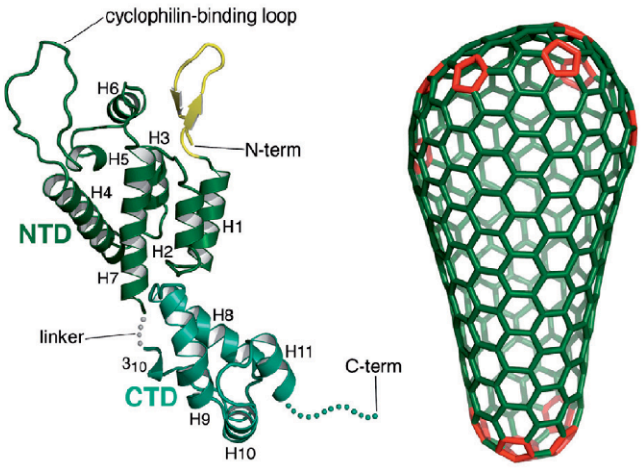


Figure 1.37. Proposed structure of the capsid protein (Ganser-Pornillos, Yeager et al. 2008; Ganser-Pornillos, Yeager et al. 2008) and the viral capsid. The structure of CA is predicted from the structures of the N-terminal domain (NTD) and the C-terminal one (CTD) determined separately (Berthet-Colominas, Monaco et al. 1999). α-helixes are marked as H1 - H11. The right figure shows

the model for the conical capsid, with CA hexamers (green) and pentameric declinations (red) (Ganser-Pornillos, Yeager et al. 2008).

Nucleocapsid (NC) is the third component of the Gag polyprotein. NC is a small basic protein of 55 amino acids characterized by two highly conserved zinc finger motifs, connected by a basic sequence. Functions and structure of NC will be described below.

p6 is a small 6-7kDa protein consisting of 51 amino acids, present in the mature virion. It has two important functional regions whose sequences are conserved, called Region-L (L for "late"), since these motifs play a critical role in the late steps of the viral cycle by recruiting cellular proteins necessary for virus budding (Freed 2002). A mutation within these motifs significantly reduced the production of viral particles (Huang, Chopra et al. 1998). On the other hand, p6 is able to bind to Vpr (Viral protein R), this interaction allowing encapsidation of Vpr into viral particles (Paxton, Connor et al. 1993; Kondo, Mammano et al. 1995).

SP1 and **SP2** also known as p2 and p1, respectively are peptides formed after cleavage of the Gag precursor. They are required during the assembly of the Gag and Gag-Pol polyproteins, and contribute to the recognition of the encapsidation signal of the viral genomic RNA.

Enzymatic proteins

Enzymatic proteins are products of the pol gene. After self catalysis, the viral protease (that belongs to the Gag-Pol precursor) cleaves the polyprotein precursor Gag-Pol and allows the production of protease, reverse transcriptase and integrase.

Protease (PR) is a 99 amino acid homodimer (Figure 1.36), generated by the cleavage of $Pr160^{Gag\text{-}Pol}$ through a self-catalysis mechanism (Pettit, Simsic et al. 1991). PR is a key enzyme essential for the maturation of the virus. Its inhibition induces the production of non-infectious virus particles (Kaplan, Manchester et al. 1994). Its active site is at the interface between the two subunits. PR cleaves at several polyprotein sites to produce the final MA, CA, NC, and p6, p2, p1 proteins from Gag and PR, RT, and IN proteins from Pol. The cleavage of HIV polyproteins by PR induces conformational rearrangements within the particle, which are needed to produce the mature infectious viruses. Some of these "maturation" events may occur simultaneously with assembly and budding. Protease is a prime target of anti-retroviral therapies, due to its importance in the virus infectivity.

Reverse Transcriptase (RT) is a product of $Pr160^{Gag\text{-}Pol}$ cleavage. In its mture form is an heterodimer composed of p66 and p51 subunits (Figure 1.36 and 1.38). The p66 subunit is composed of 5 domains: i) the RNase H domain, which corresponds to the C-terminus of p66 and is absent in p51, ii) three domains called - fingers, palm and thumb that form the polymerase catalytic site and iii) t the connexion domain which connects the polymerase catalytic site and the RNase H domain. Although the four domains are structurally similar in both subunits, their relative position differs from one subunit to another, so that p51 has no polymerase activity but plays a structural role. RT is an RNA- and DNA-dependent DNA polymerase that possesses also an endonuclease activity (RNaseH) that is polymerase dependent or independent (Wisniewski, Balakrishnan et al.

2000; Wisniewski, Chen et al. 2002). The RNaseH activity cleaves specifically the RNA in the DNA/RNA duplexes formed during reverse transcription.

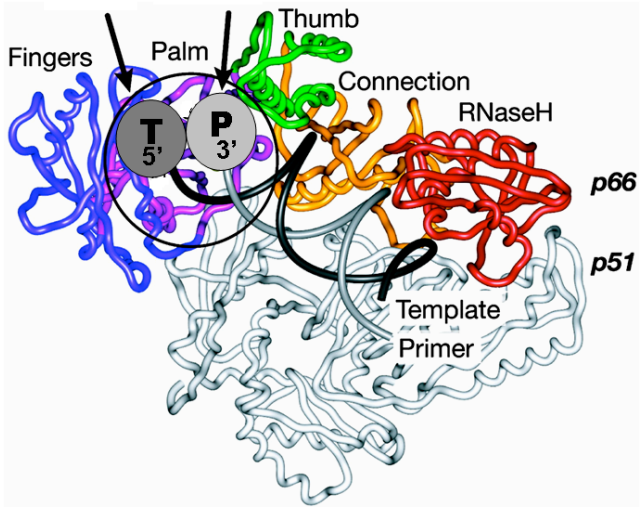


Figure 1.38. Structure of HIV-1 RT with a DNA primer/template (Pata, Stirtan et al. 2004). The DNA primer (light gray) and template (dark gray); the fingers (blue), palm (purple), thumb (green), connection (yellow), and RNaseH (red) subdomains of the p66 subunit of HIV-1 RT; and p51 subunit (white) are in ribbon representation. The region circled includes the polymerase active site.

Integrase (IN) is a 32-kDa (288 amino acids) protein encoded by the end of the Pol gene. IN is composed of three structurally and functionally distinct domains (Figure 1.36) which are required for each step of the integration reaction: the N-terminal domain (1-50) that binds zinc and possibly favors protein oligomerization, the central catalytic core domain (51-212) and a DNA-binding C-terminal domain (Delelis, Carayon et al. 2008). This enzyme is responsible for two consecutive reactions: 3' processing and strand transfer, that lead to the insertion of the proviral DNA within the genome of the host cell (Delelis, Carayon et al. 2008).

Regulatory and accessory proteins

Regulatory (Tat, Rev) and accessory (Nef, Vif, Vpr, Vpu) proteins are also encoded by the genome of HIV-1.

Transactivating regulatory protein (Tat) protein is a 101 amino-acid protein that also exists as an 86 amino acid form in some virus strains (Jeang, Xiao et al. 1999). Functions and structure of Tat will be described below.

Regulatory of Virion Expression (Rev) is an essential accessory protein that mediates the nuclear export of single-spliced and unspliced RNA, thus enabling formation of new infectious virus particles (Arya, Guo et al. 1985; Sodroski, Patarca et al. 1985). This protein (116 amino acids) consists of several motifs that play important roles in its cellular function. The leucine-rich motif (LRM, residues 75–83), located at the C terminus operates as a nuclear export signal (NES). The arginine-rich motif (ARM, residues 34–50) acts as a nuclear-localization signal (NLS) and as a

binding site for the Rev response element (RRE), which is present in all incompletely spliced viral mRNAs (Heaphy, Finch et al. 1991).

Viral protein R (Vpr) is an 96 amino acids accessory protein incorporated into virus particles (Muller, Tessmer et al. 2000) through interaction with the p6 domain of Pr55Gag (Paxton, Connor et al. 1993; Bachand, Yao et al. 1999; Selig, Pages et al. 1999). Its C-terminal part is involved in the G2 cell cycle arrest (Planelles, Jowett et al. 1996; Muthumani, Choo et al. 2005), apoptosis (Jacotot, Ravagnan et al. 2000) and the interaction with NC and nucleic acids (Schuler, Wecker et al. 1999; Bourbigot, Beltz et al. 2005). The N-terminal part of the protein (1-51) is more acidic and is required for virion incorporation, nuclear localization and Vpr oligomerization (Zhao, Wang et al. 1994; Jenkins, Pornillos et al. 2001). This protein is also thought to be involved in the nuclear transport of the pre-integration complexes (PIC) (Le Rouzic and Benichou 2005; Suzuki and Craigie 2007), maybe by disrupting the nuclear envelope (Bukrinsky and Adzhubei 1999). This last function is strengthened by the nuclear envelop (NE) localization of Vpr, which is probably mediated by the nucleoporin (Piller, Ewart et al. 1996; de Noronha, Sherman et al. 2001; Piller, Caly et al. 2003) and by accumulation of Vpr in the nucleus (Coeytaux, Coulaud et al. 2003). As the PIC presents a too important size for nuclear penetration through channels, its uptake could be explained by a local Vpr-dependent architectural membrane disruption or formation of larger channels.

Viral protein U (Vpu) is an 81-residue integral membrane protein with an N-terminal hydrophobic membrane-spanning domain and a C-terminal cytoplasmic tail (Fischer 2003; Strebel 2007). Vpu induces the degradation of CD4 receptors co-localized with the Env proteins in the endoplasmic reticulum (Willey, Maldarelli et al. 1992). The C-terminal domain of Vpu interacts with CD4 (to direct them to a degradation pathway), liberating the Env proteins that could be transported to the cell surface for assembly into viral particles. Vpu also facilitates the budding of viral particles on the surface of infected cells (Gottlinger, Dorfman et al. 1993), particularly in cells expressing tetherin, which specifically inhibits virion release from the cells.

Virion Infectivity Factor (Vif) is a small highly basic, 192 amino acid protein that is essential for in vivo infectivity and pathogenesis, functioning as an auxiliary protein. Vif promotes proteasome degradation of cellular cytidine deaminase APOBEC family (APO-lipoprotein B mRNA-editing Enzyme, Catalytic polypeptide). APOBEC3G is responsible for hypermutation of the viral genome during the reverse-transcription. It is therefore a host defense against viral infection. This defense is made ineffective by the presence of Vif. On the other hand, Vif stimulates reverse transcription by increasing, in synergy with NCp7, the processivity of reverse transcriptase and activating the first strand break (Henriet, Sinck et al. 2007). In addition, Vif interacts with Gag, the viral protease and the genomic RNA. These interactions are likely important for the correct particle assembly. Moreover, a chaperone activity was also reported for Vif (Bernacchi, Henriet et al. 2007; Strebel 2007).

Negative Factor (Nef) is a 206 amino acid protein that increases the infectivity of HIV-1. Nef also promotes the survival of infected cells by down modulating the expression of several surface molecules, such as the: CD4 and CD28 present on CD4+ T cells and the major histocompatibility complex-I (MHC I) and MHC II (Schwartz, Marechal et al. 1996; Swigut, Iafrate et al. 2000) present on antigen presenting cells (APCs) and involved in the recognition of infected cells by the immune system (Schwartz, Marechal et al. 1996; Yang, Nguyen et al. 2002).

1.10.4. HIV-1 life cycle

Knowledge of the different steps of the replicative cycle is essential for the understanding of the physiopathology of HIV and the development of therapeutics. The HIV-1 life cycle (Figure 1.39) could be divided in two main phases (Turner and Summers 1999; D'Souza and Summers 2005):

Pre-integration phase, which begins with the recognition of the host cell, entry of the virus and its decapsidation followed by synthesis of proviral DNA in the reverse transcription step, the nuclear import of the viral genome, and ends with the integration of proviral DNA. These steps are carried out mainly by the viral proteins present in the virion.

Post-integration phase, including the proviral DNA transcription into mRNA, viral protein synthesis, assembly, budding and maturation of the virion which lead to the formation of new infectious viral particles. During this phase, the virus diverts the cellular machinery of transcription, splicing, transport, and translation, which are regulated by viral proteins (Tat, Rev ...).

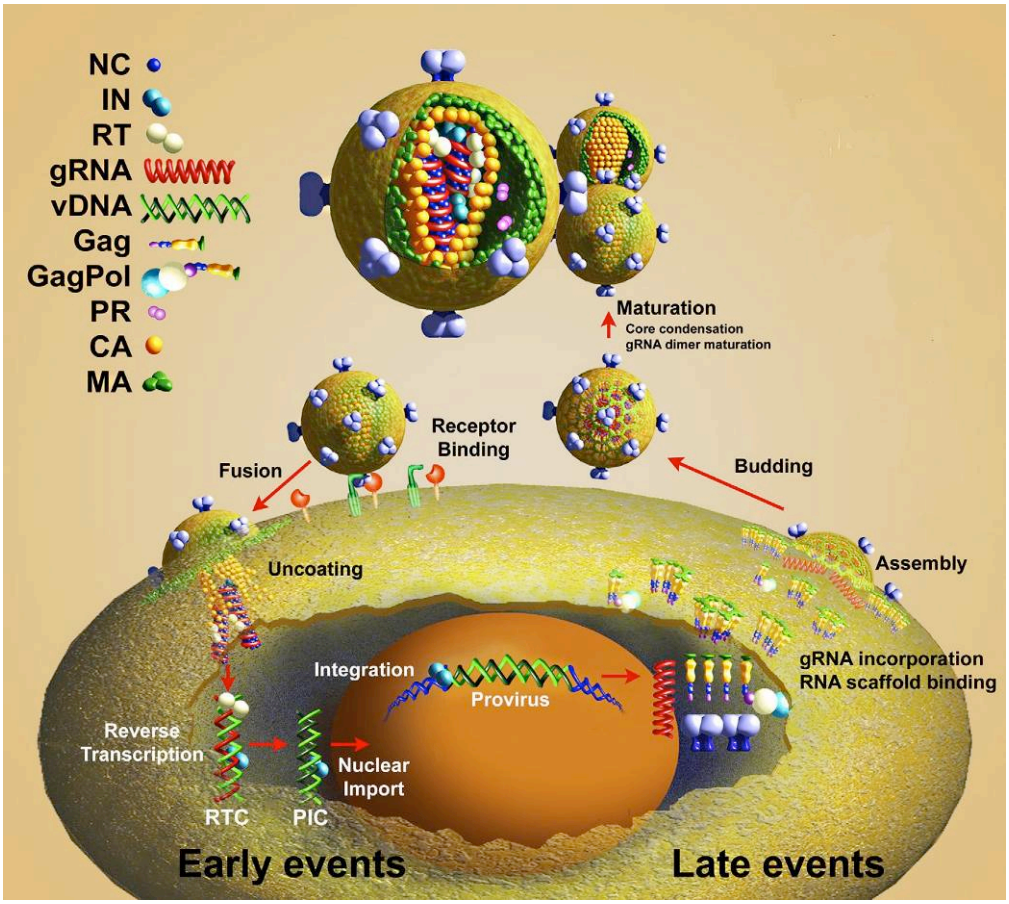


Figure 1.39. HIV-1 viral life cycle. During the early phase of replication (left part), the virus binds to CD4 and chemokine receptors on the cell surface, triggering fusion of the viral and cellular membranes. After cytoplasmic entry, the core begins to uncoat, a RTC forms and the viral RNA genome is reverse-transcribed to proviral DNA. The RTC is then transformed into a PIC, which is transported to the nucleus and integrated into the cellular DNA. The RTC and PIC are shown in the figure without NC or other factors for clarity. Late events (right part) consist of transcription of

messenger RNA from the provirus, translation of viral proteins, assembly, budding and maturation. Figure adapted from (Thomas and Gorelick 2008).

Pre-integration phase

1) Viral binding, fusion and entry

Viral entry starts with the adsorption of the virus at cell membranes. This step is promoted by the envelope glycoproteins at the surface of the virus. The gp120 proteins bind to CD4 and then with chemokine receptors (CCR5 and CXCR4) (Figure 1.39) forming a ternary gp120-CD4-coreceptor complex. Next, conformational changes occur in gp41 which trigger membrane fusion and entry of the viral core into the cytoplasm of the host cell. Virions are partially uncoated in the cytoplasm, releasing the reverse transcription complex (RTC). The RTC is composed mainly by the viral genomic RNA coated with 1500 - 2000 molecules of NCp7. It contains also RT and other proteins such as MA, CA, Vpr or IN (Fassati and Goff 2001). In fact, this schematic view is challenged by recent observations showing that the core of HIV-1 may remain intact in the cytoplasm of infected cells until the nuclear import and uncoating occur at the nuclear pore (McDonald, Vodicka et al. 2002; Arhel, Munier et al. 2006; Arhel, Souquere-Besse et al. 2007; Warrilow and Harrich 2007).

2) Reverse transcription

To enable the integration of the viral genome into the genome of infected cells, genomic RNA must be converted into a double-stranded DNA, which is realized during the reverse transcription step. It is a complex process catalyzed by the reverse transcriptase (RT) and facilitated by NCp7.

Transcription starts after the annealing of $tRNA^{Lys3}$ with the Primer Binding Site sequence (PBS) of viral RNA, generating a double-stranded sequence RNA/RNA (18 nucleotides) recognized by the RT (Figure 1.40, step 1). Minus-strand strong-stop DNA (ss-DNA) synthesis starts at the 3' end of $tRNA^{Lys3}$ and continues to the 5' end of the viral RNA, while the polymerase-dependent RNase H concomitantly degrades the RNA template (step 2). The next step consist in the transfer of (-)ssDNA to the 3' end of the viral genomic RNA (step 3). This transfer is mediated through the repeated R sequences at both sides of viral RNA. The TAR sequence plays a key role in this transfer. This step is facilitated by NCp7. Reverse transcription is resumed to generate the (-)DNA (step 4). During the synthesis of this strand, the RNase H activity of RT degrades the RNA template with the exception of two short PPT sequences (Poly Purine Tract) which are resistant. Plus-strand DNA synthesis is primed by the PPT sequences, and continues to the end of the minus-strand template (step 5). The $tRNA^{Lys3}$ primer is then degraded by the RNase H activity of RT. Hybridization of the (+)PBS and (-)PBS sequences constitutes the second strand transfer (step 6) which is stimulated by NCp7. This step allows the completion of the proviral DNA (steps 7 and 8).

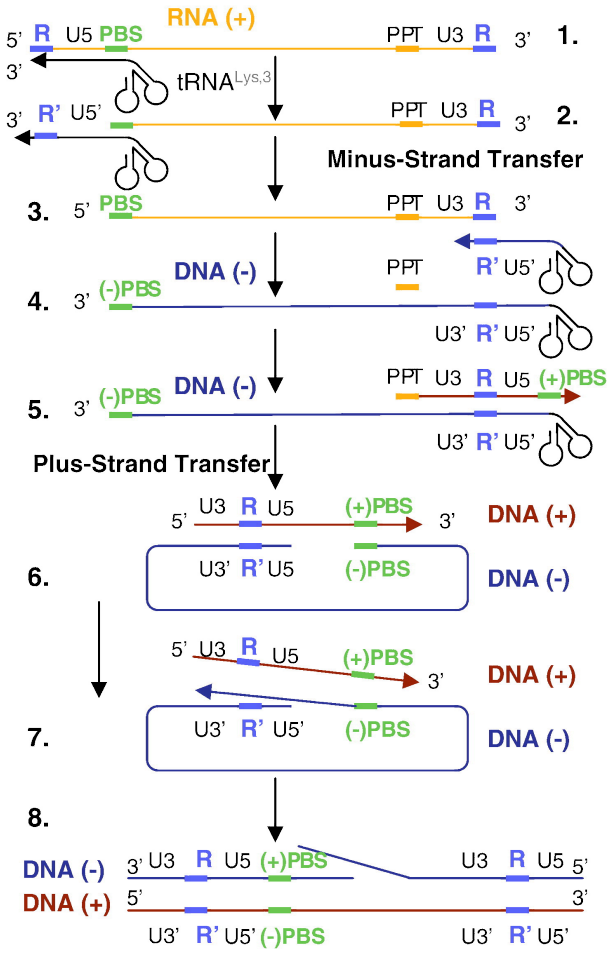


Figure 1.40. Scheme of the main steps of reverse transcription (Bourbigot, Ramalanjaona et al. 2008).

3) Transport to the nucleus

The reverse transcription product is transported to the nucleus as a component of the preintegration complex (PIC) composed of the newly synthesized proviral DNA, different viral proteins, including IN, NC, MA, Vpr and cellular proteins. The molecular mechanism by which PICs cross the nuclear envelope is still poorly understood. It probably involves the α and β importins and Ran-GTP proteins and requires the nuclear localization signals of MA and Vpr (Sherman and Greene 2002). The mechanism of nuclear import could be similar to that described for the Rev protein. It seems that the DNA flap is a major determinant of this mechanism (Figure 1.41), involved in the uncoating of the viral core and formation of the PIC (Zennou, Petit et al. 2000; Arhel, Souquere-Besse et al. 2007; Riviere, Darlix et al. 2010).

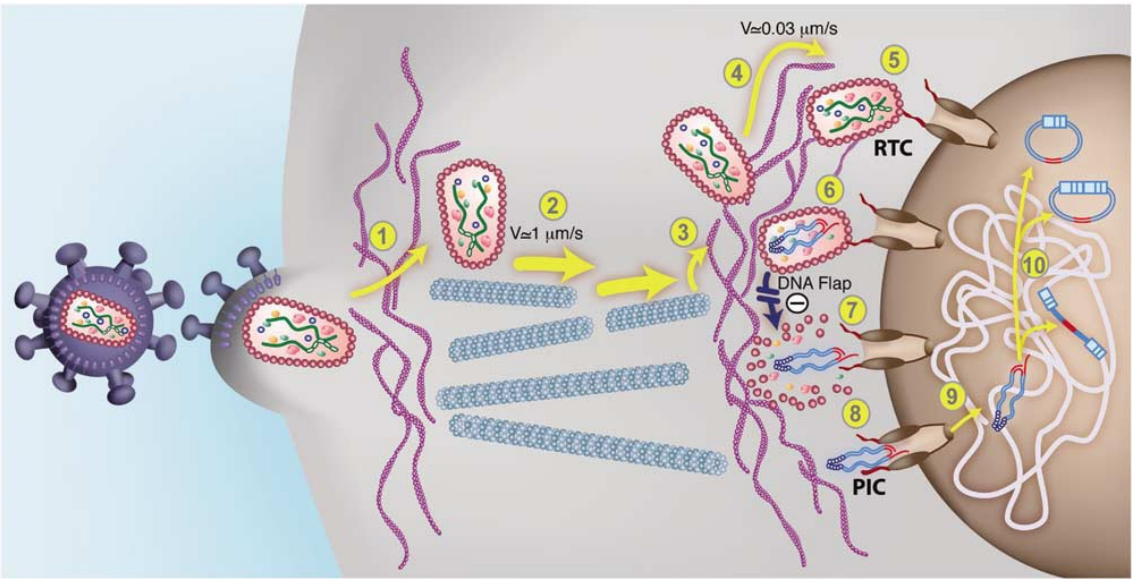


Figure 1.41. Intracellular transport and nuclear import of HIV-1 (Arhel, Souquere-Besse et al. 2007).

1. Entry of HIV-1 and transport on the actin cytoskeleton.
2. The core of HIV-1 is directed to the nucleus via microtubules.
3-6. The core attaches to nuclear pore complexes.
7. Uncoating and formation of the pre-integration complex.
8-9. Nuclear import via nuclear pores.
10. Integration of proviral DNA into the chromatin of the host cell or circularization.

4) Integration of proviral DNA

After nuclear import, the proviral DNA is integrated into the genome of the host cell (Figure 1.41, step 10). This reaction is catalyzed by the viral integrase which specifically recognizes the two LTR ends of the viral DNA. The process of integration can be divided into two sequential steps. First step is an endonucleotide cleavage called 3' processing in which two nucleotides are removed from the 3' ends of newly formed viral DNA resulting in sticky ends. The second step is called strand transfer in which the viral DNA is inserted covalently into the genome of the host cell (Engelman, Mizuuchi et al. 1991). The integrase alone can catalyze this step. However, other proteins such as NCp7 (Carteau, Gorelick et al. 1999), HMGI (Y) and LEDGF increase its efficiency.

Post-integration phase

1) Transcription and translation

The proviral DNA is transcribed by the cellular RNA polymerase II, starting from the promoter located in the 5' long terminal repeat. The Tat protein activates transcription by binding to the internal loop of the TAR stem-loop sequence in 5' of viral RNA (Jeang, Xiao et al. 1999; Gatignol 2007). Then, a large number of different viral mRNAs is generated by an alternative splicing mechanism (Schwartz, Felber et al. 1990). Unspliced mRNAs encode Gag and Pol or are packaged. Mono-spliced mRNAs (4 kb) encode Vif, Env, Vpr and Vpu. Finally, multi-spliced mRNAs (2.4 kb) encode Tat, Rev and Nef. The multi-spliced mRNA can freely leave the nuclear

compartment, while the non-spliced or single-spliced mRNA require the Rev protein for nuclear export (Malim, Tiley et al. 1990). In the cytoplasm, they can be translated or packaged.

The translation of mRNAs depends only on the cellular machinery. The synthesis of the different HIV proteins results from the existence of different mRNA transcripts. In the cytoplasm, the mono-spliced mRNAs are translated into the polyprotein precursor gp160 (Freed, Myers et al. 1990), while the unspliced mRNAs lead to the production of the polyprotein precursors $Pr55^{Gag}$ and $Pr160^{Gag\text{-}Pol}$ which give rise to all other structural, enzymatic and functional proteins of the virus.

2) Assembly and budding (De Guzman, Wu et al. 1998; Amarasinghe, De Guzman et al. 2000)

The formation of a viral particle requires as a first step, the transport of the Gag precursor at the plasma membrane and its association with various viral and cellular partners. In fact, $Pr55^{Gag}$ polyproteins alone can produce virus-like particles (VLPs) (Gheysen, Jacobs et al. 1989), although the presence of genomic RNA, enzymes encoded by pol and envelope glycoproteins are necessary for the production of infectious viruses (Adamson and Freed 2007). A dimeric RNA is encapsidated in a viral particle. Its packaging into virus particles depends on the specific interaction between the NC domain of the Gag precursor and a region located 5' of the viral RNA (Cimarelli, Sandin et al. 2000; Grigorov, Decimo et al. 2007). The binding of RNA-Gag complexes to the plasma membranes is ensured by the MA region of the $Pr55^{Gag}$ through its myristoylated N-terminal part and its cluster of basic residues. The budding mechanism still remains highly debated. It was proposed that the budding of infectious particles can occur at late endosomes (Nydegger, Foti et al. 2003; Pelchen-Matthews, Kramer et al. 2003; Ono and Freed 2004; Grigorov, Arcanger et al. 2006; Gousset, Ablan et al. 2008). Recent observations suggest that the assembly is targeted to the plasma membrane and that the viral particles in the endosomes may result from endocytosis deficient buds (Jouvenet, Neil et al. 2006; Finzi, Orthwein et al. 2007; Jouvenet, Bieniasz et al. 2008), while compartments associated with late endosomes could correspond to invaginations of the plasma membrane (Denek, Pelchen-Matthews et al. 2007; Welsch, Keppler et al. 2007).

Formation of infectious particles then requires the recruitment of Gag-Pol (Shehu-Xhilaga, Crowe et al. 2001; Shehu-Xhilaga, Hill et al. 2002), Env (Adamson and Freed 2007; Corbin, Grigorov et al. 2008), Vpr, Vif and Nef and cellular proteins (Ott 2002; Swanson and Malim 2008; Wolf and Goff 2008). The recruitment by the Gag-Pol precursor of the protease, RT and IN, is mediated through an interaction with the capsid domain of $Pr55^{Gag}$ (Muriaux, Darlix et al. 2004), while the envelope proteins are recruited by the MA domain of Gag (Cosson 1996). The virion also encapsidates the $tRNA^{Lys3}$ primer (Khorchid, Javanbakht et al. 2000; Kleiman 2002). Budding occurs from the plasma membrane and/or through membrane fusion of intracellular compartments, probably at the level of lipid rafts (Nguyen and Hildreth 2000) or tetraspanin-enriched membrane microdomains (Nydegger, Khurana et al. 2006; Grigorov, Attuil-Audenis et al. 2009).

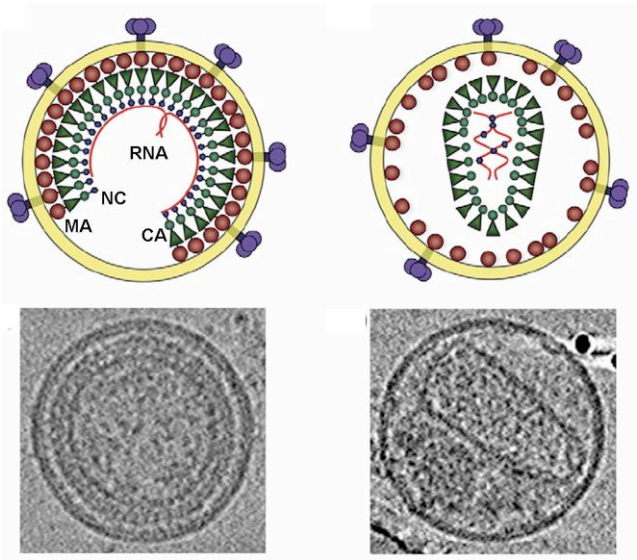


Figure 1.42. Immature (left) and mature (right) HIV-1 virions. Schematic models (top) and central slices through tomograms of HIV-1 particles derived by electron cryotomography (bottom) (Berthoux, Pechoux et al. 1997).

3) Maturation

Once the virion is released from the host cell, the virus is matured (Adamson and Freed 2007) which leads to changes in the morphology of the viral particles. The immature virion is a roughly spherical shell of radially extended Gag molecules (Figure 1.42). The N-terminal MA domains of the Gag polyproteins are bound to the inner viral membrane, and the C-terminal domains of the Gag molecules project into the center of the virus. Maturation involves proteolytic cleavage of Gag and Gag-Pol polyproteins by the viral protease (PR). These newly processed proteins then reassemble to form the distinct layers of the mature virion: MA remains associated with the inner viral membrane (the 'matrix' layer), NC coats the viral RNA genome (the 'nucleocapsid' layer), and CA assembles into the conical capsid that surrounds the nucleocapsid and its associated enzymes, reverse transcriptase (RT) and integrase (IN). This maturation step allows obtaining an infectious particle.

1.10.5. Nucleocapsid protein (NCp7)

Structure of NCp7

In the HIV-1 viral life cycle, the nucleocapsid protein (NC) is first generated as a functional domain of the Gag precursor, $Pr55^{Gag}$ (Figure 1.43). The mature form of NC is then generated by a successive series of protein cleavage by PR.

The initial PR cleavage of $Pr55^{Gag}$ results in the formation of NCp15, a peptide comprising the NCp7, p1 (SP2) and p6 domains (Shehu-Xhilaga, Kraeusslich et al. 2001). The NCp15 protein must be bound to RNA for PR to cleave it (Sheng, Pettit et al. 1997; Mirambeau, Lyonnais et al. 2007). The cleavage of NCp15 leads to NCp9 (NCp7 with the SP2 peptide still attached). The last cleavage allows the release of NCp7. Thus, during the proteolysis, the NC exists in two intermediate forms, NCp15 and NCp9, before taking its final form of 55 amino acids which is the

predominant form detected in mature HIV-1 particles (Henderson, Bowers et al. 1992; Tanchou, Decimo et al. 1998; Coren, Thomas et al. 2007). It is important to note that in several studies, the 71 amino acid precursor of NC that contains the p2 fragment was used (de Rocquigny, Ficheux et al. 1991; Morellet, Jullian et al. 1992; Morellet, de Rocquigny et al. 1994).

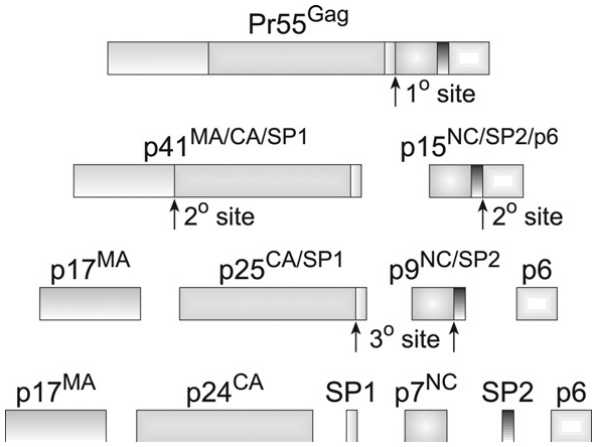


Figure 1.43. Proteolytic cleavage of HIV-1 Gag by PR (Thomas and Gorelick 2008).

NCp7 is a small protein characterized by N-terminal and C-terminal domains rich in basic residues, flanking two motifs called zinc fingers of the sequence – Cys-X_2-Cys-X_4-His-X_4-Cys (CCHC) (where X denotes variable amino acids) (Berg 1986). These CCHC motifs chelate Zn^{2+} with high affinity (K = 10^{-13} - 10^{-14} M) (South, Blake et al. 1990; Khan and Giedroc 1992; Mely, de Rocquigny et al. 1996; Bombarda, Grell et al. 2007). The binding of zinc to the fingersallows the transition from a random structure to a folded and highly constrained form, which supports the NC biological activities (Morellet, Jullian et al. 1992; Summers, Henderson et al. 1992; Morellet, de Rocquigny et al. 1994). Many studies have shown that mutations affecting the zinc fingers have a negative effect on viral replication. Substitution of His23 by a Cys modifies the folding of the proximal finger and leadsto the production of noninfectious viral particles (Demene, Dong et al. 1994; Tanchou, Decimo et al. 1998). The spacer between the two zinc fingers is highly conserved (Figure 1.40) and plays also an important role in the biological activity of the protein (Morellet, de Rocquigny et al. 1994). The folding the zinc fingers also lead to the appearance on their top of a hydrophobic plateau involving Val13, Phe16, Thr24, Ala25, Trp37, Gln45 and Met46 amino acids (de Guzman, Wu et al. 1998; Morellet, Demene et al. 1998; Amarasinghe, De Guzman et al. 2000; Beltz, Clauss et al. 2005; Godet, de Rocquigny et al. 2006; Bourbigot, Ramalanjaona et al. 2008). Among these residues, Phe16 and Trp37 play a crucial role, being very important for specific interaction with nucleic acids (Darlix et al, 2011). Moreover, the spatial proximity allows these residues to stack with neighbor nucleic bases (Figure 1.44) (Mely, Piemont et al. 1993; Lam, Maki et al. 1994). Mutation of these residues causes the formation of noninfectious viral particles. The structure of the NC protein solved by NMR spectroscopy (Levin, Guo et al. 2005) is given on Figure 1.44B.

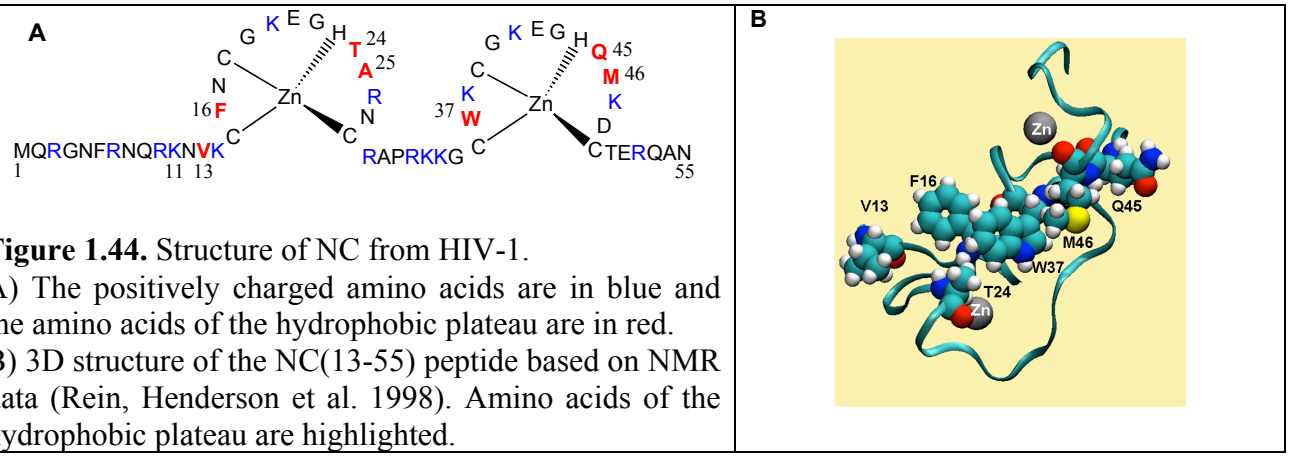


Figure 1.44. Structure of NC from HIV-1.
A) The positively charged amino acids are in blue and the amino acids of the hydrophobic plateau are in red.
B) 3D structure of the NC(13-55) peptide based on NMR data (Rein, Henderson et al. 1998). Amino acids of the hydrophobic plateau are highlighted.

Role of NCp7 protein in the viral life cycle

Role of NC in reverse transcription

The conversion of gRNA to full-length dsDNA is an essential step in retroviral replication, which requires at least two strand-transfers (de Rocquigny, Ficheux et al. 1991), and the RNA- and DNA-dependent polymerase transcription of RNA and cDNA sequences, respectively (Figure 1.40). Reverse transcriptase (RT) alone can accomplish many of these steps, albeit inefficiently.

In the context of the virus, RT must perform its functions in the context of gRNA being completely coated with NC. Through its chaperone properties, NC can initiate RNA conformational changes by directing the rearrangement of nucleic acid molecules into their most stable conformation and can thus, enhance strand annealing (Rein, Henderson et al. 1998). This activity has also been shown to dramatically reduce pausing during synthesis by RT (Ji, Klarmann et al. 1996; Wu, Henderson et al. 1996; Drummond, Mounts et al. 1997; Klasens, Huthoff et al. 1999). Pause sites are positions where RT synthesis is stalled, often by the presence of a secondary structure such as a hairpin. Mutational analysis of NC function indicated that the conserved features of the NC protein, including the basic amino acids flanking the first zinc finger, as well as the zinc finger structures and certain residues therein, all contribute to the ability of NC to reduce pausing. Due to the chaperone properties, NC is thought to assist the reverse transcriptase to convert the HIV-1 RNA genome into a linear double-stranded DNA (Levin, Guo et al. 2005). NC plays key functions in several steps of reverse transcription:

Initiation

The initiation step of reverse transcription requires the hybridization of the primer cellular $tRNA^{Lys3}$ selectively encapsidated (Kleiman 2002), to the complementary 18 nucleotides of the PBS sequence (Primer Binding Site) located in the 5' region of the viral genome. This annealing is stimulated by the presence of NC in the Gag (Cen, Huang et al. 1999; Feng, Campbell et al. 1999) or NCp7 in its mature form (de Rocquigny, Gabus et al. 1992; Lapadattapolsky, Pernelle et al. 1995; Feng, Copeland et al. 1996; Hargittai, Mangla et al. 2001; Tisne, Roques et al. 2003; Hargittai, Gorelick et al. 2004). The mature form appears more effective for hybridization of $tRNA^{Lys3}$, suggesting that the chaperone properties of NCp7 are greater than those of Gag (Guo,

Saadatmand et al. 2009). This annealing step involves the partial destabilization of the tRNA structure so that its 3' 18 nucleotides can form base pairs with the PBS region on the gRNA. A similar annealing can also be performed using heat to destabilize the intramolecular base pairing of these sequences, indicating the utility of NC in performing this at physiological temperatures. Reverse transcription from this primer to the 5' end of the RNA genome forms the DNA product called minus-strand strong-stop DNA (Thomas and Gorelick 2008).

Minus-strand-transfer

The minus-strand-transfer is a necessary step to the complete formation of the minus-strand strong-stop DNA, complementary to the genomic RNA. The transfer of synthesized (-)ssDNA to the RNA R region at the 3' end of the genome, facilitated by NC, is called minus-strand-transfer (Figure 1.36, Step 3) (Peliska and Benkovic 1992). NCp7 plays a major role in this strand-transfer by promoting the annealing of 5'-R cDNA to the 3'-R region of the genome and stimulating the RNase H activity. Since NC facilitates the formation of nucleic acid structures having the maximum number of base pairs, it will favor the displacement of the RNA fragments by the 3'-R region. Hybridization of the TAR sequence (3' of viral RNA) with the complementary cTAR sequence on (-)ssDNA is the rate-limiting step of the transfer (Guo, Wu et al. 2002; Beltz, Clauss et al. 2005). In fact, NC activates the transient opening (fraying) of cTAR terminal base-pairs, leading to a partial melting of the cTAR ends. In its free form, TAR is in equilibrium between closed and partially melted forms. NC binds preferentially to single strands and shifts the equilibrium toward the “open” form.

Second strand-transfer

During the second strand-transfer (Basu, Song et al. 2008), the minus-strand PBS anneals to the plus-strand PBS, forming a circular intermediate. This annealing reaction relies on the ability of NC to destabilize the PBS stem-loops (SLs) (Johnson, Turner et al. 2000; Egele, Schaub et al. 2004), exposing nucleotides that are sequestered in the stem and loop, and activating the fraying of the terminal G–C base pair (Egele, Schaub et al. 2004). PBS can bind up to three NCp7 molecules, with a preferential binding to the 6-TGC-8 sequence in the loop. The binding of NCp7 streches the loop and exposes the loop bases to the solvent, favouring the annealing with the complementary (+)PBS sequence through the loops (Johnson, Turner et al. 2000; Ramalanjaona, de Rocquigny et al. 2007; Bourbigot, Ramalanjaona et al. 2008). Furthermore, NCp7 dramatically restricts the local dynamics of the PBS loop bases at the picosecond and nanosecond timescale, suggesting that NCp7 "freezes" the conformations competent for an annealing through the loops. The changes in the structure and dynamics of the PBS loop which govern the annealing mechanism were strictly dependent on the integrity of the hydrophobic platform at the top of the two folded zinc fingers. As a consequence, the chaperone properties of NCp7 carried by the two zinc fingers are necessary to ensure the specificity of the annealing reaction involved in the second strand transfer (Godet, Ramalanjaona et al. 2011).

Role of NC in nuclear entry

The presence of NC in the PIC is expected to some extent as a consequence of its presence in the RTC. Moreover, experiments have shown nuclear localization of NC after infection (Gallay,

Swingler et al. 1995; Zhang and Crumpacker 2002). The PIC nucleoprotein complex is actively transported into the nucleus of non-dividing cells, but a controversy exists on the mechanism of nuclear entry in quiescent cells (Yamashita and Emerman 2005). Karyophilic properties have been identified in MA (Bukrinsky, Haggerty et al. 1993; von Schwedler, Kornbluth et al. 1994), Vpr (Zhao, Mukherjee et al. 1994; Fouchier, Meyer et al. 1998), IN (Gallay, Hope et al. 1997; Ao, Huang et al. 2007), and the central flap region (Zennou, Petit et al. 2000; Zennou, Serguera et al. 2001; de Rijck and Debyser 2006). It has also been observed that most of these signals can be blocked or eliminated, but this does not completely eliminate the nuclear entry and productive infection (Fouchier, Meyer et al. 1997; Gallay, Hope et al. 1997; Petit, Schwartz et al. 2000; Dvorin, Bell et al. 2002; Limon, Devroe et al. 2002; Limon, Nakajima et al. 2002; Yamashita and Emerman 2006; Marsden and Zack 2007). These results can be explained by several possibilities. First, the different studies have used widely differing virus titers to infect cells, implying that a large number of infectious events can overcome the engineered blocks. Another possibility is that there may be other nuclear localization signals not yet identified. In vitro experiments have identified the central flap as a site of strong NCp7-nucleic acid affinity (Mirambeau, Lyonnais et al. 2007). The central flap with tightly bound NCp7 molecules could play a role in nuclear entry. The role of NCp7 in nuclear entry is totally unknown.

Role of NC in integration

Integration of the full-length vDNA into the chromosomal DNA of the infected cell to form the provirus is the final step of early infection. In vitro, only three components are strictly required for integration; the integrase (IN), the viral LTR ends containing the requisite attachment sites, and a DNA substrate in which to integrate the vDNA (Sinha and Grandgenett 2005). Nevertheless, in vivo infection experiments performed with NC mutant viruses have provided indirect evidence for the involvement of NC in integration (Carteau, Batson et al. 1997; Carteau, Gorelick et al. 1999; Gao, Gorelick et al. 2003; Poljak, Batson et al. 2003). The role of NC in the integration process is not yet clearly determined and three main hypothesis have been raised (Thomas and Gorelick 2008) i) NC enhances the enzymatic steps of integration; ii) NC is involved the formation of a functional intasome, possibly by assisting IN to bind the LTR ends or by stabilizing the IN nucleoprotein complex at the LTR ends; iii) NC may be required for the proper structure of the preintegration complex, notably through binding to the central flap (Arhel, Souquere-Besse et al. 2007).

Role of NC in virus assembly

RNA packaging

During virus assembly, NC is present mainly as a part of the Gag polyprotein. Genetic analyses have demonstrated that the NC domain in the Gag precursor is critical for specific recognition and packaging of gRNA (Aldovini and Young 1990; Dorfman, Luban et al. 1993). Mutations of the zinc fingers or the basic residues in NC can significantly reduce packaging of HIV-1 gRNA. NMR experiments have shown that the NC zinc fingers are responsible for specific interactions with the encapsidation Ψ sequence of the gRNA, while the basic residues contribute to stabilize the complex (Muriaux, Darlix et al. 2004). Specific binding to the Ψ sequence is required for the selective recognition of viral RNA among the large excess of cellular RNAs (Amarasinghe,

De Guzman et al. 2000). NC binds specifically to three of the four step-loop sequences (SL1, SL2 and SL3) (Vuilleumier, Bombarda et al. 1999; Amarasinghe, De Guzman et al. 2000; Amarasinghe, Zhou et al. 2001; Shubsda, Paoletti et al. 2002) that constitute the Ψ sequence.

Structures of NC-SL2 and NC-SL3 complexes were solved by NMR (de Guzman, Wu et al. 1998; Amarasinghe, De Guzman et al. 2000) (Figure 1.45). In both complexes, the main driving forces are the interactions of the loop residues at the top of the stem-loops with the conservatively substituted hydrophobic residues Val13, Phe16, Thr24 and Ala25 at the top of the proximal finger and the Trp37 and Met46 residues at the top of the distal finger. In the NC-SL2 complex, the nucleotide bases of G11 and G9 bind to the hydrophobic cleft of the N- and C-terminal fingers, respectively (Darlix, Lapadattapolsky et al. 1995). The GXG motif of the SL3 loop is also crucial for NC binding.

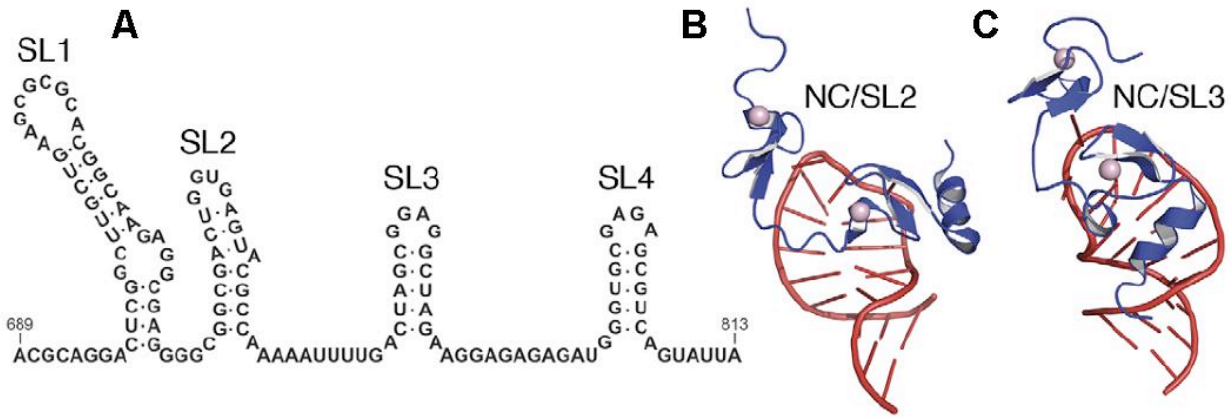


Figure 1.45. (A) Encapsidation sequence of the HIV-1 genome recognized by the Gag polyprotein. (B) and (C) 3D structure of NC complexes with SL2 and SL3. (Figure adopted from (Bourbigot, Ramalanjaona et al. 2008))

Genome dimerization

A necessary step for virus particle maturation is the formation of a stable dimer between the two molecules of gRNA present in the retrovirus particle (D'Souza and Summers 2005). It has been demonstrated that NC is critical for the formation of this complex. The mechanism of its stabilization is thought to begin with a limited base pairing between the genomes, termed a "kissing loop structure", which, with the assistance of NC expands to form a more extended duplex. This process, termed "dimer maturation," results in the formation of a more thermostable RNA dimer, and occurs during the maturation of the virus particle, after budding. NC facilitates the RNA maturation via its nucleic acid chaperone activity, which assists the RNA to find the most thermodynamically stable annealed structure.

Role of NC in genome protection. Non specific RNA/DNA binding

Viral particles contains about 1500-2000 molecules of NC proteins associated with RNA (Krishnamoorthy, Roques et al. 2003). Due to the small size of the virus, this correponds to a concentration of NCp7 in the virus, which is in the mM range. Since NCp7 is able to bind to nearly any sequence of five to seven nucleotides (Fisher, Rein et al. 1998; Vuilleumier, Bombarda et al.

1999), with dissociation constants ranging from 10^{-3} to 10^{-7} M, the high concentration of protein in the virus easily explains the full coating of the genomic RNA by this protein (Darlix, Godet et al. 2011). This coating leads to RNA condensation that is necessary for formation of the viral particle (Lapadattapolsky, Derocquigny et al. 1993; Tanchou, Gabus et al. 1995). It has been observed in vitro that nucleic acids coated with NC are protected against nucleases (You and McHenry 1993; Mely, de Rocquigny et al. 1995; Fisher, Rein et al. 1998; Urbaneja, Kane et al. 1999). Moreover, infection and in vitro experiments further suggested that NC also stabilizes nascent vDNA in the cytoplasm (Krishnamoorthy, Roques et al. 2003). Interestingly, a model of a probably weakly specific interaction was provided by the NMR structure of a NCp7 peptide bound to a short oligonucleotide d(ACGCC) (Morellet, Demene et al. 1998). In this complex, the pentanucleotide sugar-phosphate axis is almost perpendicular to the interfinger linker (Figure 1.46). This interaction is directed by the stacking of Trp37 with G (guanine) that plays a major role in the complex stability. The key role of the Trp-G stacking has been confirmed, using variants of the d(ACGCC) sequence, where each nucleotide has been systematically varied (Urbaneja, Kane et al. 1999; Vuilleumier, Bombarda et al. 1999).

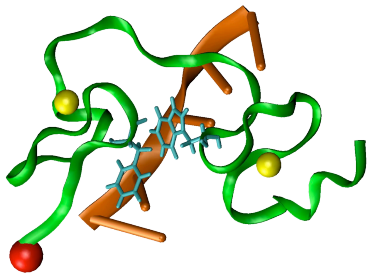


Figure 1.46. NMR-derived structure of NC(13-53) bound to the d(ACGCC) oligonucleotide. Atom coordinates are from (Morellet, Demene et al. 1998).

1.10.7. Tat protein

Structure of Tat

The HIV-1 trans-activator of transcription (Tat) is an 86-101 residue regulatory protein (9–11 kDa) (Jeang, Xiao et al. 1999). The full length Tat protein is divided into six different functional regions (Kuppuswamy, Subramanian et al. 1989) (Figure 1.47):

Region I (residues 1–20) is a proline-rich region and has a conserved Trp11.

Region II (residues 21–40) has seven well conserved cysteines at positions 22, 25, 27, 30, 31, 34 and 37 which are required for the trans-activating activity of Tat (Gregoire, Peloponese et al. 2001), with the exception of Cys31 (Loret, Georgel et al. 1992).

Region III (residues 41–48) has a conserved Lys41 residue that is essential for the trans-activation, whereas other residues Lys-X-Leu-Gly-Ile-X-Tyr play an auxiliary role in its activity (Pantano and Carloni 2005; Shojania and O'Neil 2006). This region is critical for the binding of Tat to tubulin (Karn 1999).

Region IV (residues 49–57) is rich in basic residues and has the rather well conserved sequence ^{49}RKKRRQRRR. It contains the pattern of Tat binding to TAR RNA sequence (Karn

1999), which is essential for the function of trans-activation of transcription of Tat (Wu-Baer, Lane et al. 1995). This region is also involved in the capture of Tat by cells (Herrmann and Rice 1995).

Region V (residues 58–72) is the glutamine-rich region that plays a supporting role in the trans-activating activity of Tat (Karn 1999). This region is involved in the induction of apoptosis of T-lymphocytes via the interaction with tubulin (Ensoli, Buonaguro et al. 1993; Westendorp, Frank et al. 1995).

Region VI constituting the C-terminus of Tat is encoded by the second exon. It contains a RGD motif (Arg78-Gly-Asp80). This region is responsible for some of the cell penetration properties of the protein (Belliard, Romieu et al. 2003).

Based on mutational analysis, Tat appears to be composed of two functional regions. The region responsible for transcription activation and binding of cofactors (residues 1-47) is functionally autonomous (Frankel and Pabo 1988; Ensoli, Buonaguro et al. 1993; Westendorp, Li-Weber et al. 1994; Chang, Samaniego et al. 1997). The second region is the RNA binding region (residues 48-86) (Barillari, Gendelman et al. 1993).

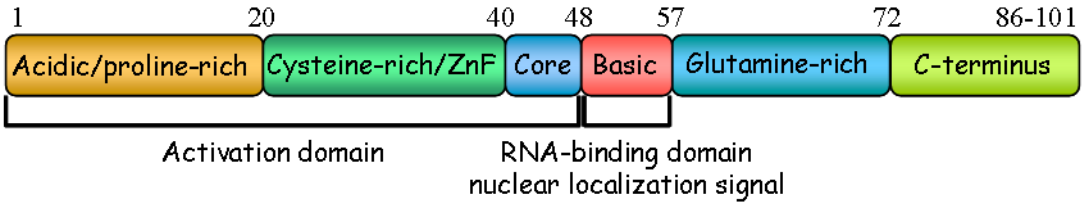


Figure 1.47. Functional domains of Tat protein. Figure adapted from (Tahirov, Babayeva et al. 2010).

Structural studies have been conducted by by circular dichroism and nuclear magnetic resonance (NMR) with Tat proteins of several HIV-1 variants (Gregoire, Peloponese et al. 2001; Watkins, Campbell et al. 2008), including Tat Z2 (Peloponese, Gregoire et al. 2000), Tat Bru (Bayer, Kraft et al. 1995) (Figure 1.48). All studied proteins have more than 70% identity in sequence but present different conformations, which cannot be superimposed.

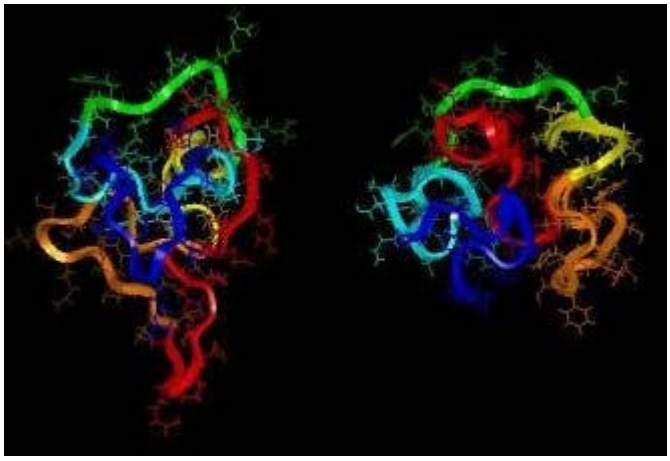


Figure 1.48. NMR studies of Tat proteins (Campbell and Loret 2009). 3D structures of Tat Z2 (left) and Tat Bru (right) obtained from NMR studies (Bayer, Kraft et al. 1995; Peloponese,

Gregoire et al. 2000). Region I is depicted in red, region II (cysteine-rich region) in orange, region III in yellow, region IV (basic region) in green, region V in light blue, region VI (residues 73–86/87) in blue. The Tat Z2 variant used had chemically modified cysteines, which affect the biological activity and 3D structure.

The first NMR structural study was performed at pH 6.3 under reducing conditions using an 86-residue Tat Z2 variant in the presence of dithiothreitol (DTT) (Bayer, Kraft et al. 1995), which leads to the loss of activity of the protein. Protein Tat Z2 (Zaire isolate, subtype D) has a rigid core formed by the region III (central area) and the region V, with the N-terminal region sandwiched between these two domains. In this model, the C-terminal domain is in contact with region V, while the cysteine-rich region (region II) and the basic region (region IV) are very flexible. In addition, the RGD motif forms a loop that is exposed to solvent. Later study of the 86-residue Tat Bru (Peloponese, Gregoire et al. 2000) was performed in the absence of reducing agents at pH 4.5. The folding of Tat Bru (subtype B) shows that region III (38-48) and Region V (60-72) Tat Bru are more structured, than the basic region (49-59). Functional residues, (except Cys37), in the basic region (IV) and cysteine-rich region (II), are exposed to solvent.

In contrast, a more recent NMR study showed that the 72-residue long Tat exists predominantly in a random coil conformation at pH 4.1 (Shojania and O'Neil 2006). Dynamics analysis suggests that the most likely region to fold is the Cys-rich region, and formation of disulfide bonds could stabilize this region locally. However, the high positive charge density, paucity of hydrophobic residues, and dynamics analysis suggest that the remainder of the protein is unlikely to form a stable conformation even at neutral pH. The absence of stable structure in Tat was confirmed by circular dichroism, dynamic light scattering, and small angle X-ray scattering methods on a chemically synthesized Tat(1-101) protein (Foucault, Mayol et al. 2010). Altogether, these results suggest that the 72-, 86- and 101-residue long Tat proteins are highly flexible and probably not structured in aqueous solution.

Functions of Tat

A) Tat activates transcription

The HIV-1 Tat is essential for the productive and processive transcription of full-length proviral DNA from the HIV-1 long terminal repeat (LTR) promoter (Jeang, Xiao et al. 1999). The absence of Tat decreases the trans-activation rate several hundred times. First, Tat recruits the human positive transcription elongation factor b (P-TEFb) to the trans-activating response (TAR) RNA element at the 5′ end of nascent transcripts through its direct interaction with cyclin T1 (CCNT1) (Figure 1.49) (Fujinaga, Cujec et al. 1998; Wei, Garber et al. 1998). P-TEFb is a heterodimer, composed of the cyclin-dependent kinase, CDK9, and its partner cyclin T1 or the minor forms T2a or T2b (Peng, Zhu et al. 1998; Wei, Garber et al. 1998). Tat forms a ternary complex with TAR and P-TEFb, which has been proposed to be necessary and sufficient for transcriptional elongation. The CDK9 kinase activity of the recruited P-TEFb heterodimer results in hyperphosphorylation of the carboxyl-terminus domain (CTD) of the largest subunit RNAP II, leading to efficient elongation (Isel and Karn 1999). Moreover, it has recently been reported that Tat transcriptional activity is regulated by lysine methylation (Van Duyne, Easley et al. 2008), and

that Tat interacts also with a histone chaperone nucleosome assembly protein (Vardabasso, Manganaro et al. 2008).

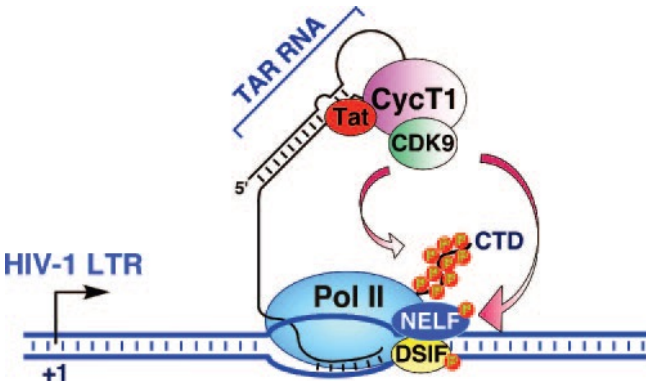


Figure 1.49. Schematic representation of Tat transactivation of HIV-1 transcription (Qiang and Yik 2006).

B) Cell penetrating activity of Tat

Several years ago, Frankel and Pabo (Frankel and Pabo 1988) and Green and Lowenstein (Green and Loewenstein 1988) demonstrated that extracellular HIV Tat can cross the plasma membrane, enter the cell and reach the nucleus. The ability of Tat to cross membranes was attributed to its C-terminus (residues 37-72) (Fawell, Seery et al. 1994). In fact, the minimal sequence sufficient for membrane translocation is Tat (48-57) (Vives, Brodin et al. 1997; Wender, Mitchell et al. 2000). Membrane translocation is also favoured by the RGD domain of Tat that can bind integrin receptor on the cell membrane (Vogel, Lee et al. 1993). Alternatively, membrane translocation of the Tat peptide is mediated by heparan sulphate (HS) proteoglycans (HSPGs), which are expressed on the surface of almost all cell types (Tyagi, Rusnati et al. 2001; Richard, Melikov et al. 2003; Vives, Richard et al. 2003). In the endocytosed vesicles, HS is degraded by heparinase, which releases the Tat protein. It was observed that TAT was internalized by means of all three endocytotic pathways (macropinocytosis, caveolae-mediated and clathrin dependent endocytosis), depending on peptide concentration (Duchardt, Fotin-Mleczek et al. 2007). The Tat peptide can be used as a powerful tool for delivery of biologically active molecules (Tunnemann, Martin et al. 2006). For instance, the cell membrane penetration capacity of Tat was applied for the transport of fluorophores, nucleotides, proteins (Dietz, Valbuena et al. 2006), drugs, supermagnetic nanoparticles (Zhao, Kircher et al. 2002), imaging and radiotherapeutic agents (Polyakov, Sharma et al. 2000) and genes into cells (Patel, Zaro et al. 2007).

C) Apoptosis induction by Tat

There are two major cell death-signalling pathways governing apoptosis: the extrinsic (or death receptor) pathway and the intrinsic pathway. The extrinsic pathway is triggered by external stimuli or ligands, sensed by cell-death receptors. Mitochondria on the other hand play a pivotal role in the intrinsic pathway by releasing apoptosis-inducing molecules, mainly cytochrome c, followed by activation of caspases-9 and -3, respectively (Ruwanpura, McLachlan et al. 2008). HIV-1 Tat is able to induce the intrinsic pathway of apoptosis in a number of human cell lines. In

vitro studies further suggest that Tat promotes apoptosis during HIV-1 infection (Kim, Avraham et al. 2003; Giacca 2005). Extracellular Tat, produced by adjacent infected macrophages, microglia or astrocytes, has also been shown to induce neuronal death by interacting with the low density lipoprotein receptor-related protein (LRP) on the surfaces of neurons (Kim, Martemyanov et al. 2008).

A link between microtubule polymerization and the pro-apoptotic effect of Tat has been suggested since the interaction of Tat with both non-polymerized subunits of microtubules (αβ-tubulin dimers) and polymerized microtubules, leads to mitochondrial induction of apoptosis (de Mareuil, Carre et al. 2005; Giacca 2005). Regions II and III of Tat including the conserved Cys37 and Phe38 as well as binding of zinc to the Cys-rich domain are crucial for Tat-tubulin interactions. Since cells are dynamic units requiring constant polymerization and depolymerisation of microtubules, this change in microtubule stabilization can act as a powerful inducer of the intrinsic pathway of cellular apoptosis (Campbell, Pesquier et al. 2004; de Mareuil, Carre et al. 2005).

Tat has also been shown to upregulate the tumour necrosis factor-related apoptosis-induced ligand (TRAIL), and caspases-3 and -8 in human monocytes (Miura, Misawa et al. 2003). TRAIL is a member of death inducing ligands of the tumour necrosis factor (TNF) family that activates caspases-3 and -8, resulting in apoptosis. Tat also upregulates Fas, Fas ligand (FasL/CD95L) and TNF-a in human astrocytes (Chauhan, Turchan et al. 2003). The Fas/FasL system comprises an important mechanism for regulating apoptotic cell death. The ability of Tat to recruit many apoptosis-inducing proteins suggests that Tat induces apoptosis through multiple mechanisms (Dabrowska, Kim et al. 2008), including microtubule stabilization by Tat as well as recruitment of various apoptosis-inducing mechanisms, such as the upregulation of Fas, FasL, TRAIL, Bax and caspase-8 (Chauhan, Turchan et al. 2003; Giacca 2005; Romani, Engelbrecht et al. 2010).

RESEARCH OBJECTIVES

In order to understand the molecular basis of various diseases, and establish assays for high throughput screenings and molecular diagnostics, the characterization and quantification of interactions of proteins with different targets such as nucleic acids, proteins and membrane in various media (solutions, biological fluids, micro-arrays, cells…) is strongly demanded. Fluorescence spectroscopy is the technique of choice to reach these objectives due to its capability to decrease the measurements down to the single molecule level and the exquisite sensitivity of many fluorophores to their environment. However, this technique is limited by the availability of optimized fluorophores that combine small size and ability to monitor sensitively and site-specifically conformational changes during interactions. Due to the limitations of the available fluorescent dyes, two-colour fluorophores of the 3-hydroxychromone (3HC) family are highly promising for biomolecular studies due to a convenient ratiometric signal and additional information channels for the analysis of the labeling site properties. These fluorophores exhibit two excited state forms: the initially excited normal (N*) and the tautomeric (T*) forms due to an excited state intramolecular proton transfer (ESIPT). Since the ESIPT reaction in these dyes is strongly sensitive to the polarity, H-bonds and electric fields, the ratio of the two emission bands could be used to sensitively monitor environmental changes. A critical advantage of ratiometric dyes over conventional one-band dyes is that this ratio does only depend on the environmental changes and not on the local concentration of the dyes or the instrumentation.

In this context, the objective of the thesis is to develop a methodology for sensing interactions of proteins with oligonucleotides and membranes based on environment-sensitive labels from the 3-hydroxychromone family. For this purpose, synthesis of new fluorescent dyes based on 3HC dye with improved properties is required. Also in order to reveal the role of individual residues in peptide-ODN interactions, fluorescent amino acid and nucleoside analogues should be synthesized and applied to any position of peptide and ODN, respectively. The HIV-1 NCp7 protein was selected for studying such interactions due to its relatively small size and its ability to bind various DNA/RNA targets. Moreover, an additional objective was to design new ratiometric probes to site-selectively investigate the interaction of NCp7 with lipid membranes, in order to validate the latter as a potential NCp7 target and provide new clues for understanding the pleiotropic role of NCp7 in the viral life cycle.

CHAPTER 2

RESULTS AND DISCUSSION

2. RESULTS AND DISCUSSION

2.1. Investigation of peptide – oligonucleotide interactions by 3-hydroxychromone label

2.1.1. Fluorescent amino acid analogue for studying peptide – oligonucleotide interactions

Biomolecular interactions usually decrease the polarity at the labeling site due to screening from water by the binding of the protein partner. This change in the polarity can be readily detected by environment-sensitive labels through changes in their fluorescence intensity or shifts in their emission maximum. However, applications of these dyes for sensing interactions of peptides with oligonucleotides (ODNs) have not been described so far, likely because the ODN environment is relatively polar.

Recently, a 2-furyl-3-hydroxychromone label (FCL) has been developed in our laboratory for studying peptide-oligonucleotide interactions due to its high sensitivity to polar solvents and H-bonding interactions (Shvadchak, Klymchenko et al. 2009). FCL label undergoes excited-state intramolecular proton transfer (ESIPT), resulting in the emission of both the normal (N*) excited state and the ESIPT product photo-tautomer (T*). Using solid phase peptide synthesis, FCL was coupled to the N-terminus of the nucleocapsid protein of HIV-1 (NC) (Figure 2.1). Most functions of NC are connected with its DNA/RNA binding properties. The truncated nucleocapsid protein NC(11-55) was used in these studies to avoid nucleic acid aggregation which is caused by the non-folded N-terminal domain of NC. It was shown that the label at the N-terminus of NC does not alter the nucleic acid binding and chaperone properties of the protein.

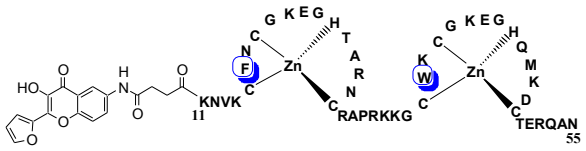


Figure 2.1. Structure of FCL-NC(11-55). The aromatic amino acids of NC are marked in blue.

The FCL label reports on the NC-oligonucleotide interaction by a change in the ratio of its two emission bands which results from a decrease in the exposure of the labeling site to bulk water induced by its proximity to nucleic bases. The response of the labeled peptide on interaction with SL2 RNA, SL3 RNA (Figure 2.2), Δ(-)PBS DNA and d(ACGCC) was investigated and correlated with the known 3D structure of these complexes. In contrast to most dyes reported so far, which are strongly quenched on oligonucleotide binding, no quenching accompanies the binding of oligonucleotides to the FCL-labeled NC. Even more interestingly, the intensity ratio of the two emission bands of FCL-NC(11-55) strongly depends on the ODN sequence. This ratio could be unambiguously correlated with the proximity of the NC N-terminus to the ODN and the possibility of the FCL probe to stack with the ODN bases, showing that the interaction could be site-specifically characterized. Moreover, by simply measuring the ratio at the two emission wavelengths, it was possible to determine the peptide-ODN binding parameters and distinguish multiple binding sites in ODNs, which is rather difficult using other fluorescence methods.

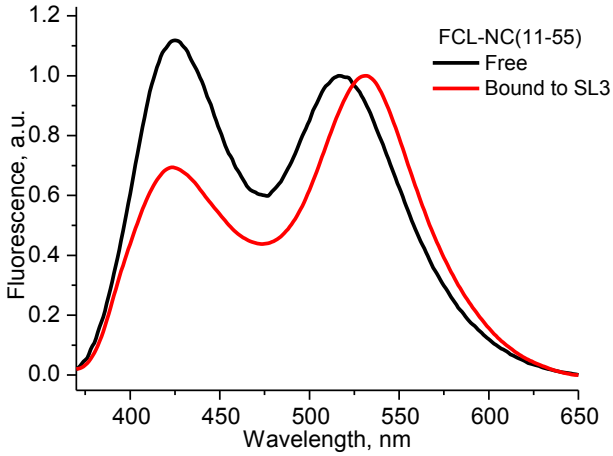


Figure 2.2. Effect of SL3 RNA interaction on the fluorescence spectra of FCL-NC(11-55). The spectra were normalized at the maximum of the long wavelength T* band.

In the present work, a fluorescent amino acid analogue (Figure 2.3) was synthesized and applied to different positions of NC in order to site-selectively sense the proximity of nucleic acid bases in peptide-nucleic acid complexes through a change in the color of its two band emission. The organic synthesis was started from L-tyrosine to get the L-amino acid bearing the 3HC fluorophore protected with the Fmoc group, which is required for solid phase peptide synthesis. As a next step, the enantiomeric purity of the obtained fluorescent amino acid analogue (3HCaa) was checked by using (R)- and (S)-Mosher agents. The NMR spectra of the diastereomeric Mosher derivatives were compared and no signal of the opposite diastereomer was found in each spectrum, indicating that the synthesis did not produce any racemization.

To demonstrate the proof of principle of the use of this fluorescent amino acid, we introduced it at three different positions of a NC(11-55) peptide: NC(11-55)-F16-3HCaa, NC(11-55)-A30-3HCaa and NC(11-55)-W37-3HCaa, where the 3HCaa was substituted for Phe16, Ala30 and Trp37 residues, located in the first zinc finger, the linker and the second zinc finger, respectively (Figure 2.3). The fluorescent amino acid was shown to preserve the structure, the activity and the nucleic acid binding properties of the peptide. Labeled peptides in aqueous solutions showed dual emission where the short- and long-wavelength bands can be assigned to N* and T* forms, respectively. The observed intensity ratio of the two emission bands, N*/T*, is close to that for 3HCaa in methanol, indicating that the 3HCaa fluorophore is partially screened from water by the peptide.

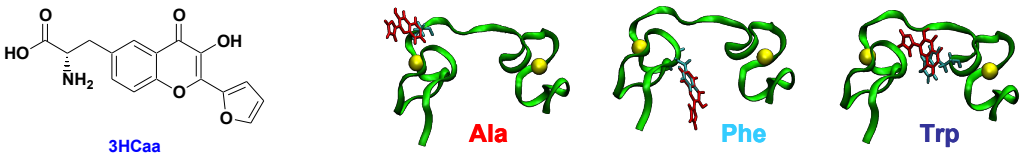


Figure 2.3. Structure of the fluorescent amino acid (3HCaa) and 3D structures of NC with Ala30, Phe16 and Trp37 residues substituted by a 3HCaa. Zn atoms are presented as yellow spheres. Structures of peptides were drawn based on NMR data.

The ODNs selected for binding to the labeled NC peptides were SL2 and SL3, two stem-loops of the HIV-1 RNA encapsidation sequence and ΔP(-)PBS, the cDNA copy of the primer binding site, deleted of its 5' overhang. These three ODNs bind NC with high affinity and 1:1 stoichiometry, and their structures in complex with NC are known. The fluorescence ratiometric response obtained for these DNA/RNA targets was found to nicely correlate with the 3D structure of the NC/ODN complexes, indicating that the changes in the intensity ratios could be used to site-selectively characterize the binding sites and determine the binding parameters of the peptide to the ODNs. As an example, Figure 2.4 shows that in the complex of NC with SL3 RNA, the three positions substituted by the 3HCaa are not equivalent. The high N*/T* ratio observed for the peptide NC(11-55)-A30-3HCaa, suggests a highly exposed localization of 3HC in the complex, in line with an absence of direct interaction with the ODN. In contrast, the much lower values observed for NC(11-55)-W37-3HCaa and NC(11-55)-F16-3HCaa suggest an efficient stacking of the 3HC fluorophore with the ODN bases. The obtained results established a correlation between the fluorescence response (N*/T* ratio) of the amino acid analogue and its proximity to the interacting ODN bases (available from NMR data), providing thus a new methodology for site-specific monitoring of peptide-ODN proximity.

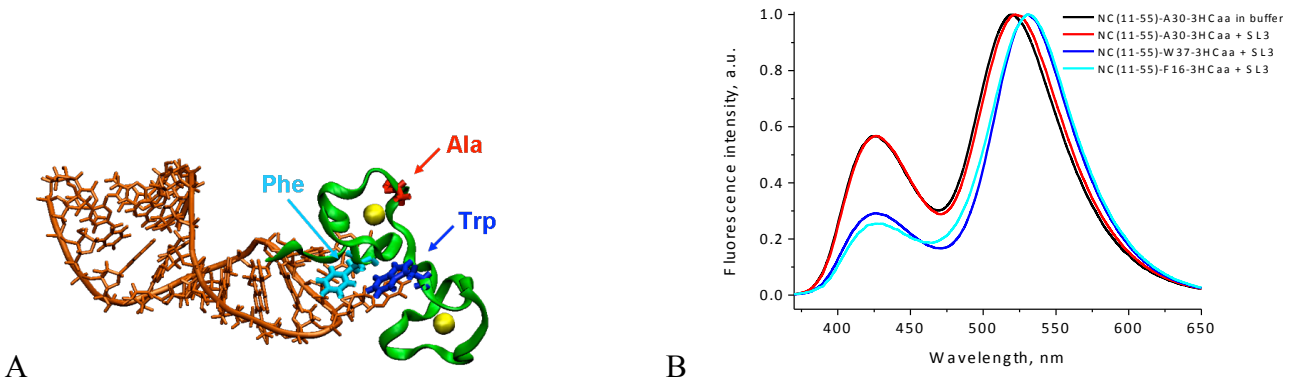


Figure 2.4. (A) 3D structure of NC-SL3 complex based on NMR data. (B) Normalized fluorescence spectra of the NC(11-55)-F16-3HCaa, NC(11-55)-A30-3HCaa, NC(11-55)-W37-3HCaa peptides in buffer and in complex with SL3 RNA.

Thus, this amino acid provides a new and universal approach for site-specific monitoring of peptide-nucleic acid proximity, opening a broad range of applications in the characterization of peptide-oligonucleotide interactions.

All these results are presented in manuscript (**Publication 1**).

Publication 1

Two-color fluorescent amino acid for proximity measurements in peptide-nucleic acid complexes

Fluorescent amino acid

DOI: 10.1002/anie.200((will be filled in by the editorial staff))

Two-Color Fluorescent Amino Acid for Proximity Measurements in Peptide-Nucleic Acid Complexes

Aleksandr V. Strizhak#, Viktoriia Y. Postupalenko#, Volodymyr V. Shvadchak, Nelly Morellet, Eric Guittet, Vasyl G. Pivovarenko, Andrey S. Klymchenko, Yves Mély**

Non-natural amino acids are important tools to investigate peptides and proteins, since they can be incorporated at any position, and thus, serve as local probes for site-selective monitoring of protein properties and functions. Probably, the most popular ones are those bearing fluorine[1] or fluorophores.[2] Fluorescent amino acids are particularly attractive, due to the ultimate sensitivity of fluorescence-based methods. Since the natural fluorescent amino acid tryptophan suffers from poor fluorescence properties, continuous efforts have been done to design amino acids with improved fluorophores. Of special interest in this respect are environment-sensitive (or solvatochromic) fluorophores, which change their emission properties in response to changes in their environment (Figure 1). Biomolecular interactions commonly decrease the polarity at the labeling site due to screening from water and thus, can be readily monitored by environment-sensitive fluorescent amino acids.[2c] For instance, a Prodan-based amino acid was used for monitoring binding of S-peptide with ribonuclease S[3] and δ-opioid receptor with antagonists,[4] and to estimate the local dielectric constant of the B1 domain of the staphylococcal protein G.[2a] Moreover, an amino acid based on 4-dimethylaminophthalimide (4-DMAP) fluorophore was applied to sense binding of labeled octapeptides to the 14-3-3bp protein.[5] Finally, improved analogues of 4-DMAP, such as 6-dimethylaminonaphthalimide, were recently applied to investigate the SH2 phosphotyrosine binding domains,[6] major histocompatibility complexes (MHC) at the cell surface[2c] and peptide-calmodulin interactions.[7]

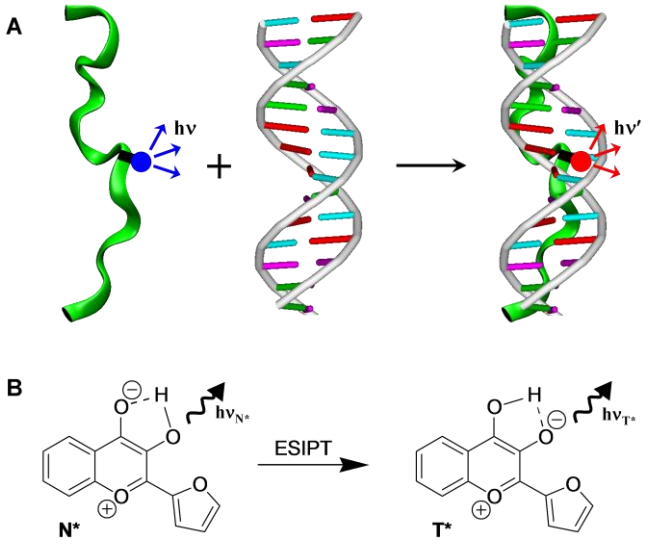


Figure 1. Monitoring of peptide/oligonucleotide interaction using solvatochromic fluorescent amino acids. (A) Principle: the interaction of the labeled peptide (in green) with oligonucleotides changes the emission color of the fluorescent amino acid. (B) ESIPT reaction in 3-hydroxychromone dyes.

However, most examples in the literature are focused on protein-protein interactions, while applications of fluorescent amino acids for sensing peptide-oligonucleotide (ODNs) interactions have been poorly explored (Figure 1A). To address this problem, we selected 3-hydroxychrome (3HC) fluorophores which undergo excited-state intramolecular proton transfer (ESIPT),[8] resulting in the emission of both a normal (N*) excited state and an ESIPT product tautomer (T*) (Figure 1B). The dual emission of 3HC dyes is highly sensitive to polarity and H-bonding interactions.[9] The 2-(2-furyl)-3HC label is particularly interesting due to the high sensitivity of its dual emission to polar environments.[9a, 9c, 10] Being attached to the N-terminus of peptides and oligocations, it shows strong changes in its dual emission upon interaction with ODNs.[11] This prompted us to develop an L-amino acid based on the 2-(2-furyl)-3HC fluorophore.

As a target protein for labeling, we selected the nucleocapsid protein (NC) of the Human Immunodeficiency Virus, type 1 (HIV-1), which plays a crucial role in the viral life cycle.[12] NC is a small (55 amino acids) basic protein, characterized by two rigid zinc fingers connected by a flexible basic linker and flanked by poorly folded N- and C-terminal basic domains.[13] NC binds both specifically[14] and non specifically,[14c, 15] to a large range of nucleic acid sequences. Specific binding is mainly mediated through the folded finger motifs, with a key role being played by the hydrophobic residues in the fingers.[14b, 14d, 14f]

In the present work, we synthesized an L-amino acid analogue based on 2-(2-furyl)-3HC dye and incorporated it at three

[*] A.V. Strizhak#, V.Y. Postupalenko#, Dr. V.V. Shvadchak, Dr. A.S. Klymchenko, Prof. Y. Mély
Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74, Route du Rhin, 67401 ILLKIRCH Cedex, France
Fax: (+) 33 368 854313
E-mail: andrey.klymchenko@unistra.fr, yves.mely@unistra.fr
Homepage: http://www-lbp.unistra.fr/

A.V. Strizhak, Prof. V.G. Pivovarenko
Department of Chemistry, Kyiv National Taras Shevchenko University, 62A, Volodymyrs'ka street, 01033 Kyiv, Ukraine

Dr. N. Morellet, Dr. E. Guittet
CNRS UPR 2301, Chimie et biochimie structurales, Institut de Chimie des Substances Naturelles, Centre de recherche de Gif-sur-Yvette, 1, Avenue de la Terrasse, 91190 Gif-sur-Yvette, France

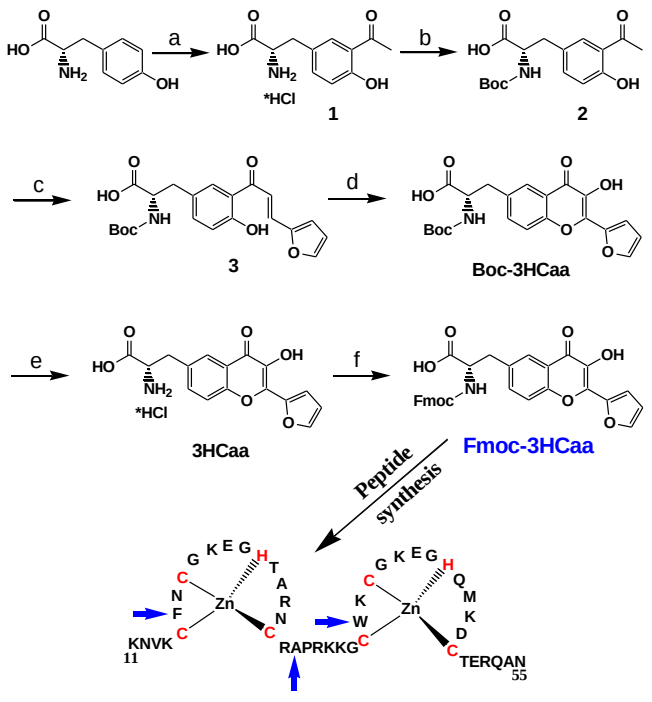
[#] These authors contributed equally to the work.

[**] This work and VYP were supported by ANR (ANR-07-BLAN-0287), ANRS, CNRS and Université de Strasbourg. AVS was supported by ARCUS Alsace. We thank to Cyril Antheaume and Patrick Wehrung for NMR and Mass measurements, respectively.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

different positions of the NC(11-55) peptide corresponding to the finger domain of NC, which allowed us to site-selectively investigate the interaction of NC(11-55) with ODNs. The ODNs selected for binding (Figure S1, Supporting Information) to the labeled peptides were SL2 and SL3, two stem-loops of the HIV-1 RNA encapsidation sequence[14a, 14c, 14f] and ΔP(-)PBS, the cDNA copy of the primer binding site, deleted of its 5' overhang.[14d] These three ODNs bind NC with high affinity and 1:1 stoichiometry, and their structures in complex with NC are known.[14b, 14d, 14f] The results showed that the response of this amino acid probe allows site-selective characterization of peptide-ODN proximity.

To obtain the L-amino acid bearing the 3HC fluorophore, we started with L-tyrosine that was first acetylated using a described procedure[16] based on Fries reaction (Scheme 1). Then, the amino group was protected with the Boc group, and the product was condensed with furaldehyde in the presence of base. The obtained chalcone was further converted into a 3HC derivative using Algar-Flynn-Oyamada reaction. Then, the Boc group was removed to obtain the desired amino acid (3HCaa). Finally, it was protected with the Fmoc group, which is required for solid phase peptide synthesis. We checked that the basic conditions used in the synthesis do not alter the optical purity of the obtained amino acid. For this purpose, the obtained Boc-protected amino acid was first benzylated, then the amino group was deprotected and finally modified with (R)- and (S)-Mosher agents (Scheme S2, Supporting Information). The ^{19}F-NMR spectra of both products showed only one fluorine peak corresponding to pure diastereomers, indicating that the synthesis did not produce any racemization. This conclusion was confirmed by ^{1}H-NMR spectra of these two diastereomers.



(a) AcCl, $AlCl_3$, $PhNO_2$; (b) Boc_2O
(c) Furfural, NaOH, EtOH/H_2O; (d) H_2O_2, NaOH, EtOH
(e) HCl, dioxane; (f) FmocOSu, $NaHCO_3$, MeCN/H_2O

Scheme 1. Scheme of synthesis of the fluorescent amino acid and the labeled peptide. The substituted positions of the NC(11-55) peptide are shown by blue arrows.

Then, the fluorescence properties of the new amino acid in the form of Boc-protected 3HCaa were characterized in different solvents. The dual emission of the new amino acid was found to depend strongly on solvent polarity (Figure S2, Table S1, Supporting Information), showing an increase of the relative intensity of the N* band in more polar solvents due to an inhibition of the ESIPT reaction.[9h]

Using the Fmoc-protected 3HCaa, we synthesized three labeled NC(11-55) peptides: NC(11-55)-F16-3HCaa, NC(11-55)-A30-3HCaa and NC(11-55)-W37-3HCaa, where the fluorescent amino acid was substituted for Phe16, Ala30 and Trp37 residues, located in the first zinc finger, the linker and the second zinc finger, respectively (Scheme 1). Importantly, the 3HCaa was compatible with the standard protocols of peptide synthesis and purification. The NOESY spectra show that the internal structure of the zinc fingers is maintained in the three labeled NC(11-55) peptides (Figure S3, Supporting Information), since similar medium and long distance NOEs were found as for the native peptide. Thus, the incorporated amino acid analogue 3HCaa does not significantly disturb the native folding in the three labeled peptides.

Next, we tested the activity of the labeled peptides, by comparing their ability to promote the annealing of two complementary DNA strands with that of the native peptide. Remarkably, all three labeled peptides exhibited annealing-promoting activities[17] comparable to that of the native peptide (Figure S4, Supporting Information), indicating that the nucleic acid chaperone activity is preserved in the labeled peptides.

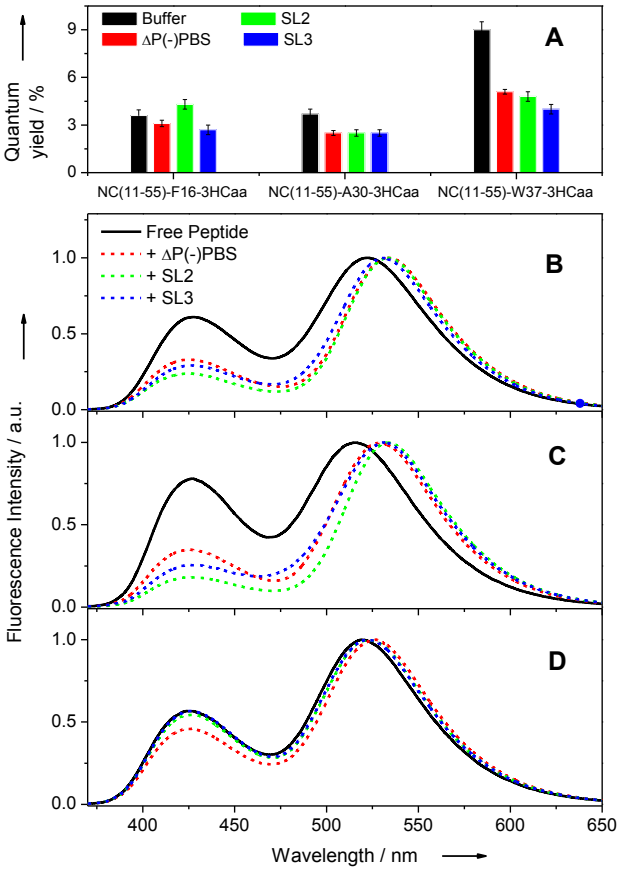


Figure 2. Effect of ODN interaction on the fluorescence quantum yields (A) and spectra (B-D) of the 3HCaa-labeled NC(11-55) peptides. Fluorescence spectra of NC(11-55)-W37-3HCaa (B), NC(11-55)-F16-3HCaa (C), NC(11-55)-A30-3HCaa (D) were normalized at the T* band. Peptide and ODN concentrations were 1 μM. Measurements were done in 10 mM phosphate buffer, 30 mM NaCl, pH 6.5. Excitation wavelength was 350 nm.

Labeled peptides in aqueous solutions (Figure 2) showed dual emission, characteristic of 3HC dyes, where the short- and long-wavelength bands can be assigned to N* and T* forms, respectively. The observed intensity ratio of the two emission bands, N*/T*, is close to that for 3HCaa in methanol (Tables S1 and S2, Supporting Information), indicating that the 3HCaa fluorophore is partially screened from water by the peptide.

On interaction with SL2, SL3 and ΔP(-)PBS, a strong change in the 3HCaa fluorescence spectra was observed, with notably a large drop in the N*/T* ratio for the NC(11-55)-W37-3HCaa and

NC(11-55)-F16-3HCaa peptides and a decrease in the fluorescence quantum yield for the NC(11-55)-W37-3HCaa peptide (Figure 2). The drop in the N*/T* ratio can be assigned to a decrease in the local polarity of the 3HCaa environment, likely connected to the stacking of the 3HC fluorophore with the ODN bases, which leads to an efficient screening of this fluorophore from water.[11] This conclusion is in line with the NMR-derived 3D structures of the three NC-ODN complexes (Figure 3), showing that both Phe16 and Trp37 residues strongly interact with the ODN bases.[14b, 14d, 14f] In contrast, binding of NC(11-55)-A30-3HCaa to the three ODNs led to a slight decrease in its quantum yield with either no change (with SL2 and SL3) or a limited decrease (with ΔP(-)PBS) in the N*/T* ratio. The absence of changes in the N*/T* ratio observed with the SL2 and SL3 sequences can be rationalized by the large distance between the A30 residue and the ODN bases in the NC-ODN complexes (Figure 3). As a consequence, the 3HCaa at position 30 in the complexes with SL2 and SL3 can not interact with the ODN and remains highly exposed to the buffer, as in the free peptide. Finally, the moderate change in the N*/T* ratio observed when NC(11-55)-A30-3HCaa was complexed with ΔP(-)PBS is fully consistent with the closer proximity of the Ala30 residue to the ΔP(-)PBS bases, as compared to that in the complexes with SL2 and SL3.[14b, 14d]

From the available NMR data, we estimated the distances between the substituted positions (from their β-carbon atom) and the closest ODN base in the complexes of NC with SL2, SL3 and ΔP(-)PBS. We found that these distances correlate with the observed values of the N*/T* ratio for all three peptide mutants (Figure 4), likely as a consequence of the level of stacking of the introduced fluorescent amino acid with the ODN bases that modulates the dual emission of the 3HC fluorophore. Thus, by introducing 3HCaa at different positions of a peptide, the proximity of the labeled position with the closest ODN base can be evaluated from the N*/T* ratio value. This approach appears thus as a new method for site-specific monitoring of peptide-ODN proximity. However, for very short distances (< 5 Å) the correlation is less clear, probably because the size, orientation and dynamics of the fluorophore and the nucleic bases start playing an important role, providing less predictable N*/T* ratio values.

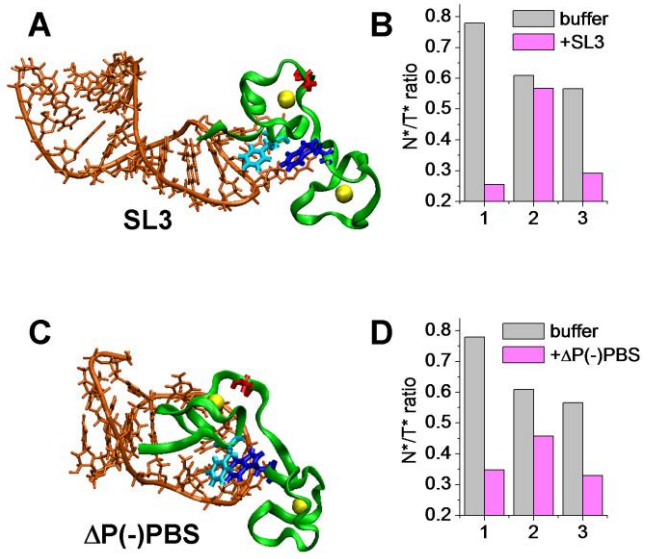


Figure 3. Comparison of 3D structures of NC-ODN complexes with the response of the 3HCaa-labeled peptides on ODN binding. 3D structures of NC complexes with SL3 (A) and ΔP(-)PBS (C), drawn based on NMR data.[14b, 14d] Only the NC(11-55) part is shown on the figures. Zn atoms are represented as yellow spheres, Phe16, Ala30 and Trp37 residues are in cyan, red, and blue, respectively. Fluorescence intensity ratio N*/T* for NC(11-55)-F16-3HCaa (1), NC(11-55)-A30-3HCaa (2), NC(11-55)-W37-3HCaa (3) peptides in buffer and bound to SL3 (B) or ΔP(-)PBS (D).

Finally, using the changes in the N*/T* ratio or the quantum yield of 3HCaa as an analytical signal of interaction, we further measured the binding constant of the three labeled peptides to SL3. The obtained binding constants matched closely that obtained with the native peptide (Table S3, Supporting Information),[14c] indicating that the incorporated 3HCaa probe did not change the binding affinity of the peptides and could thus be also used for quantification of peptide-ODN interaction.

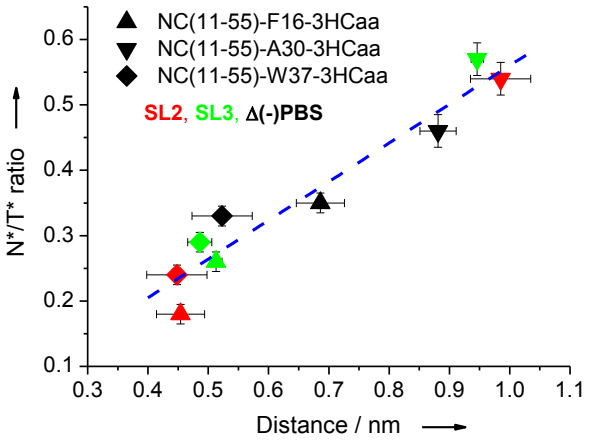


Figure 4. Correlation between the fluorescence response of the 3HCaa probe and its distance to the closest ODN base in the NC-ODN complexes. The fluorescence intensity ratio N*/T* of NC(11-55)-W37-3HCaa (◆), NC(11-55)-F16-3HCaa (▲) and NC(11-55)-A30-3HCaa (▼) is plotted as a function of the distance from the beta-carbon of the corresponding substituted amino acid in these peptides and the nearest heterocycle atom of the closest base in the complexes with SL2 (red), SL3 (green) and ΔP(-)PBS (black).

In the present work, we synthesized an enantiomerically pure fluorescent L-amino acid analogue based on 2-(2-furyl)-3-hydroxychromone dye. This amino acid was incorporated at three different positions of the NC(11-55) peptide without affecting its native folding and activity. Interaction of the labeled peptides with target ODNs changed dramatically the dual emission of the introduced amino acid and these changes depended on its position in the peptide and on the ODN sequence. The obtained results established a correlation between the fluorescence response of the amino acid probe and its proximity to the interacting ODN bases. This amino acid appears as a universal tool for site-selective proximity measurements in peptide-nucleic acid complexes.

Received: ((will be filled in by the editorial staff))
Published online on ((will be filled in by the editorial staff))

Keywords: Environment-sensitive dyes · excited state proton transfer · fluorescent probes · peptide-nucleic acid interactions · peptides

[1] a) P. K. Mikhailiuk, S. Afonin, A. N. Chernega, E. B. Rusanov, M. O. Platonov, G. G. Dubinina, M. Berditsch, A. S. Uirich, I. V. Komarov, *Angew. Chem. Int. Ed.* **2006**, *45*, 5659; b) S. Mizukami, R. Takikawa, F. Sugihara, M. Shirakawa, K. Kikuchi, *Angew. Chem. Int. Ed.* **2009**, *48*, 3641; c) M. Ieronimo, S. Afonin, K. Koch, M. Berditsch, P. Wadhwani, A. S. Ulrich, *J. Am. Chem. Soc.* **2010**, *132*, 8822 ; d) J. de Villiers, L. Koekemoer, E. Strauss, *Chem. Eur. J.* **2010**, *16*, 10030.

[2] a) B. E. Cohen, T. B. McAnaney, E. S. Park, Y. N. Jan, S. G. Boxer, L. Y. Jan, *Science* **2002**, *296*, 1700; b) M. P. Brun, L. Bischoff, C. Garbay, *Angew. Chem. Int. Ed.* **2004**, *43*, 3432; c) P. Venkatraman, T. T. Nguyen, M. Sainlos, O. Bilsel, S. Chitta, B. Imperiali, L. J. Stern, *Nat. Chem. Biol.* **2007**, *3*, 222.

[3] M. Nitz, A. R. Mezo, M. H. Ali, B. Imperiali, *Chem Commun (Camb)* **2002**, 1912.
[4] H. Chen, N. N. Chung, C. Lemieux, B. Zelent, J. M. Vanderkooi, I. Gryczynski, B. C. Wilkes, P. W. Schiller, *Biopolymers* **2005**, *80*, 325.
[5] M. E. Vazquez, D. M. Rothman, B. Imperiali, *Org. Biomol. Chem.* **2004**, *2*, 1965.
[6] M. E. Vazquez, J. B. Blanco, B. Imperiali, *J Am Chem Soc* **2005**, *127*, 1300.
[7] G. Loving, B. Imperiali, *J Am Chem Soc* **2008**, *130*, 13630.
[8] P. K. Sengupta, M. Kasha, *Chem. Phys. Lett.* **1979**, *68*, 382.
[9] a) R. Das, A. S. Klymchenko, G. Duportail, Y. Mely, *Photochem. Photobiol. Sci.* **2009**, *8*, 1583; b) V. V. Shynkar, A. S. Klymchenko, E. Piemont, A. P. Demchenko, Y. Mely, *J. Phys. Chem. A* **2004**, *108*, 8151; c) K. Enander, L. Choulier, A. L. Olsson, D. A. Yushchenko, D. Kanmert, A. S. Klymchenko, A. P. Demchenko, Y. Mely, D. Altschuh, *Bioconjugate Chem.* **2008**, *19*, 1864; d) P.-T. Chou, M. L. Martinez, H. Clements, *J. Phys. Chem.* **1993**, *97*, 2618; e) S. M. Ormson, R. G. Brown, F. Vollmer, W. Rettig, *J. Photochem. Photobiol., A* **1994**, *81*, 65; f) T. C. Swinney, D. F. Kelley, *J. Chem. Phys.* **1993**, *99*, 211; g) A. S. Klymchenko, A. P. Demchenko, *Phys. Chem. Chem. Phys.* **2003**, *5*, 461; h) A. P. Demchenko, Y. Mely, G. Duportail, A. S. Klymchenko, *Biophys. J.* **2009**, *96*, 3461; i) A. J. G. Strandjord, P. F. Barbara, *J. Phys. Chem.* **1985**, *89*, 2355; j) D. McMorrow, M. Kasha, *J. Phys. Chem.* **1984**, *88*, 2235.
[10] a) A. S. Klymchenko, A. P. Demchenko, *New J. Chem.* **2004**, *28*, 687; b) D. A. Yushchenko, J. A. Fauerbach, S. Thirunavukkuarasu, E. A. Jares-Erijman, T. M. Jovin, *J. Am. Chem. Soc.* **2010**, *132*, 7860.
[11] a) A. S. Klymchenko, V. V. Shvadchak, D. A. Yushchenko, N. Jain, Y. Mely, *J. Phys. Chem. B* **2008**, *112*, 12050; b) V. V. Shvadchak, A. S. Klymchenko, H. De Rocquigny, Y. Mely, *Nucleic Acids Res.* **2009**, *37*, e25.
[12] a) J. A. Thomas, R. J. Gorelick, *Virus Res.* **2008**, *134*, 39; b) J. L. Darlix, J. Godet, R. Ivanyi-Nagy, F. P., M. O., Y. Mely, *J Mol Biol* **2011**, in press.
[13] a) M. F. Summers, L. E. Henderson, M. R. Chance, J. W. Bess Jr, T. L. South, P. R. Blake, I. Sagi, G. Perez-Alvarado, R. C. Sowder Iii, D. R. Hare, L. O. Arthur, *Protein Sci.* **1992**, *1*, 563; b) N. Morellet, N. Jullian, H. De Rocquigny, B. Maigret, J. L. Darlix, B. P. Roques, *EMBO J.* **1992**, *11*, 3059.
[14] a) J. Clever, C. Sassetti, T. G. Parslow, *J. Virol.* **1995**, *69*, 2101 ; b) R. N. De Guzman, Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer, M. F. Summers, *Science* **1998**, *279*, 384 ; c) C. Vuilleumier, E. Bombarda, N. Morellet, D. Gerard, B. P. Roques, Y. Mely, *Biochemistry* **1999**, *38*, 16816 ; d) S. Bourbigot, N. Ramalanjaona, C. Boudier, G. F. J. Salgado, B. P. Roques, Y. Mely, S. Bouaziz, N. Morellet, *J. Mol. Biol.* **2008**, *383*, 1112; e) M. F. Shubsda, A. C. Paoletti, B. S. Hudson, P. N. Borer, *Biochemistry* **2002**, *41*, 5276; f) G. K. Amarasinghe, R. N. De Guzman, R. B. Turner, K. J. Chancellor, Z. R. Wu, M. F. Summers, *J. Mol. Biol.* **2000**, *301*, 491.
[15] a) R. J. Fisher, A. Rein, M. Fivash, M. A. Urbaneja, J. R. Casas-Finet, M. Medaglia, L. E. Henderson, *J. Virol.* **1998**, *72*, 1902 ; b) H. Beltz, C. Clauss, E. Piemont, D. Ficheux, R. J. Gorelick, B. Roques, C. Gabus, J. L. Darlix, H. de Rocquigny, Y. Mely, *J. Mol. Biol.* **2005**, *348*, 1113; c) Y. Mely, H. de Rocquigny, M. Sorinasjimeno, G. Keith, B. P. Roques, R. Marquet, D. Gerard, *J. Biol. Chem.* **1995**, *270*, 1650.
[16] C. D. Hufford, B. O. Oguntimein, J. N. Shoolery, *J. Org. Chem.* **1987**, *52*, 5286.
[17] a) J. Godet, H. de Rocquigny, C. Raja, N. Glasser, D. Ficheux, J. L. Darlix, Y. Mely, *J. Mol. Biol.* **2006**, *356*, 1180; b) J. Godet, Y. Mely, *RNA Biol* **2010**, *7*, 687.

Supporting information

Two-Color Fluorescent Amino Acid for Proximity Measurements in Peptide-Nucleic Acid Complexes

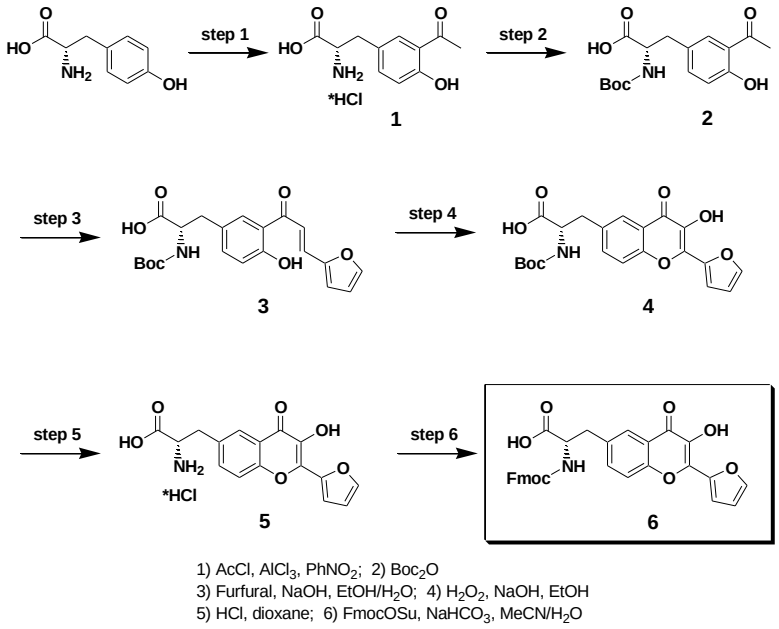
Aleksandr V. Strizhak, Viktoriia Y. Postupalenko, Volodymyr V. Shvadchak, Nelly Morellet, Eric Guittet, Vasyl G. Pivovarenko, Andrey S. Klymchenko, Yves Mély

Table of contents

Materials and methods

All the solvents and chemicals were purchased from Sigma-Aldrich Chemical Company. For absorption and fluorescence studies, the solvents were of spectroscopic grade. ΔP(-)PBS DNA, SL2 and SL3 RNAs were synthesized and HPLC-purified by IBA GmbH (Germany). Concentrations of the ODNs were calculated from their absorbance using the molar extinction coefficients at 260 nm specified by the supplier.
NMR spectra were recorded on a 400 MHz and 500 MHz Bruker Avance III, BBFO+ at room temperature. Mass spectra were obtained on a Bruker HCT Ultra and Agilent Technologies Accurate-Mass Q-TOF LC/MS 6520 mass spectrometers.

Synthesis of Fmoc-3HCaa



Scheme S1. Scheme of the synthesis of the fluorescent amino acid.

The synthesis of the compounds **1** and **2** (Scheme S1) was performed as described previously.[1]

*N-(tert-butoxycarbonyl)-3-[(2E)-3-(2-furyl)prop-2-enoyl]-L-tyrosine (**3**).* 6.425 g (19.89 mmol) of 3-acetyl-N-(*tert*-butoxycarbonyl)-L-tyrosine **2** were dissolved in 48 ml of ethanol upon stirring under Ar-atmosphere. To the reaction mixture, 48 ml of degassed 25% solution of sodium hydroxide was added upon cooling in cold water bath. Then, 2.409 g of furfural (25.09 mmol, 1.26 eq.) were added and the mixture was stirred for 12 h at RT under Ar. Additional 0.5 g of furfural was added and the mixture was stirred overnight. The reaction mixture was acidified with HCl and diluted twice with water. The formed precipitate was filtered and washed with diluted ethanol and water, dried and used in the next step without further purification. Yield 7.62 g (95%) of chalcone **3** as a yellow powder. **^{1}H-NMR** (500 MHz, MeOD) δ 1.33 (s, 9H), 2.87-2.92 (dd, *J* = 13.9 Hz, *J* = 8.85 Hz, 1H), 3.16-3.19 (dd, *J* = 13.9 Hz, *J* = 4.43 Hz, 1H), 4.31-4.34 (dd, *J* = 8.85 Hz, *J* = 4.43 Hz, 1H), 6.61-6.62 (dd, *J* = 3.16 Hz, *J* = 1.89 Hz, 1H), 6.87-6.89 (d, *J* = 8.21 Hz, 1H), 6.94 (d, *J* = 3.8 Hz, 1H), 7.39-7.41 (dd, *J* = 8.22 Hz, *J* = 1.9 Hz, 1H), 7.64-7.71 (dd, *J* = 15.16 Hz, *J* = 4.42 Hz, 2H), 7.71 (d, *J* = 1.9 Hz, 1H), 7.86-7.87 (d, *J* = 1.9 Hz, 1H). **^{13}C-NMR** (500 MHz, MeOD) δ 28.66, 38.11, 55.49, 80.47, 114.01, 118.33, 118.84, 119.02, 120.97, 129.45, 131.65, 132.31, 138.58, 147.17, 152.98, 157.68, 163.09, 175.69, 194.74 (two carbon signals are absent due to incidental equivalence). **HRMS** (EI) *m/z* calcd for $C_{21}H_{23}NO_7$ 401.14745, found: 401.147.

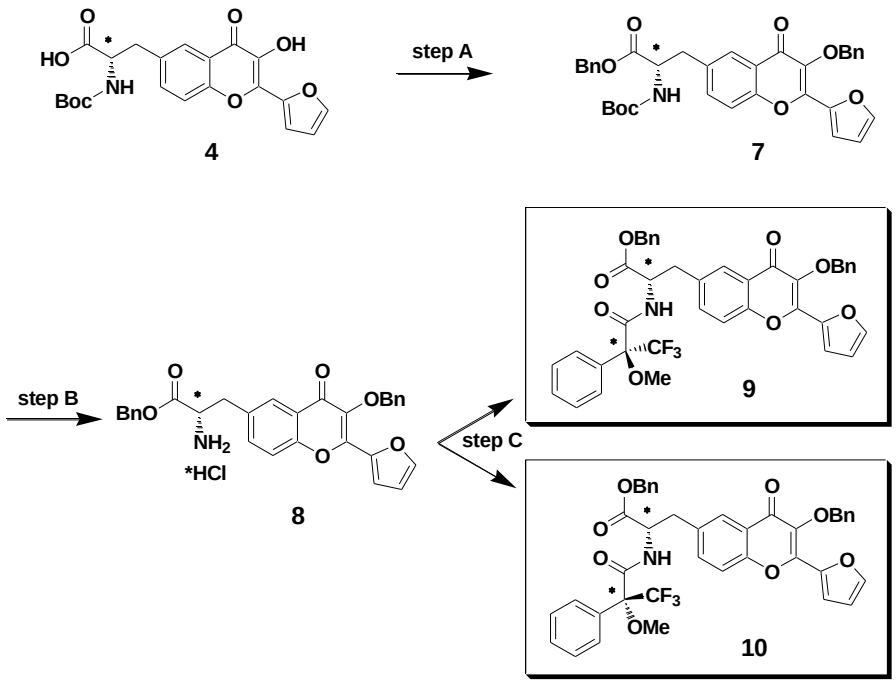
*N-(tert-butoxycarbonyl)-3-[2-(2-furyl)-3-hydroxy-4-oxo-4H-chromen-6-yl]-L-alanine (**4**).* 7.62 g (19 mmol) of chalcone **3** was dissolved in 75 ml of ethanol, cooled in ice bath. Then, 75 ml of 1M solution of sodium hydroxide was added with stirring. Next, 4.73 ml (41.8 mmol) of 30% hydrogen peroxide was added and the mixture was stirred in ice bath for 4 h. The reaction was monitored by TLC (silica, EtOAc/MeOH 9:1). After completing the reaction, the mixture was acidified and the formed precipitate was filtered. Yield of the crude 3-hydroxychromone **4** was 3.56 g (45%). It was purified to about 95% purity by crystallization several times from toluene and then from acetonitrile. Final yield ~ 1.5 g (20%) of yellow crystals. **^{1}H-NMR** (500 MHz, MeOD) δ 1.33 (s, 9H), 2.98-3.03 (dd, *J* = 13.9 Hz, *J* = 9.48 Hz, 1H), 3.26-3.30 (m, *J* = 5.05 Hz, 1H), 4.36-4.39 (dd, *J* = 9.48 Hz, *J* = 5.05 Hz, 1H), 6.67 (m, 1H), 7.33 (d, *J* = 3.16 Hz, 1H), 7.51-7.53 (d, *J* = 8.85 Hz, 1H), 7.60-7.62 (d, *J* = 8.85 Hz, 1H), 7.78 (s, 1H), 7.97 (d, *J* = 1.27 Hz, 1H). **^{13}C-NMR** (500 MHz, MeOD) δ 28.6, 38.27, 56.26, 80.52, 113.58, 116.91, 119.29, 122.83, 126.26, 135.77, 135.97, 138.44, 141.29, 145.86, 146.08, 155.09, 157.77, 173.81, 175.11 (two carbon signals are absent due to incidental equivalence). **HRMS** (EI) *m/z* calcd for $C_{21}H_{21}NO_8$ 415.12672, found: 415.12659.

*3-[2-(2-furyl)-3-hydroxy-4-oxo-4H-chromen-6-yl]-L-alanine hydrochloride (**5**).* 944 mg of 3-hydroxychromone **4** was dissolved in 10 ml of dioxane. Then, 10 ml of conc. HCl was added to the reaction mixture in cold water bath. After 5-10 min, the bath was removed and the mixture was stirred for 2-3 h. The reaction mixture was concentrated under reduced pressure to give the desired product **5** as a dark yellow powder, yield 0.769 g (96%). **^{1}H-NMR** (500 MHz, MeOD) δ 3.28-3.34 (dd, *J* = 14.56 Hz, *J* = 7.78 Hz, 1H), 3.42-3.48 (dd, *J* = 14.56 Hz, *J* = 5.77 Hz, 1H), 4.34-4.37 (dd, *J* = 7.53 Hz, *J* = 5.77 Hz, 1H), 6.7-6.72 (dd, *J* = 3.52 Hz, *J* = 1.76 Hz, 1H), 7.37-7.38 (d, *J* = 3.51 Hz, 1H), 7.63-7.65 (d, *J* = 8.79 Hz, 1H), 7.68-7.71 (dd, *J* = 8.79 Hz, *J* = 2.26 Hz, 1H), 7.81-7.82 (d, *J* = 1.00 Hz, 1H), 8.05-8.06 (d, *J* = 1.76 Hz, 1H). **^{13}C-NMR** (500 MHz, MeOD) δ 36.74, 54.95, 113.70, 117.23, 120.29, 123.34, 126.81, 132.56, 135.79, 138.64, 141.67, 145.83, 146.33, 155.71, 171.00, 173.74. **HRMS** (EI) *m/z* calcd for $C_{16}H_{13}NO_6$ 315.07429, found: 315.07419.

*N-[(9H-fluoren-9-ylmethoxy)carbonyl]-3-[2-(2-furyl)-3-hydroxy-4-oxo-4H-chromen-6-yl]-L-alanine (**6**).* 614 mg (1.744 mmol) of amino acid hydrochloride **5** was dissolved in a solution of 500 mg (5.95 mmol) of sodium bicarbonate in 10 ml of water. Then, 30 ml of acetonitrile and 588 mg (1 eq.) of FmocOSu were added and the mixture was stirred for 24 h. The reaction mixture was filtered and the precipitate was dissolved in hot water (about 100 ml per 1 g) with further acidification with HCl. Yield 900 g (96%) of yellow powder. **^{1}H-NMR** (500 MHz, DMSO-d_6) δ 3.00-3.05 (m, 1H), 3.22-3.26 (dd, *J* = 14.04 Hz, *J* = 4.27 Hz, 1H), 4.12-4.19 (m, 3H), 4.26-4.30 (m, 1H), 6.77-6.78 (dd, *J* = 3.05 Hz, *J* = 1.83 Hz, 1H), 7.21-7.25 (m, 2H), 7.27-7.28 (d, *J* = 3.05 Hz, 1H), 7.34-7.37 (t, *J* = 7.32 Hz, 2H), 7.57-7.58 (d, *J* = 7.32 Hz, 2H), 7.59-7.61 (d, *J* = 8.55 Hz, 1H), 7.68-7.70 (d, *J* = 8.55 Hz, 1H), 7.75-7.76 (d, *J* = 8.55 Hz, 1H), 7.82-7.83 (d, *J* = 7.32 Hz, 2H), 8.01 (s, 1H), 8.04 (s, 1H), 9.88 (s, 1H), 12.76 (br s, 1H). **^{13}C-NMR** (500 MHz, DMSO-d_6) δ 35.80, 46.54, 55.39, 65.64, 112.82, 115.27, 118.03, 120.04, 121.56, 124.95, 125.11, 125.18, 126.98, 127.55, 134.68, 137.20, 139.19, 140.63, 143.67, 144.09, 145.17, 152.96, 156.01, 171.78, 173.07 (six carbon signals are absent due to incidental equivalence). **HRMS** (EI) *m/z* calcd for $C_{31}H_{23}NO_8$ 537.14237, found: 537.14187.

Proof of optical purity

To test the enantiomeric purity of the obtained chromone-amino acid derivative, we prepared the dibenzyl ether-ester of amino acid (step A, B). Then, the obtained aminoester was acylated by two enantiomers of Mosher's acid chloride (step C). The NMR spectra of the diastereomeric Mosher amides were then compared. No signal of the opposite diastereomer was found in each spectrum.



A) BnBr, K_2CO_3, $(C_4H_9)_4NI$, DMF; B) conc. HCl, dioxane; C) Mosher, $EtN(i\text{-}Pr)_2$, THF

Scheme S2. Scheme of synthesis of the diastereomeric Mosher amides to prove the optical purity of the synthesized fluorescent amino acid.

*Benzyl 3-[3-(benzyloxy)-2-(2-furyl)-4-oxo-4H-chromen-6-yl]-N-(tert-butoxycarbonyl)-L-alaninate (**7**).* Benzylation was performed in soft conditions to prevent possible racemisation.[1b] 600 mg of Boc-protected amino acid **4** was dissolved in 8 ml of DMF, followed by the addition of 500 mg of benzyl bromide, 438 mg of anhydrous potassium carbonate and 54 mg of tetrabutylammonium iodide. The mixture was stirred at RT for 4 h, and controlled by TLC (EtOAc or EtOAc:heptane 1:1). The mixture was diluted with water and the product was extracted 3 times with EtOAc. The combined organic layers were dried with Na_2SO_4 and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc:heptane 1:1) to give the desired product **7** as a yellow oil.

*Benzyl 3-[3-(benzyloxy)-2-(2-furyl)-4-oxo-4H-chromen-6-yl]-L-alaninate (**8**).* Boc deprotection was performed by dissolving Boc-dibenzylderivate **7** in dioxane (1 ml per 100 mg), adding conc. HCl (1ml per 100 mg) and stirring during ~ 2 h. The reaction was controlled by TLC (EtOAc:Heptane 1:1). The solution was concentrated under reduced pressure and the residue was recrystallized from MeOH/ether. Pale yellow solid, yield ~ 90%.

*Mosher conjugates (**9**, **10**).* 33 mg of dried chlorohydrate **8** was suspended in 1 ml of THF. Then, 26 mg of Hunig's base and ~ 20 mg of Mosher's chloride were added. The mixture was stirred for 24 h, diluted with water, and stirred for 1 h to hydrolyze the excess of Mosher's chloride. The reaction mixture was then extracted with dichloromethane. The organic layer was washed with bicarbonate solution followed by 1M HCl, brine, dried with Na_2SO_4 and concentrated under reduced pressure to give the desired products **9** and **10** as yellow oil.

Benzyl 3-[3-(benzyloxy)-2-(2-furyl)-4-oxo-4H-chromen-6-yl]-N-[(2S)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoyl]-L-alaninate (9). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ 3.14-3.19 (dd, *J* = 14.24 Hz, *J* = 7.12 Hz, 1H), 3.20 (s, 3H), 3.24-3.29 (dd, *J* = 14.24 Hz, *J* = 5.60 Hz, 1H), 4.89-4.94 (dd, *J* = 14.25 Hz, *J* = 6.62 Hz, 1H), 5.05-5.08 (d, *J* = 12.21 Hz, 1H), 5.11-5.14 (d, *J* = 12.21 Hz, 1H), 5.24 (s, 2H), 6.48-6.49 (dd, *J* = 3.56 Hz, *J* = 1.52 Hz, 1H), 7.18-7.44 (m, 17H), 7.59 (d, *J* = 1.01 Hz, 1H), 7.96 (d, *J* = 1.53 Hz, 1H). **^{19}F-NMR** (376 MHz, $CDCl_3$) δ -69.5 (**9**), -72.03 (Mosher acid).

Benzyl 3-[3-(benzyloxy)-2-(2-furyl)-4-oxo-4H-chromen-6-yl]-N-[(2R)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoyl]-L-alaninate (10). **^{1}H-NMR** (400 MHz, $CDCl_3$) δ 3.02-3.07 (dd, *J* = 14.05 Hz, *J* = 7.53 Hz, 1H), 3.16-3.21 (dd, *J* = 14.05 Hz, *J* = 5.52 Hz, 1H), 3.35 (d, *J* = 1.50 Hz, 3H), 4.94-5.00 (m, *J* = 5.52 Hz, *J* = 7.28 Hz, 1H), 5.07-5.10 (d, *J* = 12.05 Hz, 1H), 5.14-5.17 (d, *J* = 12.05 Hz, 1H), 5.20-5.23 (d, *J* = 10.54 Hz, 1H), 5.23-5.26 (d, *J* = 10.54 Hz, 1H), 6.49-6.51 (dd, *J* = 3.52 Hz, *J* = 1.76 Hz, 1H), 7.01-7.04 (dd, *J* = 8.79 Hz, *J* = 2.26 Hz, 1H), 7.05-7.07 (d, *J* = 8.53 Hz, 1H), 7.17-7.32 (m, 15H), 7.42-7.45 (dd, *J* = 8.29 Hz, *J* = 1.76 Hz, 1H), 7.60 (dd, *J* = 1.76 Hz *J* = 0.75 Hz, , 1H), 7.84 (d, *J* = 2.26 Hz, 1H). **^{19}F-NMR** (376 MHz, $CDCl_3$) δ -69.41(**10**), -72.03 (Mosher acid).

Peptide synthesis

The NC(11-55) peptide was synthesized by solid phase peptide synthesis on a 433A synthesizer (ABI, Foster City, CA) as previously described.[2] The synthesis was performed at a 0.1 mmol scale using the standard fluorenylmethoxycarbonyl (Fmoc)-amino acid-coupling protocol starting from 0.44 mmol/g Wang LL resin. At selected positions of the peptide, the fluorescent amino acid analogue was incorporated by the following procedure. In a flask, 2-4 mole equivalents of the Fmoc-3HCaa were mixed with 4 eq. of HBTU/HOBt coupling solution (in DMF) and 5 eq. of DIEA. This mixture was immediately added to the peptidylresin and shaked at 37°C for 12 h. Then, the resin was washed with 1-methyl-2-pyrrolidone (NMP) and peptide synthesis was continued on the synthesizer. At the end of the synthesis, the Fmoc-deprotected peptidylresin was isolated and washed with NMP, methanol and dichloromethane.
Cleavage of the peptidylresin and deprotection were performed for 2 h using a 10 ml trifluoroacetic acid (TFA) solution containing water (5%, v/v), phenol (2%, w/v), thioanisole (5%, v/v), triisopropylsilane (2.5%, v/v) and ethanedithiol (2.5%, v/v). The peptide was precipitated using ice-cold diethyl ether and pelleted by centrifugation. The pellet was air-dried for approx. 15 minutes, solubilized with aqueous TFA (0.05 %, v/v) and lyophilized. Purification by HPLC was carried out on a C8 column (uptisphere 300A, 5µm; 250X10, Interchim, France) in water/acetonitrile mixture containing 0.05% TFA with a linear gradient 10 to 35% of acetonitrile for 30 min and monitored at 210 and 360 nm (3HC dye absorption). Obtained peptides were characterized by ESI-MS analysis. NC(11-55)-W37-3HCaa peptide: calculated M = 5245.40, found $[M+7H]^{7+}$ = 750.64 corresponding to M = 5245.38 after deconvolution; NC(11-55)-A30-3HCaa peptide: calculated M = 5360.43, found $[M+7H]^{7+}$ = 767.06 corresponding to M = 5360.37 after deconvolution; NC(11-55)-F16-3HCaa peptide: calculated M = 5287.81, found $[M+7H]^{7+}$ = 756.30 corresponding to M = 5287.03 after deconvolution.

Preparation of Zn-bound peptides

Lyophilized labeled peptides were dissolved in water (≈ 0.5 mg in 500 µL). Then, about 10 µL of this solution was used to determine the peptide concentration using an extinction coefficient of 15,000 $M^{-1}\times cm^{-1}$ at 350 nm. Next, 2.2 molar equivalents of $ZnSO_4$ were added to the peptide and pH was raised to its final value, by adding buffer. This last step was done only at the end to prevent peptide oxidization. Noticeably, a large excess of Zn^{2+} ions should be avoided since this ion could affect the 3HC fluorescence.

Spectroscopic measurements

Unless otherwise indicated, the experiments were performed in 10 mM phosphate buffer, pH 6.5, 30 mM NaCl, at 20°C. This low pH was used to prevent deprotonation of the 3-hydroxy group of the 3HC amino acid. Absorption spectra were recorded with a Cary 4000 UV-visible spectrophotometer (Varian). Fluorescence spectra were recorded on FluoroMax3 and FluoroLog spectrofluorimeters (Jobin Yvon) equipped with thermostated cell compartments. Fluorescence spectra were corrected for Raman scattering. Quantum yields were calculated using quinine sulphate in 0.5 M sulphuric acid (quantum yield, φ = 0.577) as a reference.[3] Excitation wavelength was 350 nm for the 3HCaa.

To determine the affinity of the labeled peptides for the ODNs, fixed amounts of the ODN were titrated with peptides by monitoring the two-band fluorescence of the labeled peptides. For each data point, the emission of the same concentration of labeled peptide in buffer was subtracted from the signal measured in the presence of the ODN. Affinity constants were determined from direct fitting of the corrected signal to the rewritten Scatchard equation:

$$I = I_0 - \frac{(I_0 - I_t)}{N_t} x \frac{(1+(P_t + nN_t)K_a) - \sqrt{(1+(P_t + nN_t)K_a)^2 - 4P_t nN_t K_a^2}}{2nK_a} \qquad (1)$$

where I and I_t are the signal at a given and a saturating peptide concentration, respectively, I_0 is the signal in the absence of peptide, N_t is the total ODN concentration, P_t is the total concentration of peptide, K_a is the apparent affinity constant, n is the number of binding sites. The parameters were recovered from non-linear fits of equation (1) to experimental datasets using the Microcal Origin™ 6.0 software.

To check the influence of the fluorescent amino acid on the NC(11-55) chaperone properties, the ability of the labeled peptides to promote the annealing of cTAR with its complementary dTAR sequence was compared to that of the unlabeled NC peptide.[4] The kinetic measurements were performed under pseudo-first-order conditions by using unlabeled dTAR at a concentration which was 30-fold higher than the concentration of cTAR labeled with carboxytetramethylrhodamine (TMR) at the 5′ end and with 5/6-carboxyfluorescein (Fl) at the 3′ end.[4] Excitation and emission wavelengths were 480 and 520 nm, respectively, for monitoring the Fl fluorescence. All reported concentrations correspond to those after mixing. To avoid high local concentrations during mixing, both reactants were mixed at the same volume. Peptides were added to each reactant separately at a peptide:ODN ratio of 3 :1, and then, the reaction was initiated by mixing the peptide-coated ODNs together. Experiments were performed in 25 mM Tris, 30 mM NaCl, 0.2 mM $MgCl_2$, pH 7.5 at 20 ºC.

Supporting figures

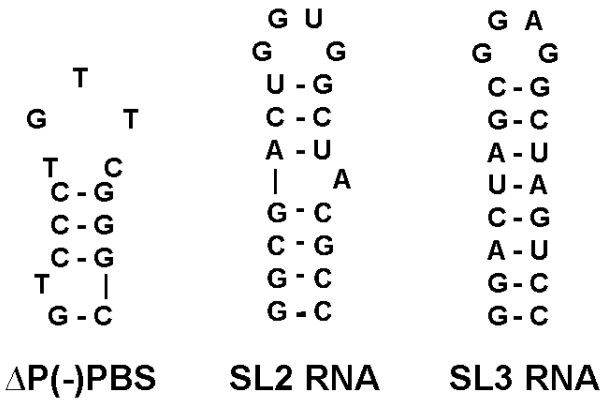


Figure S1. Sequences of the ΔP(-)PBS DNA, SL2 and SL3 RNAs used in this study.

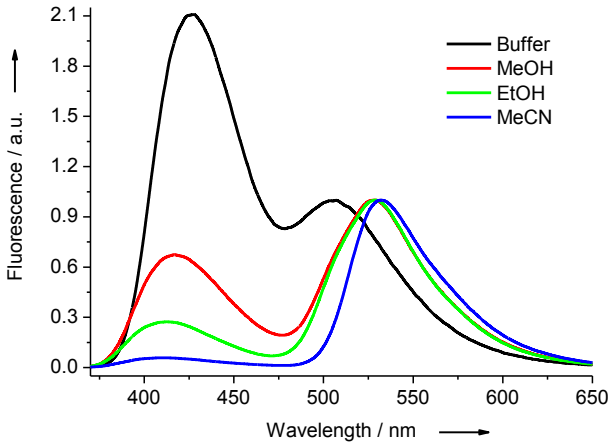


Figure S2. Fluorescence spectra of the Boc-3HCaa derivative in different solvents. The buffer was 10mM phosphate buffer, 30mM NaCl, pH = 6.5. Concentration of Boc-3HCaa was 1 μM. All the spectra were normalized at the T* band. Excitation wavelength was 350 nm.

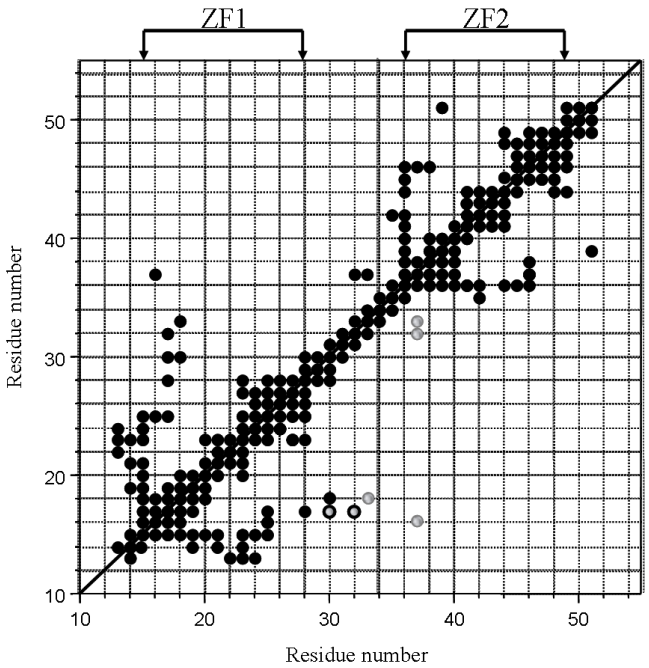


Figure S3. Diagonal plot of side-chain/side-chain Nuclear Overhauser Effects (NOEs) comparing the (13-51) domain of native NC (upper part) and the labeled NC(11-55) derivatives (lower part). ● NOEs seen for all three labeled peptides; ● NOEs seen for none of the three labeled peptides; ○ NOEs seen only for the NC(11-55)-F16-3HCaa and NC(11-55)-W37-3HCaa peptides. Analysis of NOEs shows that the internal structure is preserved for each zinc finger (residues 15 to 28 and 36 to 49 for ZF1 and ZF2, respectively) in the three labeled peptides, since the same NOEs as for the native protein are found in the spectra, resulting to a perfect symmetry for these residues from both sides of the diagonal (black discs). The differences observed for the labeled peptides compared to the native NC (gray disks) are marginal.

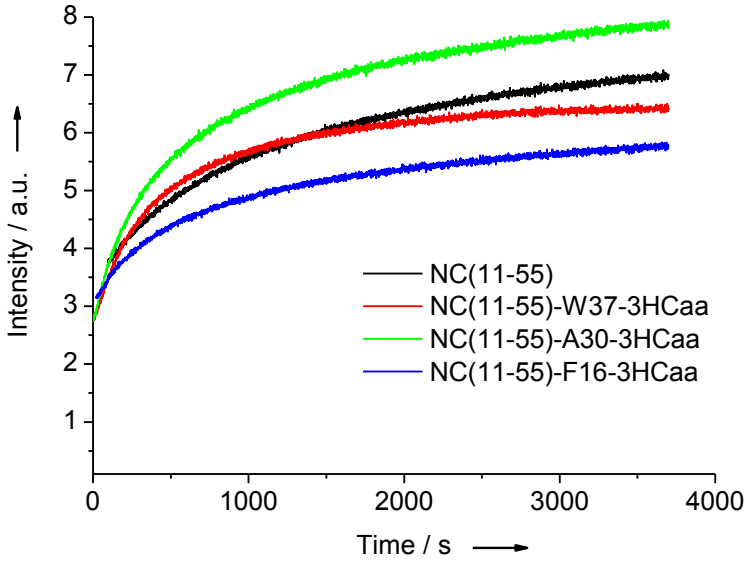


Figure S4. Nucleic acid chaperone activity of the 3HCaa-labeled NC(11-55) peptides. The annealing kinetics of 10 nM TMR-5′-cTAR-3′-Fl with 300 nM dTAR in the presence of the unlabeled and labeled NC(11-55) peptides added at a protein:ODN ratio of 3 :1 were monitored in 25 mM Tris (pH 7.5), 30 mM NaCl and 0.2 mM $MgCl_2$. Excitation and emission wavelengths were 480 and 520 nm, respectively.

Supporting tables

Table S1. Spectroscopic properties of the Boc-3HCaa label.

	λ_{ABS}	λ_{N^*}	λ_{T^*}	N*/T*	QY, %
Buffer	358	427	505	2.18	1.9
MeOH	355	419	529	0.75	3.9
EtOH	356	415	531	0.33	4.6
MeCN	350	412	534	0.07	3.0

λ_{ABS} – position of absorption maxima, λ_{N^*} and λ_{T^*} - position of the fluorescence maxima of N* and T* forms, respectively. N*/T* - ratio of the intensities of the two emission bands at their peak maxima. QY - fluorescence quantum yield. 10 mM phosphate buffer, 30 mM NaCl, pH = 6.5 was used. Excitation wavelength was 350 nm. Quinine sulfate in H_2SO_4 pH=1 (QY = 57.7%) was used as a reference for quantum yield calculation.

Table S2. Spectroscopic properties of the 3HCaa-labeled NC(11-55) peptides in complexes with oligonucleotides.

Peptide	Oligonucleotide	λ_{ABS}	λ_{N^*}	λ_{T^*}	N^*/T^*	QY, %
NC(11-55)-F16-3HCaa	-	364	427	516	0.78	3.6
	+ ΔP(-)PBS	367	425	530	0.35	3.1
	+ SL3	368	427	531	0.26	2.7
	+ SL2	369	427	533	0.18	4.3
NC(11-55)-A30-3HCaa	-	364	426	520	0.57	3.7
	+ ΔP(-)PBS	364	425	526	0.46	2.5
	+ SL3	363	426	523	0.57	2.5
	+ SL2	363	426	523	0.54	2.5
NC(11-55)-W37-3HCaa	-	366	427	522	0.61	9.0
	+ ΔP(-)PBS	367	425	534	0.33	5.1
	+ SL3	368	427	531	0.29	4.0
	+ SL2	368	425	533	0.24	4.8

λ_{ABS}, λ_{N^*} and λ_{T^*} are the maxima of absorption, N* and T* emission bands, respectively. N*/T* is the intensity ratio of the two emission bands measured at the peak maxima; QY is the fluorescence quantum yield. Excitation wavelength was 350 nm. Measurements were done in 10 mM phosphate buffer, 30 mM NaCl, pH 6.5. Peptide and oligonucleotide concentrations were 1 μM.

Table S3. Binding constants of NC(12-53) and 3HCaa-labeled NC(11-55) peptides with SL3.

Peptide	K_{bind}, M^{-1}
NC(12-53)[a]	$1(\pm0.2)\times10^6$
NC(11-55)-W37-3HCaa	$1(\pm0.3)\times10^6$
NC(11-55)-A30-3HCaa	$1(\pm0.8)\times10^6$
NC(11-55)-F16-3HCaa	$0.4(\pm0.1)\times10^6$

[a] – value taken from literature.[5] Binding constants of the 3HCaa-labeled NC(11-55) peptides with SL3 were obtained by monitoring the 3HC fluorescence. Excitation wavelength was 350 nm. Measurements were done in 50 mM Hepes buffer, 100 mM NaCl, pH 6.5. Concentration of SL3 was 0.5 μM.

References:

[1] a) C. D. Hufford, B. O. Oguntimein, J. N. Shoolery, *J. Org. Chem.* 1987, *52*, 5286; b) C. Chen, Y. F. Zhu, K. Wilcoxen, *J. Org. Chem.* 2000, *65*, 2574.

[2] H. De Rocquigny, D. Ficheux, C. Gabus, M. C. Fournie-Zaluski, J. L. Darlix, B. P. Roques, *Biochem. Biophys. Res. Commun.* 1991, *180*, 1010.

[3] J. W. Eastman, *Photochem. Photobiol.* 1967, *6*, 55.

[4] J. Godet, H. de Rocquigny, C. Raja, N. Glasser, D. Ficheux, J. L. Darlix, Y. Mely, *J. Mol. Biol.* 2006, *356*, 1180.

[5] C. Vuilleumier, E. Bombarda, N. Morellet, D. Gerard, B. P. Roques, Y. Mely, *Biochemistry* 1999, *38*, 16816.

2.1.2. Development of new ratiometric fluorophores for monitoring peptide – oligonucleotide interactions

In previous studies, site-specific monitoring of the interaction of labeled peptides (Shvadchak, Klymchenko et al. 2009) with their oligonucleotide targets was performed with 2-(2-furyl)-3HC fluorophore (FC) that shows a high sensitivity of its dual emission to environment changes in polar media, corresponding to that of peptides and proteins. Importantly, no fluorescence quenching of this probe on peptide-ODN interactions was observed but significant changes in its dual emission (N*/T* intensity ratio), which were assigned to a decrease in the environment polarity after intercalation of the probe between the ODN bases (Klymchenko, Shvadchak et al. 2008; Shvadchak, Klymchenko et al. 2009). However, the FC fluorophore shows some drawbacks such as a low pKa value and low fluorescence quantum yield in water.

To overcome these drawbacks, we selected in the present work the 4'-methoxyphenyl-3HC fluorophore which presents a 2-aryl group of intermediate electron donor ability between phenyl and 4'-(dialkylamino)phenyl groups (**Publication 2**). This 2-aryl group provides to the fluorophore an optimal sensitivity to environment changes in polar media and makes it the most suitable sensor unit for studying highly hydrated media. Based on this fluorophore, we developed three labels having different linker groups since the role of the linker, which may control intercalation of the dye between the ODN bases, was not explored before. Two of them (F6A and F6C) were modified at 6-position with linkers of different lengths and a third one (F4O) was conjugated with a linker from the opposite side of the fluorophore (Figure 2.5). As compared to the parent compound FC, the new labels display a higher sensitivity of the ratio of their two emission bands (N*/T*) to solvent polarity and H-bond donor ability as well as higher fluorescence quantum yields in water. Moreover, they show higher pK_a values of their 3-hydroxyl group (above 9), allowing their applications at neutral pH without interference of anionic forms.

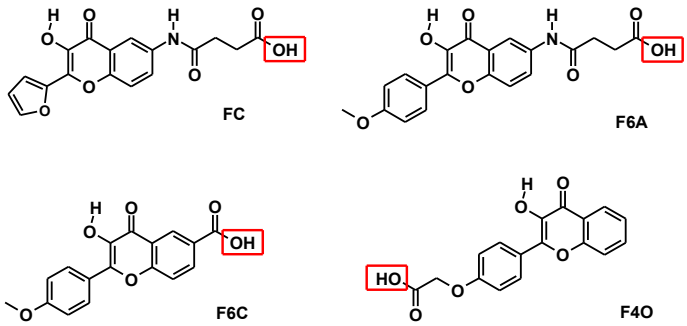


Figure 2.5. Structures of studied labels.

As a next step, these labels were covalently linked to the N-terminus of the model Tat(44-61) peptide (Kuciak, Gabus et al. 2008; Boudier, Storchak et al. 2010) that corresponds to the basic domain of the HIV-1 Tat protein, and which is one of the smallest known nucleic acid chaperones. This coupling did not modify the nucleic acid chaperone properties of the peptide. The emission spectra of the labeled peptides exhibited significant differences with the spectra of the free labels in

water which were assigned to an intramolecular screening of the label from the bulk water in conformations where the probe is in proximity to the rather apolar N-terminal part of the labeled peptide.

The fluorescence response of the new labels to the interaction of Tat(44-61) with ODNs of different size, composition and strandedness such as small single-stranded (ss) ODNs $(dA)_6$ and $(dA)_8$, the hairpin cTAR DNA sequence consisting of a 22 bp stem and a loop of 6 bases, single-stranded poly(dT), poly(dA) and the double-stranded calf thymus (CT DNA) sequences was investigated and the data were compared to those for the free FC label. For example, upon binding of the F4O-Tat(44-61) peptide to dsDNA, a dramatic decrease in the N*/T* ratio of its two emission bands was observed in comparison to the free peptide (Figure 2.6). All new labels showed at least 3-fold larger sensitivity of their dual emission to peptide-DNA interactions as compared to the parent label FC, due to their higher sensitivity to hydration (Figure 2.6). The level of probe hydration in the peptide-oligonucleotide complexes decreases in the following order: short ssDNAs >> long ssDNAs > DNA hairpins > dsDNAs. The main reason of the decrease of hydration is the more efficient stacking of the label with the DNA or RNA bases or base pairs in the various complexes. The strongest response was observed with F4O, presenting the most compact geometry for stacking with ODN bases (Figure 2.6B). The role of the linker in the label response was also described. These three labels are proposed as new improved tools for site-selectively monitoring peptide-ODN interactions.

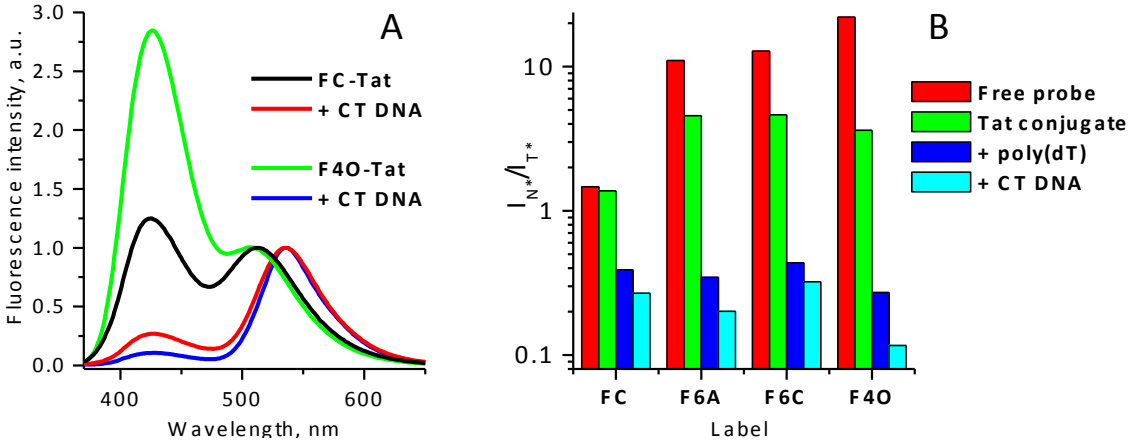


Figure 2.6. Changes in the spectroscopic properties of labeled Tat(44-61) peptides on interaction with oligonucleotides. (A) Fluorescence spectra of FC- and F4O- labeled Tat(44-61) peptides in the absence and in the presence of CT DNA. (B) N*/T* intensity ratios for free FC, F6A, F6C and F4O labels (red), for their conjugates with Tat(44-61) peptide alone (green) or in the presence of single-stranded poly(dT) (blue) or double-stranded CT DNA (cyan) in 10 mM phosphate buffer pH 7.0, 30 mM NaCl. In all cases, the concentration of DNA was adjusted to a ratio of 10 bases or base pairs per peptide.

Furthermore, we developed a new approach for the estimation of local hydrogen bond (HB) donor concentration using developed dual-fluorescence 3HF labels displaying selective sensitivity to HB donors and poor response to dipolar molecules, pH and ionic strength of the medium

(**Publication 3**). For these probes in binary solvent mixtures, a linear dependence was observed for log(N*/T*), the logarithm of the emission intensity ratio of two bands as a function of the local concentration of HB donor. The obtained calibration curve allowed us to quantitatively determine the hydration of the label, expressed as the partial volume of water in the label surrounding for different N-terminus labeled peptides and their complexes with DNAs.

Publication 2

Improved hydration-sensitive dual-fluorescence labels for monitoring peptide-nucleic acid interactions

Improved Hydration-Sensitive Dual-Fluorescence Labels For Monitoring Peptide−Nucleic Acid Interactions

Oleksandr M. Zamotaiev,[†] Viktoriia Y. Postupalenko,[‡] Volodymyr V. Shvadchak,[‡] Vasyl G. Pivovarenko,[†] Andrey S. Klymchenko,*[,‡] and Yves Mély*[,‡]

Department of Chemistry, Kyiv National Taras Shevchenko University, 01033 Kyiv, Ukraine, and Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74, Route du Rhin, 67401 ILLKIRCH Cedex, France. Received October 1, 2010; Revised Manuscript Received December 3, 2010

Environmentally sensitive labels constitute a new, attractive tool for monitoring biomolecular interactions. 3-Hydroxychromone derivatives are of particular interest because they undergo excited-state intramolecular proton transfer (ESIPT) showing dual emission highly sensitive to environmental hydration. To overcome the drawbacks of the previously developed label for sensing protein−DNA interactions based on 2-furanyl-3-hydroxychromone (FC), a series of hydration-sensitive labels based on 3-hydroxy-4′-methoxyflavone have been synthesized. As compared to FC, the new labels display higher sensitivity of the ratio of their two emission bands (N*/T*) to solvent polarity and H-bond donor ability, as well as higher fluorescence quantum yields in water. Moreover, they show higher pK_a values of their 3-hydroxyl group, allowing their application at neutral pH without interference of anionic forms. To illustrate the applications of these labels, we covalently coupled them to the N-terminus of the Tat(44−61) peptide that corresponds to the basic domain of the HIV-1 Tat protein. This coupling did not modify the nucleic acid chaperone properties of the peptide. Binding of oligonucleotides of varying length, sequence, and strandedness to the labeled peptides induced dramatic change in the N*/T* ratio of their two emission bands. This change indicated that the level of probe hydration in the peptide/oligonucleotide complexes decreases in the following order: short ssDNAs ≫ long ssDNAs > DNA hairpins > dsDNAs. The level of probe hydration was related to the ability of the probe to stack with the DNA bases or base pairs in the various complexes. The changes in the N*/T* ratio upon interaction of the labeled Tat peptides with DNA were about 3-fold larger with the new probes as compared to the parent FC label, in line with the higher sensitivity of the new probes to the environment. One of these labels, presenting the most compact geometry, showed the highest sensitivity, probably due to its optimal stacking with the DNA bases. Thus, the new hydration-sensitive labels appear as improved highly sensitive tools to site-selectively monitor the binding of peptides to oligonucleotides and nucleic acids.

INTRODUCTION

Environmentally sensitive (or solvatochromic) fluorescent dyes, which change their emission properties in response to changes in their environment, have become important tools for monitoring biomolecular interactions. In contrast to the commonly used Förster resonance energy transfer (FRET) technique, the approach based on environmentally sensitive probes requires labeling of only one of the interaction partners. Moreover, this approach appears complementary to the anisotropy technique, since it is focused on monitoring the local properties at the labeling site. In the past decade, a series of environmentally sensitive fluorescent labels, such as Prodan derivatives (*1*−*3*), dimethylaminophtalimide (*4*), dimethylaminonaphtamides (*5, 6*) and others (*7*), have been developed and applied to study protein−protein interactions and protein conformational transitions (*1, 2, 4, 5, 7*). Biomolecular interactions commonly decrease the polarity at the labeling site due to screening from water by the binding of the protein partner. This change in the polarity can be readily detected by environmentally sensitive labels through shifts in their emission maximum or changes in their fluorescence intensity (*8*).

* Corresponding authors. Yves Mély, yves.mely@unistra.fr, tel: +33 368 854263. Andrey Klymchenko, andrey.klymchenko@unistra.fr, tel: +33 368 854255, fax: +33 368 854313. Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74, Route du Rhin, 67401 ILLKIRCH Cedex, France.

[†] Kyiv National Taras Shevchenko University.

[‡] UMR 7213 CNRS.

For instance, the reactive derivative of Prodan, Acrylodan, attached at the N-terminus of peptide ligands, was used to monitor their interactions with cholecystokinin receptors, revealing receptor conformational changes associated with their ligand-driven activation (*9*). Moreover, the amino acid derivative of Prodan, Aladan, enabled monitoring of the binding of the S-peptide with ribonuclease S (*2*) and to estimate the local dielectric constant of the B1 domain of the staphylococcal protein G at different sites (*1*). In another study, aminophenoxazole maleimide, APM, was used for monitoring the conformational changes of the water-exposed domain of the β2 adrenergic receptor during its interactions with ligands and the voltage sensitive Shaker potassium channel (*7*). Moreover, a 4DMN-based amino acid was able to sense the phosphorylation-dependent binding of a synthetic peptide to the 14-3-3bp protein (*4*), while its improved analogue 6DMN was used for sensing peptide binding to proteins of the major histocompatibility complex (MHC) at the cell surface (*8*).

However, most examples in the literature are focused on protein−protein interactions, while applications of these dyes for sensing interactions of peptides with oligonucleotides (ODNs) have not been sufficiently explored. In part, this is due to the relatively polar ODN environment (*10*), so that polarity may not be dramatically affected by the interaction. In addition, most of the aforementioned environmentally sensitive dyes show relatively low quantum yields in polar protic media (*5*), and ODN bases frequently play the role of fluorescence quencher (*11, 12*). Recently, using a novel environmentally sensitive label based

10.1021/bc100434d
Published on Web 12/21/2010

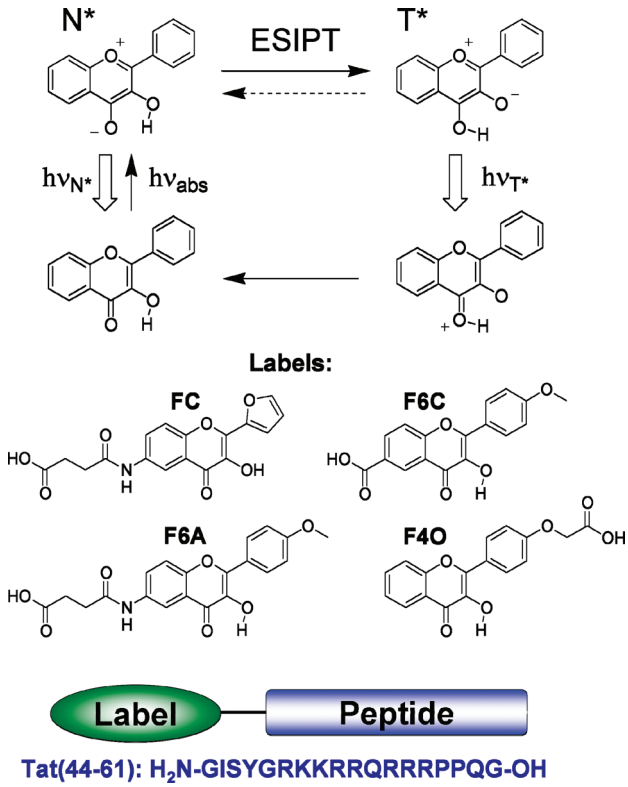


Figure 1. ESIPT reaction in 3HC dyes; structure of the studied labels and their conjugates with the Tat(44−61) peptide.

on 3-hydroxychromone (3HC), we have shown the possibility of site-specifically monitoring the interaction of labeled oligocations (*13*) and peptides (*14*) with their oligonucleotide targets. 3HC dyes undergo excited-state intramolecular proton transfer (ESIPT) (*15*) (Figure 1), resulting in the emission of both the normal (N*) excited state and the ESIPT product phototautomer (T*). The dual emission of 3HC dyes is highly sensitive to polarity and H-bonding interactions (*16*−*25*), which inhibit the ESIPT reaction and thus decrease the relative intensity of the T* band. In our previous studies, 2-(2-furyl)-3HC fluorophore was selected, since it shows high sensitivity of its dual emission to environmental changes in polar media corresponding to that of peptides and proteins (*16, 18, 26*). Importantly, we observed no quenching of fluorescence of this probe on peptide−ODN interactions but significant changes in its dual emission (N*/T* intensity ratio), which were assigned to a decrease in the environment polarity after intercalation of the probe between the ODN bases (*13, 14*). However, the 2-(2-furyl)-3HC fluorophore shows some drawbacks. At first, the pK_a value of its 3-OH group is around 8.4 (*26*); therefore, fluorescent anionic forms of the probe appear at pH > 7 and complicate the interpretation of the data. Moreover, the fluorescence quantum yield of 2-(2-furyl)-3HC in water is rather low (2%). Furthermore, the role of the linker group, which may control intercalation of the dye between the ODN bases, was not explored. Therefore, a search for new fluorophores and appropriate conjugation linkers is required to obtain improved fluorescent labels for monitoring peptide−ODN interactions, characterized by lower pH dependence and higher brightness and sensitivity.

In the present work, as an alternative to 2-(2-furyl)-3HC fluorophore, we selected the 4′-methoxyphenyl-3HC fluorophore that, similarly to 2-furyl, presents a 2-aryl group of intermediate electron donor ability between phenyl and 4′-(dialkylamino)phenyl groups. This 2-aryl group provides to the fluorophore an optimal sensitivity to environment changes in polar media and makes it the most suitable sensor unit for studying highly hydrated media. On the basis of this fluorophore, we developed three labels having different linker groups. Two of them (F6A and F6C) were modified at the 6-position with linkers of different lengths, and a third one was (F4O) conjugated with a linker from the opposite side of the fluorophore (Figure 1). The new labels were compared with the initially developed 2-(2-furyl)-3HC-based label FC. Our results show that the new labels are characterized by pK_a values above 9 and significantly improved brightness in water compared to FC. Being attached to the N-terminus of the model Tat(44−61) peptide (*27, 28*), the new labels showed a nearly 3-fold larger ratiometric response to peptide−ODN interactions, which was connected with their higher sensitivity to their environment. The role of the linker group in the label response was also described. The new labels are proposed as improved tools for monitoring site-selective peptide−ODN interactions.

MATERIALS AND METHODS

Reagents and Solvents. All the reagents were purchased from Sigma-Aldrich Chemical Co. Solvents for synthesis were of reagent-grade quality and were appropriately dried if necessary. For absorption and fluorescence studies, the solvents were of spectroscopic grade. Calf thymus DNA (CT-DNA), cTAR, and single-stranded DNA poly dA and poly dT were from Sigma. Short ODNs (6−8 oligonucleotides) were synthesized by IBA. Concentrations of the ODNs were calculated from their absorbance using the molar extinction coefficients at 260 nm specified by the supplier.

Synthesis of the Labels. Synthesis of FC and F6C compounds were published elsewhere (*14, 29*). The F6A and F4O labels were synthesized by a similar multistep pathway, through the preparation of corresponding key compounds: 4-alkoxybenzaldehydes and 2′-hydroxyacetophenones. The structure and purity of all synthesized compounds were proven by LC MS, NMR, and UV spectroscopy (see also Supporting Information). All the obtained labels were of 98% purity according to HPLC. ^{1}H NMR (300 MHz, DMSO-d_6): **F6A**: δ 2.57 (t, 2H, **CH$_2$**), 3.32 (t, 2H, N**CH$_2$**), 3.84 (s, 3H, O**CH$_3$**), 7.12 (d, 2H, J = 9 Hz, Ar**H**), 7.71 (d, 1H, J = 9 Hz, Ar**H**), 7.86 (dd, 1H, J = 9 Hz, J = 2.2 Hz, Ar**H**), 8.18 (d, 2H, J = 9 Hz, Ar**H**), 8.44 (d, 1H, J = 2.2 Hz, Ar**H**), 9.42 (s, 1H, **OH**), 10.27 (s, 1H, **NH**), 12.12 (br s, 1H, **COOH**). m/z (M+H$^+$) calculated for $C_{20}H_{18}NO_7$: 384.1; found 384.2. **F6C**: δ 3.82 (s, 3H, O**CH$_3$**), 7.09 (d, 2H, J = 8.5 Hz, Ar**H**), 7.80 (d, 1H, J = 8.5 Hz, Ar**H**), 8.17 (d, 2H, J = 8.5 Hz, Ar**H**), 8.22 (d, 1H, J = 8.5 Hz), 8.61 (s, 1H, Ar**H**), 9.71 (s, 1H, O**H**). m/z (M+H$^+$) calculated for $C_{17}H_{13}O_6$: 313.0; found 313.0. **F4O**: δ 4.79 (s, 2H, **CH$_2$**), 7.11 (d, 2H, J = 8.8 Hz, Ar**H**), 7.46 (t, 1H, J = 7.2 Hz, Ar**H**), 7.76 (m, 2H, Ar**H**), 8.10 (d, 1H, J = 7.4 Hz, Ar**H**), 8.18 (d, 2H, J = 8.8 Hz, Ar**H**). m/z (M+H$^+$) calculated for $C_{17}H_{13}O_6$: 313.0; found 313.0.

Peptide Synthesis. Peptides were synthesized by solid-phase peptide synthesis on a 433A synthesizer (ABI, Foster City, CA). The synthesis was performed at on 0.25 mmol scale using standard side-chain protected fluorenylmethoxycarbonyl (Fmoc)-amino acids and HBTU/HOBt coupling protocol. LL-HMP resin (ABI, 0.44 mmol/g reactive group concentrations) was used as a solid support. At the end of the synthesis, peptidyl resin was isolated and washed twice by NMP.

Two equivalents (0.06 mmol) of the labels (FC, F6A, F6C, F4O) was dissolved in 1 mL of NMP mixed with 2 equiv of HBTU/HOBt coupling solution (in DMF) and added to Fmoc-deprotected peptidyl resin (0.03 mmol) swelled in 1 mL of NMP. After a few minutes of shaking, 5 equiv of DIEA solution was added. Then, the reaction mixture was stirred overnight at 40 °C. Resin was filtered and washed with NMP, methanol, and DCM.

Table 1. Spectroscopic Properties of FC, F6A, F6C, and F4O Labels[a]

media (E_T(30))	label	λ_{abs}, nm	λ_{N^*}, nm	λ_{T^*}, nm	N*/T*	QY, %
Water (pH 7) (63.1)	FC	357	431	514	1.47	2.0
	F6A	351	443	514	11	10.2
	F6C	355	442	517	13	7.8
	F4O	353	437	516	22	6.4
Water (pH 11)	FC	419	508	−	−	0.8
	F6A	411	512	−	−	0.9
	F6C	409	511	−	−	0.9
	F4O	406	508	−	−	0.5
MeOH (55.4)	FC	349	423	533	0.81	5.6
	F6A	349	429	533	1.10	5.9
	F6C	354	428	535	0.97	4.65
	F4O	356	432	529	2.78	4.3
EtOH (51.9)	FC	351	420	535	0.34	6.1
	F6A	351	424	535	0.28	6.3
	F6C	356	419	533	0.23	4.9
	F4O	359	430	535	1.25	3.5
DMF (43.2)	FC	344	421	539	0.10	6.9
	F6A	345	411	539	0.15	4.55
	F6C	354	418	544	0.13	5.0
	F4O	356	432	540	0.30	2.6
THF (37.4)	FC	339	408	537	0.07	5.5
	F6A	338	409	538	0.02	6.3
	F6C	355	411	544	0.04	9.5
	F4O	350	406	537	0.04	6.7

[a] E_T(30), empiric polarity index; λ_{abs}, position of absorption maximum; λ_{N^*} and λ_{T^*}, positions of fluorescence maxima of N* and T* forms, respectively; N*/T*, intensity ratio of the two emission bands at their peak maxima (calculated after mathematical deconvolution of the spectra); QY, fluorescence quantum yield.

Cleavage and deprotection of the peptidyl resin were performed for 1.5 h using a 10 mL trifluoroacetic acid (TFA) solution containing water (5%, v/v), TIS ($(iPr)_3SiH$, 2.5%, v/v), phenol (1%, w/v), thioanisole (5%, v/v), and ethane dithiol (2.5%, v/v). The solution was concentrated in vacuo and the peptide was precipitated by using ice-cold diethyl ether and then pelleted by centrifugation. The pellet was washed with diethyl ether and dried. The peptides were solubilized with aqueous TFA (0.05%, v/v). HPLC purification was carried out on a C8 column (uptisphere 300A, 5 μm; 250 × 10, Interchim, France) in water/acetonitrile mixture containing 0.05% TFA with linear gradients depending on the peptide (typically 10−40% of acetonitrile for 30 min) and monitored at 210 and 370 nm. Molecular masses obtained by ion spray mass spectrometry were 2523, 2562, 2491, and 2491 and for FC-Tat, F6A-Tat, F6C-Tat, and F4O-Tat, respectively, in agreement with the expected theoretical masses. Prior to use, peptides were dissolved in distilled water, aliquoted, and stored at −20 °C. Concentrations of the labeled peptides were determined from the label absorbance at the band maximum (Table 1) using $\varepsilon = 15\,000\ M^{-1}\ cm^{-1}$ for FC label (*14*) and $\varepsilon = 23\,000\ M^{-1}\ cm^{-1}$ for F6A, F6C, and F4O labels, respectively.

Instrumentation. Proton NMR spectra were recorded on a 300 MHz Bruker spectrometer, and mass spectra were recorded on a Mariner System 5155 mass spectrometer using the electrospray ionization (ESI) method. All column chromatography experiments were performed on silica gel (Merck, Kieselgel 60H, Art 7736). Absorption and fluorescence spectra were recorded on a Cary 400 spectrophotometer (Varian) and FluoroMax 3.0 spectrofluorimeter (Jobin Yvon, Horiba), respectively. For fluorescence studies, the dyes were used at 0.5 to 1 μM concentrations. Excitation wavelength was at the absorption maximum for each label. Fluorescence quantum yields were determined using quinine sulfate in 0.5 M sulfuric acid as a reference (quantum yield, $\varphi = 0.57$) (*30*). For the experiments in water, 10 mM phosphate buffer containing 30 mM NaCl (pH 7.0) was used systematically.

Analysis of Tat Chaperone Activity. To check the influence of the labels on Tat(44−61) chaperone properties, the ability of the FC-, F6C-, F6A-, and F4O-labeled Tat(44−61) peptides to promote the annealing of cTAR with its complementary dTAR sequence was compared to that of the unlabeled Tat peptide (*27, 28*). The kinetic measurements were performed under pseudo-first-order conditions by using unlabeled dTAR at a concentration which was 30-fold higher than the concentration of cTAR labeled with carboxytetramethylrhodamine (TMR) at the 5′ end and with 5/6-carboxyfluorescein (Fl) at the 3′ end. Excitation and emission wavelengths were 480 and 520 nm, respectively, for monitoring the Fl fluorescence. All reported concentrations correspond to those after mixing. To avoid high local concentrations during mixing, both reactants were mixed at the same volume. Peptides were added to each reactant separately at a peptide:ODN ratio of 2, and then, the reaction was initiated by mixing the peptide-coated ODNs together. Experiments were performed in 25 mM Tris (pH 7.5) and 30 mM NaCl with 0.2 mM $MgCl_2$ at 20 °C.

Deconvolution of Spectra. Deconvolution of the fluorescence spectra of the labeled peptides into N* and T* bands was performed using the Siano software kindly provided by Prof. A. O. Doroshenko (Kharkov, Ukraine), as previously described (*31*). The program is based on an iterative nonlinear least-squares method, where the individual emission bands were approximated by a log-normal function accounting for several parameters: maximal amplitude, I_{max}; spectral maximum position, ν_{max}; and position of half-maximum amplitudes, ν_1 and ν_2; for the blue and red parts of the band, respectively. These parameters determine the shape parameters of the log-normal function, namely, the full width at the half-maximum, fwhm = $\nu_1 - \nu_2$, and the band asymmetry, $P = (\nu_1 - \nu_{max})/(\nu_{max} - \nu_2)$. All parameters were allowed to vary in the iteration process. The resulting fluorescence intensities at the peak maximum of the separated N* and T* bands were used for calculation of the N*/T* ratio.

RESULTS AND DISCUSSION

All new labels are based on the same fluorophore bearing a 4′-methoxyphenyl group at the 2-position of the 3HC unit. They were synthesized starting from the corresponding 4′-alkoxybenzaldehyde and 2′-hydroxyacetophenone, which were condensed in basic condition into corresponding chalcones that were further converted into 3HC derivatives using Algar-Flynn-Oyamada reaction.

Spectroscopic Characteristics of the Free Labels. Absorption spectra of the three studied labels (F6A, F6C, and F4O) are characterized by a single band with a maximum at 350−355 nm. These absorption spectra show limited solvatochromism, having maximum absorbance wavelengths poorly dependent on the nature of the solvent (Table 1). Interestingly, while the same absorption band was observed for these labels in 15 mM phosphate buffer at acidic and neutral pH values, a second band with a maximum absorption wavelength at 410 nm appeared at basic pH values (pH > 7.8, Figure 2). This second band corresponds to the anionic form of the probes, a product of deprotonation of their 3-OH group. From acid−base titration data, the pK_a values of their 3-OH group were found to be equal to 9.37, 9.25, and 9.37, respectively, for F6A, F6C, and F4O labels. These pK_a values are thus larger by about one unit as compared to the pK_a value (8.4) of the FC label (*26*). This constitutes a decisive advantage, which allows application of the new probes at neutral pH without interference of the anionic band.

The fluorescence spectra of the new labels were recorded in a range of solvents from moderately polar aprotic solvents, such as tetrahydrofuran, to highly polar hydrogen bond donor solvents, such as alcohols and water (Table 1, Figure 3, and Table S1 in Supporting Information). The quantum yields of the new labels appear poorly dependent on the nature of the

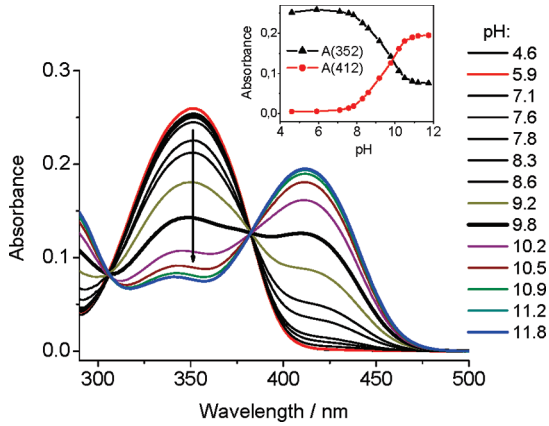


Figure 2. Absorption spectra of the F4O probe at different pH in 15 mM phosphate buffer. Inset: changes in absorbance at 352 and 412 nm versus pH.

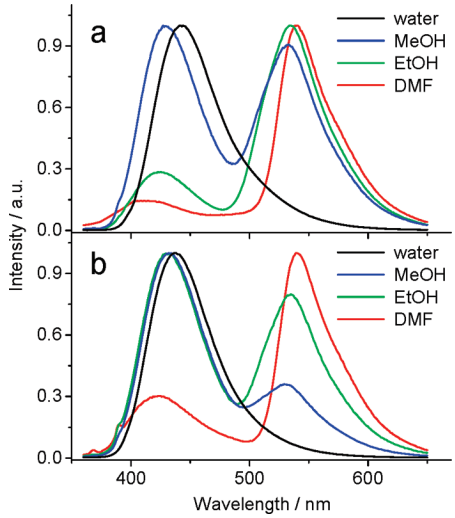


Figure 3. Normalized fluorescence spectra of F6A (a) and F4O (b) labels in different solvents. Excitation wavelength was fixed at the absorption maximum in each case.

solvent and similar to the FC ones, with the exception of water, where the quantum yields of the new labels are 3- to 5-fold higher than the FC one. Moreover, in aqueous solution at pH 11, where only anionic forms of the dyes are expected, the fluorescence quantum yields were rather small, similar to that of FC at the same pH. Therefore, we can conclude that, at neutral pH, the fluorescence of the new labels, characterized by a rather high pK_a value, will not be contaminated by the anionic form, in contrast to the case of FC. Importantly, the F6A, F6C, and F4O labels exhibit a strong sensitivity of the intensity ratio of their emission bands, N*/T*, to the solvent nature. In the moderately polar aprotic solvent tetrahydrofuran, the three labels show a very low N*/T* ratio. This ratio increases in the more polar aprotic solvent DMF. In protic solvents, the N*/T* ratio is much higher than in aprotic solvents of similar polarity, indicating that the H-bond donor ability of protic solvents inhibits the ESIPT reaction, and thus decreases the emission of the ESIPT product T* (*16, 26*). The band intensity ratio further increases with polarity in alcohols and reaches its highest value in water (Figure 3, Table 1). In water, the intensity of the T* band is so low that it is embedded in the red-edge part of the N* emission band. In this case, the T* band position is determined using a well-established spectral deconvolution procedure (*31*). Noticeably, similar values of the N*/T* ratio were found in buffer and water independent of the salt

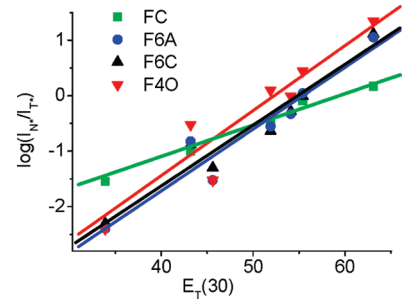


Figure 4. Dependence of log(N*/T*) on the empiric polarity index $E_T(30)$ (symbols) for FC (green), F6A (blue), F6C (black), and F4O (red) labels. The solid lines correspond to the linear fits to the experimental data for each probe. The $E_T(30)$ values (*32*) of the solvents used and band ratios are given in Table 1 and Supporting Information Table S1.

concentration (data not shown), indicating that the value of the N*/T* ratio is controlled only by water.

Similarly to the parent 3HC dyes of close structure (*16, 26*), the log(N*/T*) parameter of the three new labels was found to vary linearly with the solvent empirical polarity index $E_T(30)$ (Figure 4). The latter is an integrated parameter that accounts for both dielectric properties of solvent and their H-bond donor ability (*32*). This linear dependence confirms that the N*/T* ratio of the new labels depends on both the dielectric constant and the H-bond donor ability of solvents. Furthermore, in line with our expectations, the new labels were found to be more sensitive to the $E_T(30)$ index than FC, as could be seen from their steeper slope in Figure 4. Noticeably, in water, characterized by high dielectric constant and H-bond donor ability, we also observed an outstanding blue shift of the T* band for the three new labels in comparison with alcohols or aprotic solvents (Table 1). The similar fluorescence properties of F6A, F6C, and F4O probes in solvents further suggest that the nature of the linker group and its length and position in the 3HC fluorophore do not substantially affect its fluorescence. Due to their high sensitivity to polarity changes in polar media, the F6A, F6C, and F4O probes appear as prospective candidates for peptide labeling in order to sense and characterize interactions of peptides with polar molecules, such as ODNs and nucleic acids. Therefore, being attached to a peptide, the present labels are expected to report on the accessibility of the labeling site to bulk water by its N*/T* ratio and the position of the T* band.

Synthesis and Characterization of the Labeled Tat(44−61) Peptides. To demonstrate the ability of the synthesized labels to monitor protein−ODN interactions, we used the Tat(44−61) peptide, which corresponds to the basic domain of the Tat protein of the human immunodeficiency virus of type 1, and is one of the smallest known nucleic acid chaperones (*27, 28*). The Tat(44−61) peptide was synthesized by solid-phase chemistry (*33, 34*) and covalently labeled at its N-terminus, by any of the FC, F6A, F6C, and F4O labels. The chaperone properties of the labeled peptides were tested using previously described protocols (*28, 35*) and were found comparable to those of the nonlabeled peptide (Figure S1 in Supporting Information), indicating that the labels do not interfere with the peptide activity. The emission spectra of the labeled Tat(44−61) exhibited significant differences from the spectra of the free labels in water, with a decrease of the N*/T* ratio and a 5−9 nm red shift of the T* band (Figure 5A, Table 2). According to the data of the free probe in model solvents (Table 1), these changes indicate a decrease in the polarity and/or H-bond donor ability of the probe surrounding. This decrease was independent of the peptide concentration in the 50−1000 nM range (data not shown) and was thus interpreted as an intramolecular screening of the label from the bulk water in conformations

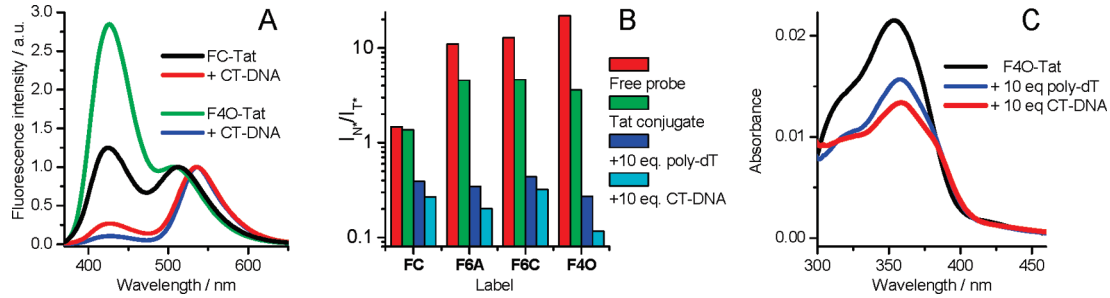


Figure 5. Changes in the spectroscopic properties of labeled Tat(44−61) peptides upon interaction with oligonucleotides. (A) Fluorescence spectra of FC- and F4O- labeled Tat(44−61) peptides in the absence and in the presence of CT-DNA. (B) N*/T* intensity ratios for free FC, F6A, F6C, and F4O labels (red) and for their conjugates with Tat(44−61) peptide in the absence (green) or in the presence of poly dT (blue) or CT-DNA (cyan). (C) Absorption spectra of the F4O-labeled Tat(44−61) peptide in the absence and in the presence of 10 equiv of single-stranded poly dT and double-stranded CT-DNA. All spectra were recorded in 10 mM phosphate buffer pH 7.0 and 30 mM NaCl. Concentration of peptide was 1 μM. In all cases, the concentration of DNA was adjusted to a ratio of 10 bases or base pairs per peptide.

Table 2. Spectroscopic Properties of FC, F6A, F6C, and F4O Labels Covalently Linked to Tat(44-61) Peptide Complexed with Various Oligonucleotides[a]

sample	λ_{abs}, nm	λ_{N^*}, nm	λ_{T^*}, nm	N*/T*	QY, %
FC-Tat	355	426	516	1.37	3.4
F6A-Tat	354	442	520	4.56	7.3
F6C-Tat	362	443	524	4.64	2.0
F4O-Tat	354	426	515	3.61	4.2
FC-Tat: poly dT	357	426	532	0.390	3.7
F6A-Tat: poly dT	358	437	534	0.345	5.0
F6C-Tat: poly dT	365	438	541	0.435	3.3
F4O-Tat: poly dT	358	424	532	0.270	5.3
FC-Tat: CT-DNA	358	429	535	0.269	2.0
F6A-Tat: CT-DNA	355	441	537	0.201	2.3
F6C-Tat: CT-DNA	366	442	544	0.322	1.4
F4O-Tat: CT-DNA	359	429	537	0.116	3.8
F4O-Tat: $(dA)_6$	357	427	526	1.24	5.8
F4O-Tat: $(dA)_8$	357	429	528	0.581	6.5
F4O-Tat: cTAR	358	425	533	0.177	3.8
F4O-Tat: poly dA	358	430	530	0.343	3.7

[a] λ_{abs}, position of absorption maxima; λ_{N^*} and λ_{T^*}, position of the fluorescence maxima of N* and T* forms, respectively; N*/T*, intensity ratio of the two emission bands at their peak maxima; QY, fluorescence quantum yield.

where the probe is in proximity to the rather apolar GISYG(44−48) N-terminal sequence of the labeled peptide. Indeed, a large collection of conformations is probably explored by the N-terminal domain of Tat(44−61) as a consequence of the flexibility of both the peptide and the spacer connecting the label to the peptide (*36*). Noticeably, both the N*/T* band ratio and the T* band position of the labeled peptides were marginally dependent on the salt concentration (data not shown), indicating that the intramolecular screening of the dye from water was poorly affected by the ionic strength.

The fluorescence response of the new labels to the interaction of Tat(44−61) with ODNs of different strandedness, length, and structure was investigated and the data were compared to those for the FC label. In a first step, we compared the effects of the single-stranded poly dT sequence and the double-stranded calf-thymus DNA (CT-DNA) on the fluorescence properties of the Tat(44−61) peptides labeled by FC, F6A, F6C, and F4O probes. Both nucleic acid sequences induced only limited changes in the position of the N* emission maximum and the fluorescence quantum yield of the probes (Table 2). In contrast, the two sequences induced a dramatic decrease in the N*/T* ratio, the decrease being more pronounced with the CT-DNA (Figure 5B, Table 2). In addition, both nucleic acid sequences induced a strong red shift (up to 22 nm) in the position of the T* band maximum, with again a more pronounced shift in the case of CT-DNA. These spectral changes strongly suggested an efficient screening of the labels from water on binding of the labeled peptides to poly dT and CT-DNA. In addition, the 3−4 nm red-shifted position of the absorption maximum together with the ∼20% decreased absorbance of all four labeled peptides in their complexes with poly dT and CT-DNA (in the case of F4O label, see Figure 5C) indicate that the labels stack with the bases of both sequences. Moreover, the stronger spectroscopic changes observed with CT-DNA as compared to those of poly dT further suggested that the screening of the labels from water was more efficient in double-stranded than in single-stranded sequences. This is likely a consequence of the lower accessibility to water for the label when intercalated in the more rigid dsDNA sequence as compared to the stacked conformation achieved in the more flexible ssDNA sequences. Thus, these labels appear to be a sensitive means to discriminate the binding of the Tat(44−61) peptide to dsDNA from that to ssDNA. Interestingly, a much more pronounced difference in the accessibility to water on binding to ss and ds sequences was obtained when the FC label was conjugated to spermine (*13*), suggesting that the interaction of the covalently bound labels with the bases primarily depends on the binding mode of the peptide to the nucleic acid sequence. Noticeably, no anionic form of the F6A, F6C, and F4O labels was perceived for the labeled Tat(44−61) peptides both in the absence and in the presence of either (ss) or (ds) sequences, in agreement with the observed high pK_a values for the new labels.

Further comparison of the labeled peptides revealed that F6A, F6C, and F4O labels show systematically stronger changes (about 3-fold larger) than the FC label in their N*/T* ratio on complexation to ODNs. Thus, the transfer of these labels from the highly polar hydrated environment of the peptide N-terminus to the less hydrated sites within the DNA bases leads to stronger spectroscopic effects, which is evidently connected to their much higher sensitivity to the environment, as evidenced from the data in organic solvents. The changes of highest amplitude were observed with the F4O label, where a 31-fold decrease in the N*/T* ratio was observed on interaction with CT-DNA. This may likely be a consequence of the more compact structure of this label that favors an efficient stacking of the 3HC moiety with the DNA base pairs (Figure 6). The label F6C shows the smallest decrease of N*/T* ratio upon binding with double-stranded CT-DNA. Since this label has a shorter spacer than F6A, a longer linker is probably required for optimal intercalation of the probe within the DNA base pairs.

To further investigate the probe response to ODN binding, we characterized the spectroscopic response of the F4O-labeled Tat(44−61) peptide on binding to small single-stranded (ss) ODNs such as $(dA)_6$ and $(dA)_8$, the hairpin cTAR DNA sequence consisting of a 22 bp stem and a loop of 6 bases, as well as the single-stranded poly dA sequence (Table 2 and Figure S2 in Supporting Information). Comparison of the spectroscopic parameters of F4O-Tat(44−61) bound to poly dA

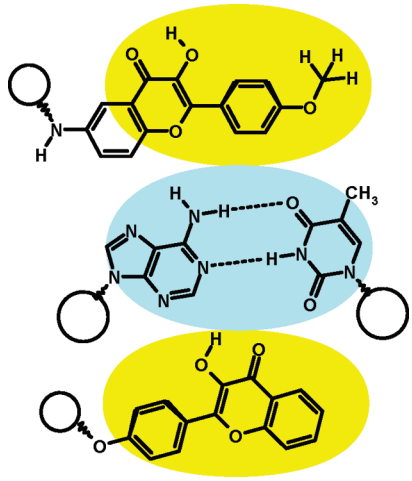


Figure 6. Comparison of the size of the F6A (top) and F4O (bottom) labels with that of an A-T base pair.

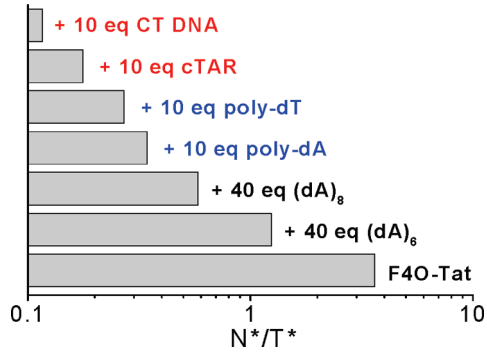


Figure 7. Fluorescence ratiometric response of F4O-labeled Tat(44−61) peptide to interaction with various oligonucleotides. Concentration of peptide was 0.5 μM in 10 mM phosphate buffer pH 7.0 and 30 mM NaCl. The ODNs were added at a ratio of 10 or 40 bases or base pairs per peptide molecule. In all cases, full complexation of the peptide by the ODNs was achieved, as no change in the N*/T* ratio was observed after a 2-fold increase in the concentration of ODNs.

and poly dT (Table 2, Figure 7) revealed that the response of the F4O showed some dependence on the nature of the DNA bases. Interestingly, the binding of F4O-Tat(44−61) peptide to the cTAR DNA stem-loop gave an N*/T* ratio and T* band position intermediate to those of ssDNA and dsDNA (Figure 6, Table 2), indicating that the peptide binds both to the single stranded loop and to the double-stranded stem of the cTAR sequence. In the case of the short $d(A)_6$ and $d(A)_8$ sequences, the rather high values of the N*/T* ratio and the blue-shifted position of the T* band indicated a significant exposure to water. In fact, with the decrease of ODN length (poly dA → $(dA)_8$ → $(dA)_6$), the N*/T* ratio increases, while the T* band shifts to the blue, indicating that stacking of F4O with the ODN bases is less favored with the smaller ODNs. This may be the result of a different binding mode of the Tat peptide to short ODNs, as well as a higher probability of getting the F4O label located at the end of ODN, with no possibility of stacking with the ODN bases.

Thus, using the N*/T* ratio of the F4O label covalently bound to Tat(44−61), it is possible to distinguish the interaction of the peptide with four types of DNAs, namely, small single-stranded ODNs, single-stranded DNAs, partially double-stranded DNAs, and fully double-stranded DNAs. Alternatively, the T* band position can also be used for the same purpose, though the changes in this parameter are less sensitive than those in the N*/T* ratio. As a consequence of the high sensitivity of the F4O label to the nature of the bound ODN, it is possible to compare the relative affinities of this peptide to ss and ds DNAs. For this aim, we compared the fluorescence spectra of F4O-Tat(44−61) peptide added to a 1:1 (as expressed in base concentrations) mixture of poly dT and CT-DNA with the spectra obtained at the same concentrations of the individual DNAs (Figure S3 in Supporting Information). The N*/T* ratio of the peptide in the presence of the mixture of poly dT and CT-DNA was in between the ratios obtained from individual DNAs. Taking into account that the fluorescence quantum yields of the corresponding peptide−DNA complexes are nearly the same (Table 2), we can conclude that Tat peptide exhibits a similar affinity for ss and ds DNAs.

CONCLUSIONS

A new series of hydration-sensitive labels, namely, F6A, F6C, and F4O probes, have been synthesized. These probes display a higher sensitivity of their N*/T* ratio to the environment as compared to their parent FC probe. They also exhibit higher pK_a values than FC, which prevents any significant interference of their anionic form at neutral pH. These labels were covalently linked to the N-terminus of the Tat(44−61) peptide and used for sensing the binding of ODNs of different size, composition, and strandedness. The N*/T* ratio of the labeled Tat peptide was observed to change in the following order: free Tat ≫ Tat/short ssODNs ≫ Tat/long ssDNAs > Tat/stem-loops > Tat/dsDNAs. The low N*/T* values and red-shifted T* band position of the complexes with dsDNAs were associated with the low water accessibility of the probe intercalated between the base pairs. A slight increase in the hydration of the probes was observed when they were stacked with the bases of the more flexible ssDNAs. Finally, the rather high N*/T* ratio and blue-shifted position of the T* band in complexes with short ssDNAs are thought to result from a different binding mode of the peptide and/or the location of the probe at one end of the peptide−ODN complex with low possibility to stack with the bases. All new labels showed at least a 3-fold larger sensitivity of their dual emission to peptide−ODN interactions compared to the parent label FC, which can be explained by their higher sensitivity to hydration. Comparison of the new labels shows that the strongest response to peptide−ODN interactions was observed with F4O, presenting the most compact geometry for stacking with ODN bases. Moreover, we found that the longer linker (5 atoms) allows better fluorophore stacking than the shorter (1 atom) one. Thus, the new generation of hydration-sensitive probes appears as a highly sensitive ratiometric tool to site-selectively monitor the binding of peptides to ODNs.

ACKNOWLEDGMENT

This work and VYP were supported by ANR (ANR-07-BLAN-0287, FLU-AA-NT), CNRS, Université de Strasbourg, and the ARCUS program between France, Ukraine, and Russia. We thank H. de Rocquigny for his help in peptide synthesis. VGP is thankful to Université de Strasbourg for support of a one month stay.

Supporting Information Available: Procedures for synthesis of F6A and F4O labels, spectroscopic properties of these labels, and characterization of labeled peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Cohen, B. E., McAnaney, T. B., Park, E. S., Jan, Y. N., Boxer, S. G., and Jan, L. Y. (2002) Probing protein electrostatics with a synthetic fluorescent amino acid. *Science 296*, 1700–1703.
(2) Nitz, M., Mezo, A. R., Ali, M. H., and Imperiali, B. (2002) Enantioselective synthesis and application of the highly fluorescent and environment-sensitive amino acid 6-(2-dimethylami-

nonaphthoyl) alanine (DANA). *Chem. Commun. 8*, 1912−1913.
(3) Vazquez, M. E., Nitz, M., Stehn, J., Yaffe, M. B., and Imperiali, B. (2003) Fluorescent caged phosphoserine peptides as probes to investigate phosphorylation-dependent protein associations. *J. Am. Chem. Soc. 125*, 10150–10151.
(4) Vazquez, M. E., Rothman, D. M., and Imperiali, B. (2004) A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis. *Org. Biomol. Chem. 2*, 1965–1966.
(5) Vazquez, M. E., Blanco, J. B., and Imperiali, B. (2005) Photophysics and biological applications of the environment-sensitive fluorophore 6-N, N-dimethylamino-2,3-naphthalimide. *J. Am. Chem. Soc. 127*, 1300–1306.
(6) Loving, G., and Imperiali, B. (2008) A versatile amino acid analogue of the solvatochromic fluorophore 4-N, N-dimethylamino-1,8-naphthalimide: A powerful tool for the study of dynamic protein interactions. *J. Am. Chem. Soc. 130*, 13630–13638.
(7) Cohen, B. E., Pralle, A., Yao, X., Swaminath, G., Gandhi, C. S., Jan, Y. N., Kobilka, B. K., Isacoff, E. Y., and Jan, L. Y. (2005) A fluorescent probe designed for studying protein conformational change. *Proc. Natl. Acad. Sci. U.S.A. 102*, 965–970.
(8) Venkatraman, P., Nguyen, T. T., Sainlos, M., Bilsel, O., Chitta, S., Imperiali, B., and Stern, L. J. (2007) Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells. *Nat. Chem. Biol. 3*, 222–228.
(9) Harikumar, K. G., Pinon, D. I., Wessels, W. S., Prendergast, F. G., and Miller, L. J. (2002) Environment and mobility of a series of fluorescent reporters at the amino terminus of structurally related peptide agonists and antagonists bound to the cholecystokinin receptor. *J. Biol. Chem. 277*, 18552–18560.
(10) Sen, S., Paraggio, N. A., Gearheart, L. A., Connor, E. E., Issa, A., Coleman, R. S., Wilson Iii, D. M., Wyatt, M. D., and Berg, M. A. (2005) Effect of protein binding on ultrafast DNA dynamics: Characterization of a DNA:APE1 complex. *Biophys. J. 89*, 4129–4138.
(11) Yang, X., Liu, W. H., Jin, W. J., Shen, G. L., and Yu, R. Q. (1999) DNA binding studies of a solvatochromic fluorescence probe 3-methoxybenzanthrone. *Spectrochim. Acta, Part A 55*, 2719–2727.
(12) Kumar, C. V., and Asuncion, E. H. (1993) DNA binding studies and site selective fluorescence sensitization of an anthryl probe. *J. Am. Chem. Soc. 115*, 8547–8553.
(13) Klymchenko, A. S., Shvadchak, V. V., Yushchenko, D. A., Jain, N., and Mely, Y. (2008) Excited-state intramolecular proton transfer distinguishes microenvironments in single-and double-stranded DNA. *J. Phys. Chem. B 112*, 12050–12055.
(14) Shvadchak, V. V., Klymchenko, A. S., De Rocquigny, H., and Mely, Y. (2009) Sensing peptide - Oligonucleotide interactions by a two-color fluorescence label: Application to the HIV-1 nucleocapsid protein. *Nucleic Acids Res. 37*, e25.
(15) Sengupta, P. K., and Kasha, M. (1979) Excited state proton-transfer spectroscopy of 3-hytdroxyflavone and quercetin. *Chem. Phys. Lett. 68*, 382–385.
(16) Das, R., Klymchenko, A. S., Duportail, G., and Mely, Y. (2009) Unusually slow proton transfer dynamics of a 3-hydroxychromone dye in protic solvents. *Photochem. Photobiol. Sci. 8*, 1583–1589.
(17) Shynkar, V. V., Klymchenko, A. S., Piemont, E., Demchenko, A. P., and Mely, Y. (2004) Dynamics of intermolecular hydrogen bonds in the excited states of 4′-dialkylamino-3-hydroxyflavones. On the pathway to an ideal fluorescent hydrogen bonding sensor. *J. Phys. Chem. A 108*, 8151–8159.
(18) Enander, K., Choulier, L., Olsson, A. L., Yushchenko, D. A., Kanmert, D., Klymchenko, A. S., Demchenko, A. P., Mely, Y., and Altschuh, D. (2008) A peptide-based, ratiometric biosensor construct for direct fluorescence detection of a protein analyte. *Bioconjugate Chem. 19*, 1864–1870.
(19) Chou, P.-T., Martinez, M. L., and Clements, H. (1993) Reversal of excitation behavior of proton-transfer vs. charge-transfer by dielectric perturbation of electronic manifolds. *J. Phys. Chem. 97*, 2618–2622.
(20) Ormson, S. M., Brown, R. G., Vollmer, F., and Rettig, W. (1994) Switching between charge- and proton-transfer emission in the excited state of a substituted 3-hydroxyflavone. *J. Photochem. Photobiol., A 81*, 65–72.
(21) Swinney, T. C., and Kelley, D. F. (1993) Proton transfer dynamics in substituted 3-hydroxyflavones: Solvent polarization effects. *J. Chem. Phys. 99*, 211–221.
(22) Klymchenko, A. S., and Demchenko, A. P. (2003) Multiparametric probing of intermolecular interactions with fluorescent dye exhibiting excited state intramolecular proton transfer. *Phys. Chem. Chem. Phys. 5*, 461–468.
(23) Demchenko, A. P., Mely, Y., Duportail, G., and Klymchenko, A. S. (2009) Monitoring biophysical properties of lipid membranes by environment-sensitive fluorescent probes. *Biophys. J. 96*, 3461–3470.
(24) Strandjord, A. J. G., and Barbara, P. F. (1985) Proton-transfer kinetics of 3-hydroxyflavone: Solvent effects. *J. Phys. Chem. 89*, 2355–2361.
(25) McMorrow, D., and Kasha, M. (1984) Intramolecular excited-state proton transfer in 3-hydroxyflavone. Hydrogen-bonding solvent perturbations. *J. Phys. Chem. 88*, 2235–2243.
(26) Klymchenko, A. S., and Demchenko, A. P. (2004) 3-Hydroxychromone dyes exhibiting excited-state intramolecular proton transfer in water with efficient two-band fluorescence. *New J. Chem. 28*, 687–692.
(27) Kuciak, M., Gabus, C., Ivanyi-Nagy, R., Semrad, K., Storchak, R., Chaloin, O., Muller, S., Mely, Y., and Darlix, J. L. (2008) The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities in vitro. *Nucleic Acids Res. 36*, 3389–3400.
(28) Boudier, C., Storchak, R., Sharma, K. K., Didier, P., Follenius-Wund, A., Muller, S., Darlix, J. L., and Mely, Y. (2010) The mechanism of HIV-1 Tat-directed nucleic acid annealing supports its role in reverse transcription. *J. Mol. Biol. 400*, 487–501.
(29) Pfister, J. R., Wymann, W. E., Schuler, M. E., and Roszkowski, A. P. (1980) Inhibition of histamine-induced gastric secretion by flavone-6-carboxylic acids. *J. Med. Chem. 23*, 335–338.
(30) Eastman, J. W. (1967) Quantitative spectrofluorimetry - the fluorescence quantum yield of quinine sulfate. *Photochem. Photobiol. 6*, 55–72.
(31) Doroshenko, A. O., Sychevskaya, L. B., Grygorovych, A. V., and Pivovarenko, V. G. (2002) Fluorescence probing of cell membranes with azacrown substituted ketocyanine dyes. *J. Fluoresc. 12*, 455–464.
(32) Reichardt, C. (1994) Solvatochromic dyes as solvent polarity indicators. *Chem. Rev. 94*, 2319–2358.
(33) Egele, C., Barbier, P., Didier, P., Piemont, E., Allegro, D., Chaloin, O., Muller, S., Peyrot, V., and Mely, Y. (2008) Modulation of microtubule assembly by the HIV-1 Tat protein is strongly dependent on zinc binding to Tat. *Retrovirology 5*, 62.
(34) 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.
(35) 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.
(36) Shojania, S., and O'Neil, J. D. (2006) HIV-1 Tat is a natively unfolded protein: The solution conformation and dynamics of reduced HIV-1 Tat-(1−72) by NMR spectroscopy. *J. Biol. Chem. 281*, 8347–8356.

BC100434D

Supporting information

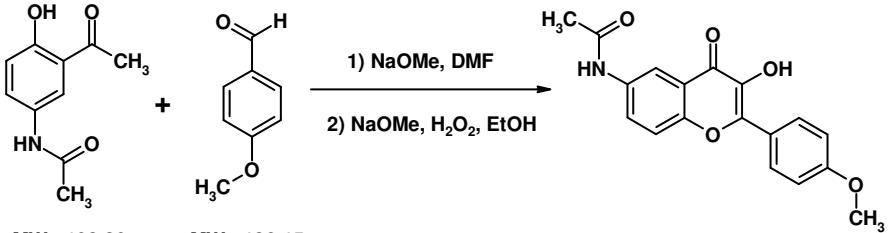
IMPROVED HYDRATION-SENSITIVE DUAL-FLUORESCENCE LABELS FOR MONITORING PEPTIDE-DNA INTERACTIONS

Oleksandr M. Zamotaiev, Viktoriia Y. Postupalenko, Volodymyr V. Shvadchak, Vasyl G. Pivovarenko, Andrey S. Klymchenko, Yves Mély

1. Synthesis of F6A and F4O labels.

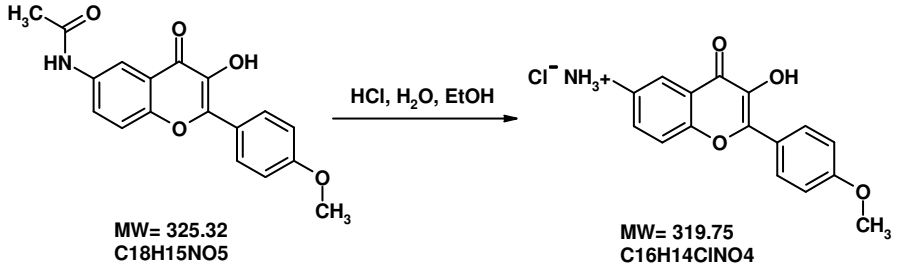
Synthesis of F6A.

1) *N*-[3-hydroxy-2-(4-methoxyphenyl)-4-oxo-4*H*-chromen-6-yl]acetamide



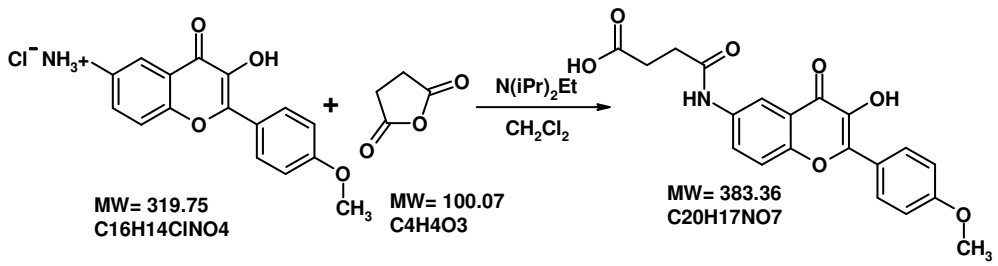
0.4 g (2.07 mmol) of 2-hydroxy-5-acetylamino-acetophenone and 0.276 g (2.27 mmol) of 4-methoxybenzaldehyde were dissolved in 5 ml of dry DMF in Ar-atmosphere. At RT, 336 mg (6.22 mmol) of sodium methoxide were added to the reaction mixture, under stirring. The solution was stirred at RT during 6 h, then diluted with 30 ml of dry ethanol. To the obtained solution, 1.68 g (31.1 mmol) of NaOMe and then 2.8 ml of 30%-H_2O_2 were added. The mixture was gently boiled during 10 min, and left stirring to cool to rt. The obtained yellow solution was poured into 100 ml of water, acidified to pH 6.5 with 10%-HCl and filtered. The precipitate was washed twice with water and dried. Both synthesis steps were monitored by TLC in CH_2Cl_2:MeOH (92:8). Yield 340 mg (50.5 %), yellow powder. ^{1}H-NMR (300 MHz, $CDCl_3$): δ 2.08 (s, 3H, NHCO**C**$\mathbf{H_3}$), 3.84 (s, 3H, O**C**$\mathbf{H_3}$), 7.12 (d, 2H, **ArH**, J = 8.2 Hz), 7.70 (d, 1H, **ArH**, J = 8.4 Hz), 7.87 (d, 1H, **ArH**, J = 8.4 Hz), 8.18 (d, 2H, **ArH**, J = 8.2 Hz), 8.41 (s, 1H, **ArH**), 9.39 (s, 1H, Ar**OH**), 10.23 (s, 1H, **NH**).

2) 3-hydroxy-2-(4-methoxyphenyl)-4-oxo-4*H*-chromen-6-aminium chloride



150 mg (0.461 mmol) of amide were dissolved in a mixture of 6 ml of water, 4 ml of 37% HCl and 10 ml of ethanol. The reaction mixture was refluxed during 1 h, and cooled. The yellow precipitate was filtered, washed with water and dried. Yield 123 mg (83.4 %). ^{1}H-NMR (300 MHz, DMSO- d_6): δ 3.84 (s, 3H, O**C**$\mathbf{H_3}$), 7.12 (d, 2H, Ar**H**, J = 9.1 Hz), 7.33 (dd, 1H, Ar**H,** J = 9 Hz, J = 2.1 Hz), 7.51 (d, 1H, Ar**H**, J = 2.1 Hz), 7.63 (d, 1H, Ar**H**, J = 9 Hz), 8.17 (d, 2H, Ar**H**, J = 9.1 Hz), 9.31 (br. s, 1H).

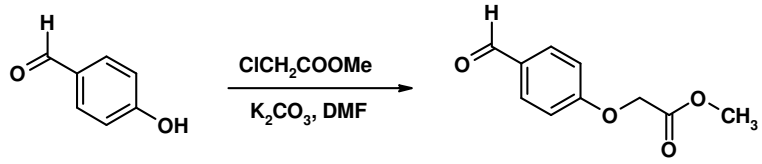
3) 4-{[3-hydroxy-2-(4-methoxyphenyl)-4-oxo-4*H*-chromen-6-yl]amino}-4-oxobutanoic acid (F6A)



50 mg (0.156 mmol) of amine hydrochloride, 0.055 ml (0.316 mmol) of DIEA and 40 mg (0.4 mmol) of succinic anhydride were dissolved in 30 ml of CH_2Cl_2. The solution was refluxed for 12 h and then cooled to RT. Flaky yellow precipitate was filtered, washed twice with CH_2Cl_2 and dried. Reaction was monitored by TLC in CH_2Cl_2:MeOH (85:15). Yield 46 mg (76.9 %). ^{1}H-NMR (300 MHz, DMSO-d_6): δ 2.57 (t, 2H, **C**$\mathbf{H_2}$), 3.32 (t, 2H, N**C**$\mathbf{H_2}$), 3.84 (s, 3H, O**C**$\mathbf{H_3}$), 7.12 (d, 2H, J = 9 Hz, Ar**H**), 7.71 (d, 1H, J = 9 Hz, Ar**H**), 7.86 (dd, 1H, J = 9 Hz, J = 2.2 Hz, Ar**H**), 8.18 (d, 2H, J = 9 Hz, Ar**H**), 8.44 (d, 1H, J = 2.2 Hz, Ar**H**), 9.42 (s, 1H, **OH**), 10.27 (s, 1H. **NH**), 12.12 (br. s., 1H, **COOH**). *m/z* ($M+H^+$) calculated for $C_{20}H_{18}NO_7$: 384.1; found: 384.2.

Synthesis of F4O.

1) Methyl (4-formylphenoxy)acetate



To a mixture of 1.22 g (10 mmol) of aldehyde and 1 g (7.24 mmol) of dry K_2CO_3 in 10 ml of dry DMF, 1.04 ml (11 mmol) of methylbromoacetate was added. Reaction mixture was stirred and heated to 95 ºC during 5 h. Reaction was monitored by TLC, CH_2Cl_2:MeOH (92:8). Then, the solvent was evaporated in vacuum and 10 ml of water was added to the obtained oil. After the oil was crystallized, the precipitate was filtered and washed with water and heptane. Yield 1.52 g (78.3 %). ^{1}H-NMR (300 MHz, $CDCl_3$): δ 3.83 (s, 3H, O**C**$\mathbf{H_3}$), 4.73 (s, 2H, $\mathbf{CH_2}$), 7.01 (d, 2H, J = 8.8 Hz, Ar**H**), 7.85 (d, 2H, J = 8.8 Hz, Ar**H**), 9.91 (s, 1H, CO**H**).

2) [4-(3-hydroxy-4-oxo-4*H*-chromen-2-yl)phenoxy]acetic acid (F4O)

0.282 ml (2.34 mmol) of 2-hydroxyacetophenone and 0.5 g (2.58 mmol) of aldehyde were dissolved in 5 ml of dry DMF in Ar-atmosphere. At RT, 0.38 g (7.04 mmol) of sodium methoxide was added to the reaction mixture, under stirring. The orange-red mixture was stirred at RT during 10 h, and then diluted with 25 ml of dry ethanol. To the obtained solution, 1.9 g (35.2 mmol) of NaOMe and then, 2.85 ml of 30%-H_2O_2 were added. The mixture was boiled during 10 min and then left stirring to cool to RT. Obtained yellow solution was poured into 100 ml of water, acidified to pH 6.5 with 10%-HCl and filtered. The yellow precipitate was washed twice with water and dried. The obtained substance was dissolved in a 1:1 mixture of i-PrOH and water with few drops of conc. HCl. The mixture was boiled and after complete dissolution, it was evaporated. Yield 50 mg (6.84 %), light yellow powder. ^{1}H-NMR (300 MHz, DMSO-d_6): δ 4.79 (s, 2H, $\mathbf{CH_2}$), 7.11 (d, 2H, J = 8.8Hz, Ar**H**), 7.46 (t, 1H, J = 7.2Hz, Ar**H**), 7.76 (m, 2H, Ar**H**), 8.10 (d, 1H, J = 7.4Hz, Ar**H**), 8.18 (d, 2H, J = 8.8Hz, Ar**H**). *m/z* ($M+H^+$) calculated for $C_{17}H_{13}O_6$: 313.0; found: 313.0.

2. Spectroscopic properties of the labels.

Table S1. Spectroscopic properties of FC, F6A, F6C and F4O labels.

Media	E_T30	Label	λ_{Abs}, nm	λ_{N*}, nm	λ_{T*}, nm	N*/T*	QY, %
Water	63.1	FC	357	431	514	1.47	2.1
		F6A	351	443	514	11	10.6
		F6C	355	442	517	13	8.1
		F4O	353	437	516	22	6.6
MeOH	55.4	FC	349	423	533	0.81	5.8
		F6A	349	429	533	1.10	6.1
		F6C	354	428	535	0.97	4.9
		F4O	356	432	529	2.78	4.5
EtOH	51.9	FC	351	420	535	0.34	6.3
		F6A	351	424	535	0.28	6.5
		F6C	356	419	533	0.23	5.1
		F4O	359	430	535	1.25	3.6
DMF	43.2	FC	344	421	539	0.10	7.1
		F6A	345	411	539	0.15	4.8
		F6C	354	418	544	0.13	5.2
		F4O	356	432	540	0.30	2.7
MF	54.1	FC	353	427	537	0.37	7.6
		F6A	351	429	535	0.47	8.1
		F6C	356	426	536	0.52	6.8
		F4O	357	430	536	0.99	4.9
Acetonitrile	45.6	FC	344	413	533	0.06	2.8
		F6A	344	419	531	0.03	5.0
		F6C	352	422	538	0.05	4.5
		F4O	347	409	529	0.03	4.0
THF	37.4	FC	339	408	537	0.07	5.5
		F6A	338	409	538	0.02	6.3
		F6C	355	411	544	0.04	9.5
		F4O	350	406	537	0.04	6.7
Dioxane	36	FC	340	411	537	0.06	6.2
		F6A	339	412	537	0.02	6.4
		F6C	354	408	541	0.02	11.1
		F4O	350	405	534	0.03	8.5

E_T30 – empiric polarity index, λ_{abs} – position of absorption maxima, λ_{N*} and λ_{T*} - position of fluorescence maxima of N* and T* forms respectively. N*/T* - ratio of the intensities of the two emission bands at their peak maxima, QY- fluorescence quantum yield.

3. Characterization of labeled peptides.

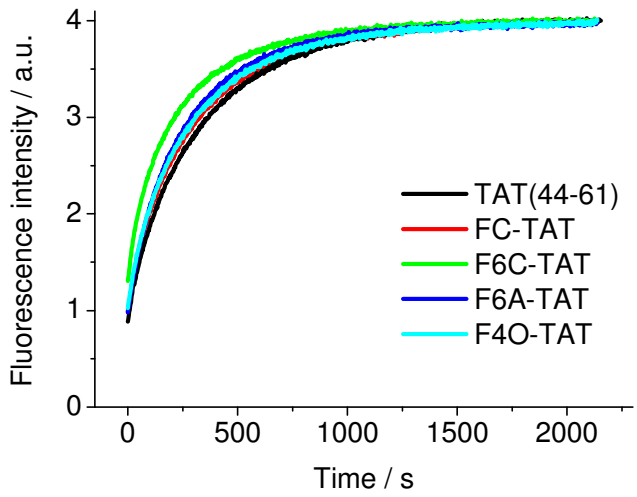


Figure S1. Nucleic acid chaperone properties of the native and FC-, F6C-, F6A-, and F4O- labeled Tat(44-61) peptides. The annealing kinetics of 10 nM TMR-5′-cTAR-3′-Fl with 300 nM dTAR in the presence of the unlabelled and labeled Tat(44-61) peptides added at a protein:ODN ratio of 2 were monitored in 25 mM Tris (pH 7.5), 30 mM NaCl and 0.2 mM $MgCl_2$. Excitation and emission wavelengths were 480 and 520 nm, respectively.

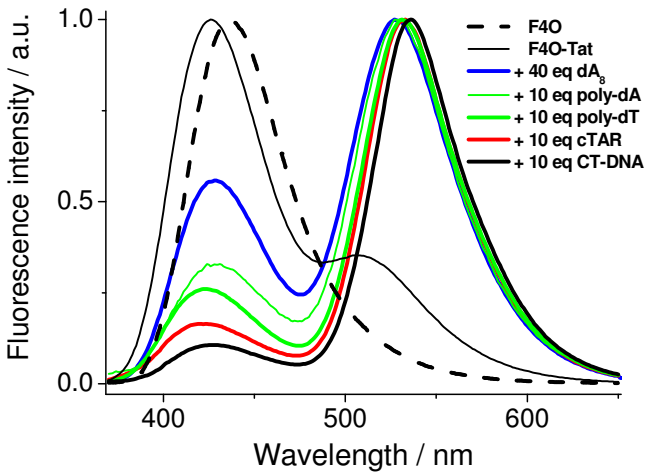


Figure S2. Fluorescence spectra of the F4O-Tat(44-61) peptide complexed to various oligonucleotides. The spectra of the labeled peptides were recorded in the presence of dA_8, poly-dA, poly-dT, cTAR and the double-stranded CT-DNA. All spectra were obtained in 10 mM phosphate buffer pH 7.0, 30 mM NaCl. The spectra of the labeled peptide (black solid) and the label (black dotted) in their free form were recorded for comparison. Concentration of peptide was 1 µM. The ODNs were added at a ratio of 10 bases or base pairs per peptide.

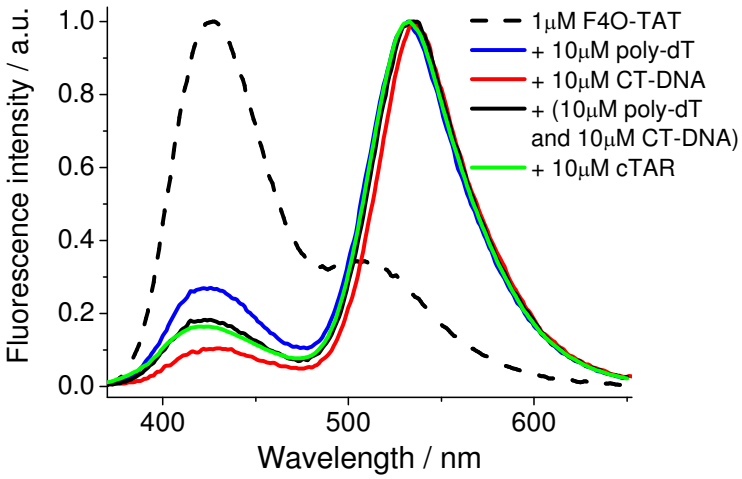


Figure S3. Fluorescence spectra of F4O- labeled Tat(44-61) in the presence of poly-dT, or double-stranded CT-DNA, or their mixture (1:1). All spectra were obtained in 10 mM phosphate buffer pH 7.0, 30 mM NaCl. Concentration of DNAs is expressed per base and base pairs, for ss and ds DNA, respectively.

Publication 3

Quantification of local water concentration in biomolecules using dual-fluorescence labels

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

Quantification of Local Hydration at the Surface of Biomolecules Using Dual-Fluorescence Labels

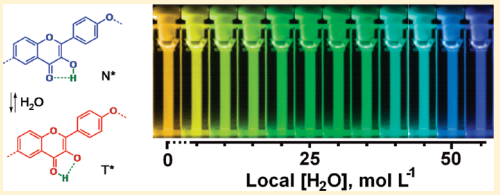
Vasyl G. Pivovarenko,*[,†,‡] Oleksandr M. Zamotaiev,[†] Volodymyr V. Shvadchak,[‡] Viktoriia Y. Postupalenko,[‡] Andrey S. Klymchenko,[‡] and Yves Mély[‡]

[†]Organic Chemistry Chair, Chemistry Faculty, Taras Shevchenko National University of Kyiv, 01601 Kyiv, Ukraine

[‡]Laboratoire de Biophotonique et Pharmacologie, UMR 7213 du CNRS, Université de Strasbourg, Faculté de Pharmacie, 74 route du Rhin, 67401 Illkirch, France

Supporting Information

ABSTRACT: By using four labels of the 3-hydroxyflavone family displaying selective sensitivity to hydrogen bond (HB) donors and poor response to other polar molecules, we developed an approach for measuring local water concentration $[H_2O]_L$ (or partial volume of water: $W_A = [H_2O]_L/55.6$) in the label surrounding both in solvent mixtures and in biomolecules by the intensity ratio of two emissive forms of the label, N^*/T^*. Using a series of binary water/solvent mixtures with limited preferential solvation effects, a linear dependence of $\log(N^*/T^*)$ on the local concentration of HB donor was obtained and then used as a calibration curve for estimating the W_A values in the surroundings of the probes conjugated to biomolecules. By this approach, we estimated the hydration of the labels in different peptides and their complexes with DNAs. We found that W_A values for the label at the peptide N-terminus are lower (0.63−0.91) than for free labels and depend strongly on the nature of the N-terminal amino acid. When complexed with different DNAs, the estimated hydration of the labels conjugated to the labeled peptides was much lower (W_A = 0−0.47) and depended on the DNA nature and linker-label structure. Thus, the elaborated method allows a site-specific evaluation of hydration at the surface of a biomolecule through the determination of the partial volume of water. We believe the developed procedure can be successfully applied for monitoring hydration at the surface of any biomolecule or nanostructure.



INTRODUCTION

Water plays a large variety of roles in biological systems at different levels of complexity, from molecules and cell organelles to tissues.[1−4] As an essential functional component, water is always present in the interior of biomolecules. At the surface of biomolecules, such as proteins, nucleic acids, or lipid aggregates, the properties of water change dramatically depending on the chemical nature of the solvated segment and the water−segment distance.[1,5−8]

Until now, main attention was paid to the study of water located in the interior of biomolecules or at the surface of lipid and protein assemblies.[1−4,9−16] IR, near-IR vibrational spectroscopy,[12,13] terahertz spectroscopy,[5] neutron and X-ray scattering,[12] and magnetic resonance methods[10] are usually applied in these studies.[1,17−25] Electron spin resonance (ESR) of molecular aggregates incorporating spin-labeled lipid units was shown to be a powerful approach[24] for determining hydrogen-bonding with water at various locations in the fluid membrane.[17,24,26]

Fluorescence probing is also an important tool for characterizing the hydration of biopolymers and its changes during biological processes.[1,4,27−32] For instance, polarity-sensitive fluorescent probes efficiently report on the hydration of supramolecular assemblies,[4] while Prodan,[28−31] phthalimide or naphthalimide,[28,32−35] and Nile Red[28,36] derivatives are common tools for lipid membranes[30,31,36] and for monitoring peptide binding to nonpolar proteins.[28,32−36] However, applications of polarity sensors to hydration studies are limited to relatively nonpolar systems, where the response of these dyes to polarity is high. In polar media, their fluorescence is quenched at higher water concentration[34] or is poorly sensitive to water.

3-Hydroxychromones (3HC) and, particularly, 3-hydroxyflavones (3HFs) appear as very attractive tools for studying hydration of biomolecules and their complexes.[37−56] These probes undergo excited-state proton transfer (ESIPT) that results in two emission bands. The rate of the ESIPT reaction and therefore the relative intensity of the normal (N^*) and tautomeric (T^*) bands depends on the properties of their molecular surroundings. The sensitivity of 3HF dyes to protic solvents is mainly due to H-bond formation with the groups involved in the ESIPT reaction (3-OH and 4-keto groups).[57,58] Because of their extreme sensitivity to traces of water in organic medium,[38,39] 4′-amino derivatives of 3HFs were used for the evaluation of hydration of lipid membranes at different locations.[40,41,45−47,49,50] Recently, improved 3HC derivatives were synthesized and successfully applied for site-selectively investigating the physicochemical properties, and notably, the

Received: October 23, 2011
Revised: February 15, 2012
Published: March 6, 2012

polarity and hydration of lipid membranes, proteins, and nucleic acids.[43,48,50−55]

However, the quantification of the local water content $[H_2O]_L$ at a given location of a biomolecular system using fluorescent probes has been never addressed so far. For most hydration-sensitive probes, quantification of $[H_2O]_L$ is a complex problem as a result of the overlapping responses of the probes to water and various parameters of their environment (e.g., polarity) as well as to their interactions with surrounding molecules and ions. Thus, to quantify the local water content, the chemical structure of the fluorescent probes needs to be optimized so that H-bonding dominates over the effects of dipole−dipole and ionic interactions and over the other physicochemical properties of the surroundings. Recently, a step toward this aim was achieved by the design of four 3HF probes that show strong sensitivity to hydration changes and negligible response to polarity and basicity.[54]

In this work, we propose a method for measuring the local water content at the labeling site of peptides, based on the ratiometric response of 3HF fluorophores and the application of this method for comparing the hydration levels of peptides in their free state and in peptide−DNA complexes.

MATERIALS AND METHODS

Reagents and Solvents. All the reagents were purchased from Sigma-Aldrich. Solvents for synthesis were of reagent quality and were appropriately dried if necessary. For absorption and fluorescence studies, the solvents were of spectroscopic grade. Calf thymus DNA (CT-DNA), cTAR, and single-stranded DNA poly(dA) and poly(dT) were from Sigma.

Synthesis of FC, F6A, and F4O compounds were published elsewhere.[54,55] Lipophilicity coefficients log *P* of the applied probes were calculated by ACD/Laboratories software. To more accurately measure the hydration parameters of the peptides labeled by the F6C probe, the amide derivative of free probe (F6Ca) was synthesized (see Supporting Information) and characterized spectroscopically.

3-Hydroxy-2-(4-methoxyphenyl)-4-oxo-4*H*-6-chromenecarboxamide (F6Ca). ^{1}H-NMR (400 MHz, DMSO-d_6). Abbreviations: s, singlet; d, doublet; m, multiplet; brs, broad singlet. δ, ppm: 3.87 (s, 3H, O**CH**$_3$), 7.05 (d, 2H, J = 8.8 Hz, Ar**H**), 7.31 (brs, 1H, N**H**), 7.70 (d, 1H, J = 8.8 Hz, Ar**H**), 8.20 (m, 4H, Ar**H**), 8.67 (brs, 1H, N**H**), 9.35 (brs, 1H, **OH**). Calc for $C_{17}H_{14}NO_5^+$ m/z ($M + H^+$): 312.087. Found: 312.087.

Peptide Synthesis. All peptides were synthesized by solid phase peptide synthesis on a 433A synthesizer (ABI, Foster City, CA) as published elsewhere.[54,60]

Sample Preparation. To the solution of probe (~2×10^{-6} mol L^{-1}) in a chosen organic solvent, aqueous solutions with equal probe concentration were added in the quantities needed for obtaining final water concentrations of 4, 8, 11, 14, 18, 22, 25, 27.5, and 30 mol L^{-1}. A second series was obtained by mixing the probe solution in organic solvent with its aqueous solution in the quantities needed for obtaining final water concentrations of 53, 51, 49, 47, 45, 42, 37, 33, 30, and 27.5 mol L^{-1}. For probes F6Ca and FC, a more close-packed set of water concentrations was applied. Fluorescence spectra of the obtained solutions including those in neat water and neat solvent were measured at 25 °C, immediately after preparation, using an excitation wavelength corresponding to the maximum of the absorption band.

Instrumentation. Proton NMR spectra were recorded on a 300 MHz Bruker spectrometer, and mass spectra were recorded on a Mariner System 5155 mass spectrometer using the electrospray ionization (ESI) method. All column chromatography experiments were performed on silica gel (Merck, Kieselgel 60H, Art 7736). Absorption and fluorescence spectra were recorded on a Cary 400 spectrophotometer (Varian) and FluoroMax 3.0 spectrofluorimeter (Jobin Yvon, Horiba), respectively. For fluorescence studies, the dyes were used at 0.5 to 1 × 10^{-6} mol L^{-1} concentrations. Excitation wavelength was on the absorption maximum for each label. Fluorescence quantum yields were determined using quinine sulfate in 0.5 mol L^{-1} sulfuric acid, as a reference (quantum yield, φ = 0.57).[59] For the experiments in water, 10^{-2} mol L^{-1} phosphate buffer containing 3 × 10^{-2} mol L^{-1} NaCl (pH 7.0) was used systematically. All spectroscopic data for labeled peptides and their complexes with DNA are from the literature.[54,55,60]

Deconvolution of Spectra. Deconvolution of the fluorescence spectra into *N** and *T** bands was performed using the Siano software kindly provided by Professor A. O. Doroshenko (Kharkiv, Ukraine), as previously described.[61] The program is based on an iterative nonlinear least-squares method, where the individual emission bands were approximated by a log-normal function accounting for several parameters: maximal amplitude, I_{max}, spectral maximum position, ν_{max}, and position of half-maximum amplitudes, ν_1 and ν_2, for the blue and red parts of the band, respectively. These parameters determine the shape parameters of the log-normal function, namely, the full width at the half-maximum, fwhm = $\nu_1 - \nu_2$, and the band asymmetry, $P = (\nu_1 - \nu_{max})/(\nu_{max} - \nu_2)$. All parameters were allowed to vary in the iteration process, which allowed a good correlation of the separated bands with the initial spectrum (number of repetitions: ≥6; $R \geq 0.998$ for the experiments with peptides and DNAs, and $R \geq 0.9998$ for solvent mixtures. See Figures 1C and S2−S5 in the Supporting Information). In the case of spectra from neat water, five separately obtained spectra were subjected to the deconvolution procedure: $R \geq 0.9$. In this case, the fwhm parameter of the *T** band was taken from the data of appropriate binary solvent mixtures and then fixed upon deconvolution. The resulting integral fluorescence intensities of the separated *N** and *T** bands were used for calculation of the *N**/*T** ratio.

RESULTS AND DISCUSSION

Selected Probes and Mechanism of Water Sensing. In contrast to numerous dialkylamino-derivatives of 3HFs[39−42,45−47,49,50] that show strong sensitivity to both nonspecific dipolar and specific H-bonding interactions, the four 3HF derivatives recently synthesized by us (Scheme 1) are mainly sensitive to H-bonding interactions.[54,55] Indeed, for the four probes, it has been shown that H-bonding with the solvent, but not the solvent polarity, governs their *N**/*T** ratio and the position of their *T** band.[54,55] Moreover, their fluorescence properties were also found to be poorly dependent on the pH[54,56] and ionic strength (Figure 1A). Thus, the four 3HF derivatives appear as suitable candidates for measuring local water concentrations in biomolecules. Since the four compounds are characterized by different linker length and position, it will be possible to select the most appropriate ones for labeling a given biomolecule and investigate its interaction with target molecules.

In biological media, water molecules are the main H-bond (HB) donors. Assuming that the mechanism of the 3HF response to water molecules is the same when the probe is

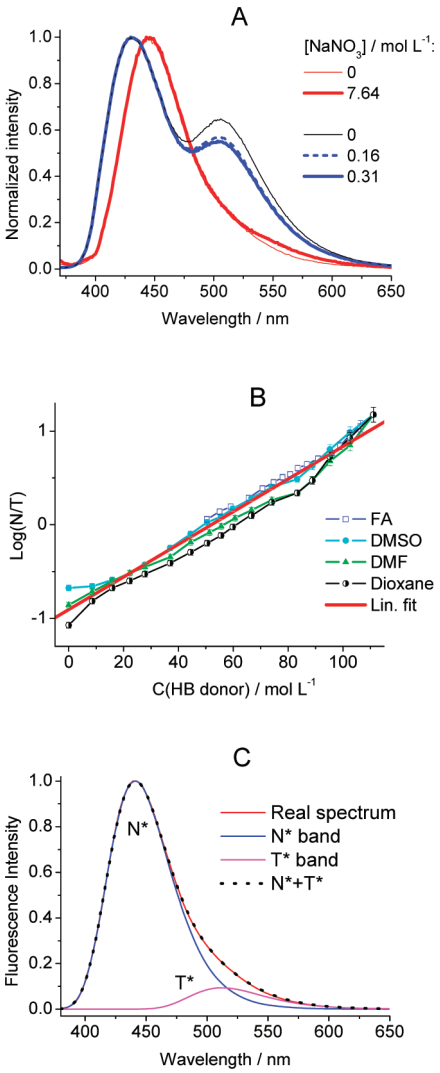
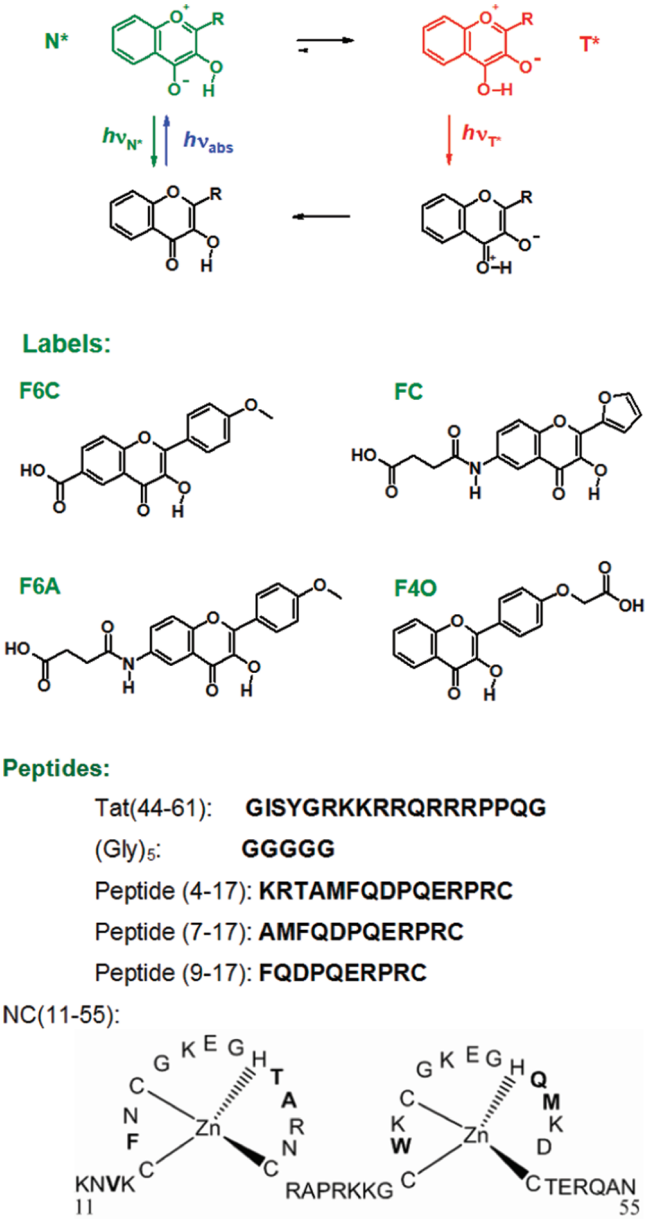


Figure 1. Dependence of the emission of 3HF probes on ionic strength and the molar concentration of hydrogen bond donor. (A) Influence of ionic strength on the emission of F6Ca (red) and FC (blue) probes in 10 mM phosphate buffer, pH 7.0. Excitation wavelength was fixed at the absorption maximum in each case. (B) Dependence of $\log(N^*/T^*)$ versus molar concentration of hydrogen bond donor in water–solvent mixtures for F6Ca probe. Since water and formamide possess two HB donor groups, we use $C_{HBdonor} = 2C_{solvent}$ for these solvents. (C) Example of deconvolution of the spectrum of F6Ca probe in water in two bands.

bound to biomolecules than when it is free, the probe response will result from the interaction of the 3HF atoms involved in ESIPT with water molecules. The mechanism is depicted in Scheme 2 (partially taken from ref 57). Through H-bonding with the 4-carbonyl group of 3HF dyes, water molecules weaken the intramolecular hydrogen bond and thus slow down or even block the intramolecular proton transfer.[57,58] Therefore, when bound to a biomolecule, an increasing exposure of the probe to water molecules within the solvation shell should increase the number of H-bonds, resulting in an increase of the N^*/T^* ratio and a blue shift of the T^* form.[54,58] Interference of HB-donating groups on the biomolecule (NH, OH, etc.) in the measurements of water concentration is unlikely, especially at the biomolecule surface, due to their relatively small concentration in respect to water.

Scheme 1. ESIPT Reaction in the 3HF Dyes and Structure of the Studied Labels and Labeled Peptides

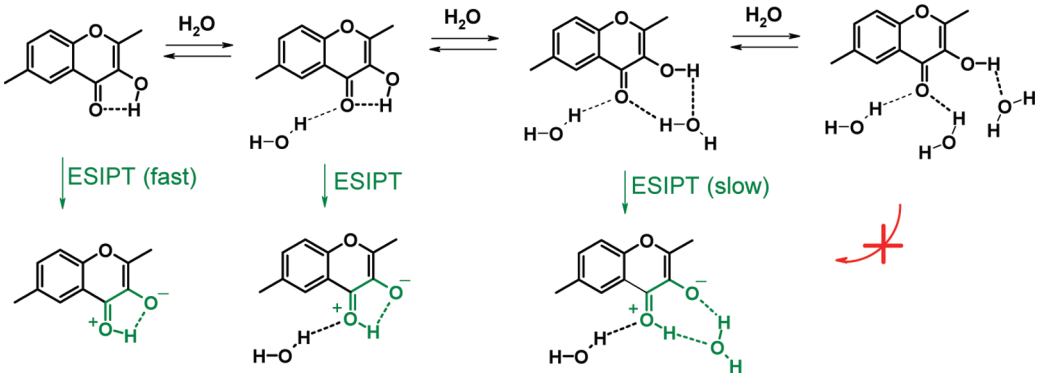


Calibration of Probe Response to H-Bond Donors. To calibrate the probe response to local water concentration $[H_2O]_L$, we measured the energy-related spectroscopic parameters of all labels in binary solvent–water mixtures and neat solvents. To ensure a more precise calibration for the peptide-bound label F6C, we transformed the F6C probe into its amide F6Ca, which is chemically closer to the form of the probe attached to the peptides. For other 3HCs, this modification should not play any role as their carboxylic group is electronically isolated from the fluorophore.

The $\log(N^*/T^*)$ parameter of the probe fluorescence was found appropriate for quantification of water concentration in the probe surroundings. In Figure 1B, as an example, the $\log(N^*/T^*)$ parameter of F6Ca is shown as a function of the total molar concentration of HB donors in the mixture. The log-scale was used in order to convert the N^*/T^* ratio into a parameter linearly correlating with the solvent concentration.[39] Linear correlation in such coordinates displays that the N^*/T^*

dx.doi.org/10.1021/jp2101732 | J. Phys. Chem. A 2012, 116, 3103–3109

Scheme 2. Chemical Mechanism of Water Sensing by the 3HF Probes[a]



[a]Increasing levels of exposure to water of the 3HF probe bound to a biomolecule is thought to result in an increase in the number of H-bonds with the water molecules within the solvation shell.

parameter, the ratio of the two emissive populations of the probe, is driven by the chemical potential of water in the applied systems. The $\log(N^*/T^*)$ parameter according to the Boltzmann equation is proportional to the energy difference between the states, when they are in equilibrium. This is the case for 3HF probes in organic solvents[62] and should be also true for water solvates, for which relaxation times are much shorter.[1,4,7]

A linear correlation was observed in mixtures of water with formamide, dioxane, DMF, and DMSO, suggesting that the chemical potential of water (or water activity,[63] more roughly, water concentration) drives the N^*/T^* ratio of the emitting tautomers. The linearity also shows that the local concentration of water $[H_2O]_L$ in the probe solvation shell corresponds to the total concentration of water $[H_2O]_T$ in the mixtures.

In contrast, for some other solvent mixtures, the water concentration in the probe solvation shell differs from that in bulk solution since the concentration of each species in the probe solvation shell depends on its energy of interaction with the probe molecule. As a result, the energy-related parameters of probe fluorescence show nonlinear dependence upon the total concentration of the active component in the mixture.[64] Such phenomenon is known as the preferential solvation effect.[64,65] Such preferential solvation of the probe by water or by the organic solvent result in substantial upward and downward deviations from the linearity in the dependence of $\log(N^*/T^*)$ as a function of the total water concentration (Supporting Information, Figure S4). For instance, the strong HB-donor water molecules dominate over acetonitrile in the solvation shell (Supporting Information, Figure S4 upward deviations), while lipophilic 2-propanol and, especially, tetrahydrofurane dominate over water (Supporting Information, Figure S4, downward deviations). The level and sign of the observed deviations are in line with the chemical properties of solvent (e.g., its lipophilicity and HB-donating and HB-accepting coefficients, Table S2, Supporting Information). Upward deviations were also observed for binary mixtures of methanol, methylformamide (Figure S4, Supporting Information), and trifluoroethanol (not shown), whose aggregates with water possess increased HB-donating ability.

On the basis of the four water–solvent mixtures providing a linear dependence of $\log(N^*/T^*)$, a linear calibration curve was determined: $\log(N^*/T^*) = 0.0164[H_2O]_L - 0.695$ ($R = 0.978$) (Figure 1B, red line). At high water concentrations, a preferential solvation effect gives small deviations from the data point in neat water. Accordingly, these data points were removed from the calibration curve and substituted by the data point in neat water.

Next, using the same approach, we determined the calibration curves for the three other probes and found out that the calibration curves for F6A, F4O, and F6Ca were similar, while that for FC exhibited a nearly 2-fold smaller slope (Figure 2), indicating that FC is less sensitive to hydration.

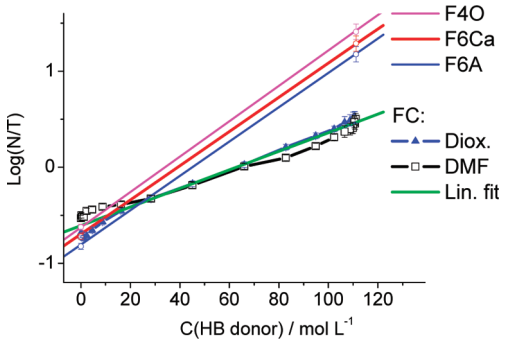


Figure 2. Dependence of $\log(N^*/T^*)$ on the concentration of hydrogen bond donor for the four 3HC derivatives. Fitted lines: F4O, $\log(N^*/T^*) = 0.0184[H_2O]_L - 0.627$; F6Ca, $\log(N^*/T^*) = 0.0177[H_2O]_L - 0.706$; F6A, $\log(N^*/T^*) = 0.0179[H_2O]_L - 0.805$; FC, $\log(N^*/T^*) = 0.0100[H_2O]_L - 0.607$.

Comparing the calibration curves with the titration curves in Figure 1, it is quite obvious that the error in the estimation of hydration will be much higher at low hydration levels, where the solvent effects other than hydration play an important role. The data points for polar (DMF and DMSO) and low polar solvents (e.g., dioxane) are close at water concentrations higher than 5 mol L^{-1}, which suggests a 5 mol L^{-1} limit for the precise estimation of water concentrations. Below this limit, non-specific dipolar interactions (i.e., solvent polarity) start interfering with the analysis. This concentration range is likely appropriate for proteins and oligonucleotides as well as their complexes, but not for lipid membrane sites, where the water concentration is usually below this range.[11,12]

From the mechanism presented in Scheme 2, it is obvious that the fluorescent probe senses only water in its close proximity, which allows quantifying the water concentration in the probe solvation shell. Since the probe response was related to the total concentration of water in our calibration procedure,

dx.doi.org/10.1021/jp2101732 | J. Phys. Chem. A 2012, 116, 3103–3109

we obtain in fact an average value of the local water concentration $[H_2O]_L$ in the probe solvation shell, that depends on the average location, orientation dynamics, and fluorescence lifetime of all peptide-conjugated probes in solution.

Hydration Measurements in Labeled Peptides and Their Complexes with Oligonucleotides. We already reported[54] that the Tat(44−61) peptide, corresponding to the basic domain of the HIV-1 protein,[66] labeled by the applied probes, showed a decrease of their N^*/T^* ratio (Table 1), when interacting with DNA sequences of various lengths and strandedness. The decrease in N^*/T^* ratio appears as a result of possible stacking of the probes with nucleic bases, which leads to a decrease of the probe hydration. Now, we show the possibility to quantify the hydration level at the location of the probe. As an application, we calculated the hydration of the FC probe bound to the N-terminus of a pentaglycine peptide,[55] the zinc finger domain of the HIV-1 nucleocapsid protein (NC),[55,67] and several peptide sequences (4−17, 7−17, and 9−17),[60] which contain the binding site recognized by an antibody fragment, $scFv1F4_{Q34S}$.

Table 1. Intensity Ratios of Fluorescence Bands and Hydration Parameters for the Applied Probes in Their Peptide Conjugates and Peptide−DNA Complexes[a]

	N^*/T^* ratio	$[H_2O]_L$	W_A
F6Ca (1.4)[b]	18	55.5	1.0
F6C−Tat	4.64	38.3	0.69
F6C−Tat + poly dT	0.44	9.4	0.17
F6C−Tat + CT-DNA	0.32	5.75	0.10
FC (−0.33)[b]	3.2	55.5	1.0
FC−Tat	1.8	43.0	0.77
FC−Tat + poly dT	0.65	20.9	0.38
FC−Tat + poly dA	0.82	26.0	0.47
FC−Tat + CT-DNA	0.51	15.6	0.28
FC−NC	1.7	41.8	0.75
FC−$(Gly)_5$	2.25	48.0	0.86
FC(4−17)	1.7	42.3	0.76
FC(7−17)	2.5	50.3	0.91
FC(9−17)	1.3	36.5	0.66
F6A (0.46)[b]	15	55. 5	1.0
F6A−Tat	6.5	43.0	0.77
F6A−Tat + poly dT	0.53	15.9	0.29
F6A−Tat + CT-DNA	0.32	10.0	0.18
F4O (0.98)[b]	25	55.5	1.0
F4O−Tat	4.7	35.3	0.63
F4O−Tat + poly dT	0.44	7.4	0.13
F4O−Tat + poly dA	0.53	9.6	0.17
F4O−Tat + cTAR	0.29	~2.5	~0.05
F4O−Tat + CT-DNA	0.19	~0	~0

[a] $[H_2O]_L$ is the local water concentration (mol L^{-1}) calculated from the calibration curves in Figure 2. W_A is the local water access coefficient. [b] Logarithm of the distribution coefficient of the amide derivatives of the probes in 1-octanol−water mixture (log P).

For biomolecules and their complexes, it is more convenient to express the level of hydration through the water access coefficient W_A of the probe:

$$W_A = [H_2O]_L/[H_2O] = [H_2O]_L/55.56$$

where $[H_2O]$ is the molar concentration of neat water.

W_A shows the fraction of space in the surrounding of the probe proton-transfer system, occupied by water molecules. W_A is also called the partial volume of water and is frequently used for characterization of hydration.[11,12]

Using this approach, we calculated the hydration level for several peptides in 10^{-2} mol L^{-1} phosphate buffer containing 3 × 10^{-2} mol L^{-1} NaCl (pH 7.0).[54,60] We found that the W_A values of the probes in their conjugates with the N-terminus of the Tat peptides was 20−40% lower than the W_A value of the probe alone (Table 1). The observed W_A values were the lowest (Figure 3A) for the F6C and F4O probes, which present the highest lipophilicity coefficients (log P = 1.4 and 0.98, respectively, Table 1). These rather low W_A values are likely the result of the stronger interaction of these dyes with the more hydrophobic components of the peptide, which reduces the

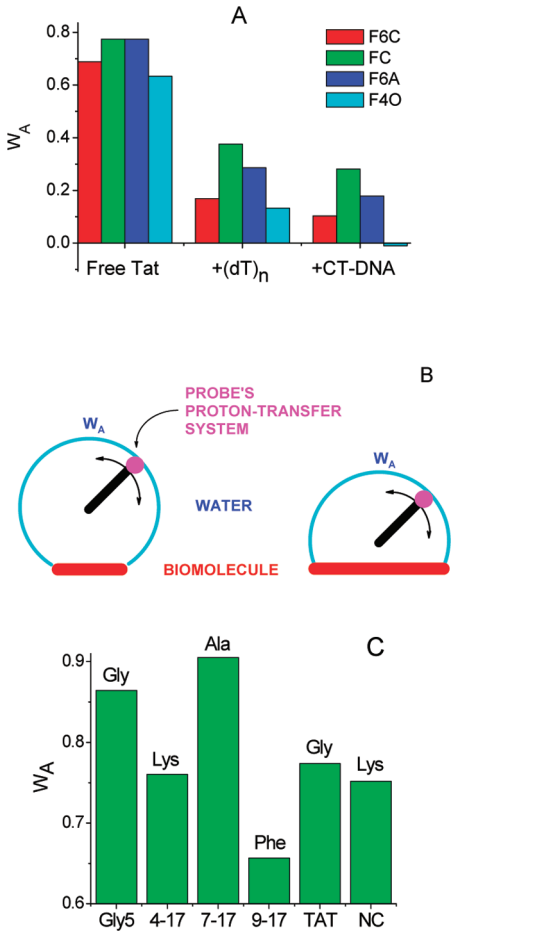


Figure 3. Water access W_A coefficients for peptides labeled by the 3HF probes. (A) The water access for the Tat(44−61) peptide labeled by either the F6C, FC, F6A, and F4O probe was determined in the absence and in the presence of poly dT or CT-DNA. (B) Representation of the decrease of local water access W_A caused by the probe lipophilicity. The rotational motion of the label is represented by an arc, whose length reflects the measured W_A value. The sphere delineated by the arc represents the accessible volume for W_A measurements. A more lipophilic label (right) is thought to interact more strongly than a less lipophilic one (left) with the biomolecule and thus lead to a smaller W_A value. (C) Dependence of the water access on the nature of the peptide and N-terminal amino acid. The W_A coefficients are given for FC-labeled $(Gly)_5$, peptide (4−17), peptide (7−17), peptide (9−17), Tat(44−61), and NC peptide. The nature of the N-terminal amino acid is indicated over the bars.

dx.doi.org/10.1021/jp2101732 | J. Phys. Chem. A 2012, 116, 3103−3109

volume of accessible water (Figure 3B). Moreover, as these probes are connected to the peptide through a shorter linker, their screening from bulk water is likely more efficient.

On the next step, we studied the effect of the peptide sequence on the probe hydration. Remarkably, the W_A values vary significantly with the N-terminally labeled peptide sequence (Figure 3C). Thus, in $(Gly)_5$ and peptide (7−17), the W_A value is the highest, while for peptide (9−17), it is the lowest. However, peptide 4−17, Tat(44−61), and NC show intermediate values of hydration. Comparison of the sequences suggests that the closest amino acid residues influence the label hydration. For instance, the low W_A value for peptide (9−17) can be correlated with the presence of the highly hydrophobic and rather bulky N-terminal Phe amino acid, which likely screens efficiently the label from bulk water, while high W_A values are observed when the N-terminal residue is Gly.

Interaction of the labeled Tat(44−61) peptides with poly dT and poly dA DNA was found to dramatically decrease the hydration level of the probes (Table 1), likely as a consequence of the stacking of the probes with the bases of these single-stranded (*ss*) nucleic acids. The W_A values in poly dT were somewhat lower than in poly dA, indicating that interaction with dT bases provided better screening from the bulk water. A further decrease in the hydration of the probes was observed when the Tat peptides were bound to the double-stranded (ds) CT-DNA (Figure 3A), likely as a result of the intercalation of the probe between the DNA base pairs.[54] The data for F6C, F6A, and FC show a correlation between the W_A and log *P* values, as the probes demonstrate a decrease in the W_A values with the increase of their lipophilicity coefficients. Thus, the intercalation between the bases is more efficient for less polar labels, which results in decreased water access W_A values (Figure 3B).

Interestingly, the W_A value for the F4O probe in the peptide−DNA complexes is much lower than for the other probes, suggesting that this probe has almost no access to water in the peptide−dsDNA complex. The peculiar behavior of this probe in respect to the three others can be rationalized by considering the possible intercalation of the various probes in the peptide−DNA complexes. We could speculate that the structure of the F4O dye favors the deep insertion of the atoms participating to ESIPT within the base pairs, allowing efficient screening from the water (Figure 4). In contrast, for the three other probes, the insertion of the ESIPT atoms within the base pairs is less deep, so that water accessibility is still possible.

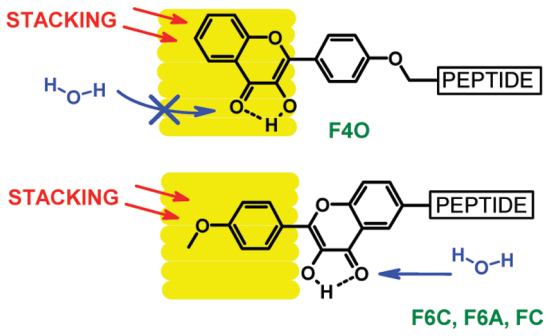


Figure 4. Directions of hydrogen bonding and stacking interactions for the various probes. The interior of DNA base pairs is marked by yellow.

CONCLUSIONS

In this study, we developed a new approach for the estimation of local water concentration using recently developed dual-fluorescence 3HF labels displaying selective sensitivity to HB donors. For these probes in binary solvent mixtures showing no preferential solvation effect, a linear dependence was obtained for $\log(N^*/T^*)$, the logarithm of the emission intensity ratio of two bands versus the local concentration of HB donor. Because of the poor sensitivity of the probes to dipolar molecules, pH, and ionic strength of the medium, the obtained calibration curve allowed us to quantitatively determine the hydration of the label, expressed as the partial volume of water in the label surroundings ($W_A = [H_2O]_L/55.6$). Using this approach, we determine the hydration level of the probes for different N-terminus-labeled peptides and their complexes with DNAs. We found that the hydration level of the labels in peptides was 10−40% lower than for the labels alone. Further comparison of the free labeled peptides shows that the lipophilicity and size of the N-terminal amino acid plays a key role on the label hydration. In complexes of Tat-peptides with ssDNA, the estimated hydration was much smaller, dependent on the label and the DNA bases. Moreover, with dsDNA, the hydration of the label was even lower indicating an efficient screening from water, likely through dye intercalation. We found that the measured value of hydration is dependent on the lipophilicity of the applied label and on the length and position of its connection to the linker since these features govern in part the location and dynamics of the label in the analyzed object. Quantitative analysis of hydration allowed us to further find out that the F4O label showed the lowest level of hydration (~0) in the complex of the Tat-peptide with dsDNA, as a consequence of its deeper insertion between the base pairs. Thus, we found that the water content in the label surrounding at the N-terminus of the peptide depends largely on the peptide nature and can reach almost zero values in complexes with dsDNA, where the label is efficiently screened from water. We believe that the developed procedure can be successfully applied for monitoring hydration of any biomolecule or nanostructure.

ASSOCIATED CONTENT

Supporting Information

Synthesis of F6Ca probe. Spectroscopic properties of F6C, F6Ca, F6A, F4O, and FC probes in different solvents. Dependence of $\log(N^*/T^*)$ versus molar concentration of hydrogen bond donor in water−solvent mixtures. Fluorescence spectra of labeled peptides alone and in the presence of DNAs. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

ARCUS program is acknowledged for financial support. V.G.P. is thankful to Université de Strasbourg for support of a one month stay.

REFERENCES

(1) Zhong, D.; Pal, S. K.; Zewail, A. H. *Chem. Phys. Lett.* **2011**, *503*, 1.
(2) Ball, P. *Chem. Rev.* **2008**, *108*, 74.
(3) Chaplin, M. *Nat. Rev. Mol. Cell. Biol.* **2006**, *7*, 861.

(4) Pal, S. K.; Zewail, A. H. *Chem. Rev.* **2004**, *104*, 2099.
(5) Fayer, M. D. *Acc. Chem. Res.* **2012**, *45*, 3.
(6) Shultz, M. J.; Vu, T. H.; Meyer, B.; Bisson, P. *Acc. Chem. Res.* **2012**, *45*, 15.
(7) Zhang, L.; Yang, Y.; Kao, Y.-T.; Wang, L.; Zhong, D. *J. Am. Chem. Soc.* **2009**, *131*, 10677.
(8) Ebbinghaus, S.; Kim, S. J.; Heyden, Y. X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20749.
(9) Chalikian, T. V.; Macgregor, R. B. *Phys. Life Rev.* **2007**, *4*, 91.
(10) Gawrisch, K.; Gaede, H. C.; Mihailescu, M.; White, S. H. *Eur. Biophys. J.* **2007**, *36*, 281.
(11) Lomidze, A. L.; Pogozheva, I. D.; Mosberg, H. I. *J. Chem. Inf. Model.* **2011**, *51*, 930.
(12) Kucerka, N.; Nagle, J. F.; Sachs, J. N.; Feller, S. E.; Pencer, J.; Jackson, A.; Katsaras, J. *Biophys. J.* **2008**, *95*, 2356.
(13) Marsh, D. *Biophys. J.* **2008**, *94*, 3996.
(14) Bagchi, B. *Chem. Rev.* **2005**, *105*, 3197.
(15) Khoshtariya, D. E.; Hansen, E.; Leecharoen, R.; Walker, G. C. *J. Mol. Liq.* **2003**, *105*, 13.
(16) Waegele, M. M.; Tucker, M. J.; Gai, F. *Chem. Phys. Lett.* **2009**, *478*, 249.
(17) Kurad, D.; Jeschke, G.; Marsh, D. *Biophys. J.* **2003**, *85*, 1025.
(18) Marsh, D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7777.
(19) Marsh, D. *Eur. Biophys. J.* **2002**, *31*, 559.
(20) Furo, I. *J. Mol. Liq.* **2005**, *117*, 117.
(21) Bryant, G.; Pope, J. M.; Wolfe, J. *Eur. Biophys. J.* **1992**, *21*, 223.
(22) Hawton, M. H.; Doane, J. W. *Biophys. J.* **1987**, *52*, 401.
(23) Volke, F.; Pampel, A. *Biophys. J.* **1995**, *68*, 1960.
(24) McCarney, E. R.; Armstrong, B. D.; Kausik, R.; Han, S. *Langmuir* **2008**, *24*, 10062.
(25) Xu, Z. C.; Ellena, J. F.; Cafiso, D. S. *Biophys. J.* **1986**, *49*, A508.
(26) Owenius, R.; Engstrom, M.; Lindgren, M.; Huber, M. *J. Phys. Chem. A* **2001**, *105*, 10967.
(27) Meulen, K. A. V.; Saecker, R. M.; Record, M. T. Jr. *J. Mol. Biol.* **2008**, *377*, 9.
(28) Loving, G. S.; Sainlos, M.; Imperiali, B. *Trends Biotechnol.* **2010**, *28*, 73.
(29) Kucherak, O. A.; Didier, P.; Mely, Y.; Klymchenko, A. S. *J. Phys. Chem. Lett.* **2010**, *1*, 616.
(30) Lee, J. C.-M.; Law, R. J.; Discher, D. E. *Langmuir* **2001**, *17*, 3592.
(31) Zhang, Y.-L.; Frangos, J. A.; Chachisvilis, M. *Biochem. Biophys. Res. Commun.* **2006**, *347*, 838.
(32) Loving, G.; Imperiali, B. *Bioconjugate Chem.* **2009**, *20*, 2133.
(33) Sainlos, M.; Imperiali, B. *Nat. Protoc.* **2007**, *2*, 3210.
(34) Loving, G.; Imperiali, B. *J. Am. Chem. Soc.* **2008**, *130*, 13630.
(35) Venkatraman, P.; Nguyen, T. T.; Sainlos, M.; Bilsel, O.; Chitta, S.; Imperiali, B.; Stern, L. J. *Nat. Chem. Biol.* **2007**, *3*, 222.
(36) Kucherak, O. A.; Oncul, S.; Darwich, Z.; Yushchenko, D. A.; Arntz, Y.; Didier, P.; Mely, Y.; Klymchenko, A. S. *J. Am. Chem. Soc.* **2010**, *132*, 4907.
(37) McMorrow, D.; Kasha, M. *J. Phys. Chem.* **1984**, *88*, 2235.
(38) Liu, W.; Wang, Y.; Jin, W.; Shen, G.; Yu, R. *Anal. Chim. Acta* **1999**, *383*, 299.
(39) Klymchenko, A. S.; Demchenko, A. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 461.
(40) Klymchenko, A. S.; Duportail, G.; Demchenko, A. P.; Mély, Y. *Biophys. J.* **2004**, *86*, 2929.
(41) Shynkar, V. V.; Klymchenko, A. S.; Piémont, E.; Demchenko, A. P.; Mély, Y. *J. Phys. Chem. A* **2004**, *108*, 8151.
(42) Klymchenko, A. S.; Kenfack, C.; Duportail, G.; Mely, Y. *J. Chem. Sci.* **2007**, *119*, 83.
(43) Klymchenko, A. S.; Shvadchak, V. V.; Yushchenko, D. A.; Jain, N.; Mely, Y. *J. Phys. Chem. B* **2008**, *112*, 12050.
(44) Das, R.; Klymchenko, A. S.; Duportail, G.; Mély, Y. *Photochem. Photobiol. Sci.* **2009**, *8*, 1583.
(45) Klymchenko, A. S.; Mely, Y.; Demchenko, A. P.; Duportail, G. *Biochim. Biophys. Acta* **2004**, *1665*, 6.
(46) Oncul, S.; Klymchenko, A. S.; Kucherak, O. A.; Demchenko, A. P.; Martin, S.; Dontenwill, M.; Arntz, Y.; Didier, P.; Duportail, G.; Mély, Y. *Biochim. Biophys. Acta* **2010**, *1798*, 1436.
(47) M'Baye, G.; Mély, Y.; Duportail, G.; Klymchenko, A. S. *Biophys. J.* **2008**, *95*, 1217.
(48) van Hell, A. J.; Klymchenko, A.; Burgers, P. P.; Moret, E. E.; Jiskoot, W.; Hennink, W. E.; Crommelin, D. J.; Mastrobattista, E. *J. Phys. Chem. B* **2010**, *114*, 11046.
(49) Demchenko, A. P.; Mély, Y.; Duportail, G.; Klymchenko, A. S. *Biophys. J.* **2009**, *96*, 3461.
(50) Postupalenko, V. Y.; Shvadchak, V. V.; Duportail, G.; Pivovarenko, V. G.; Klymchenko, A. S.; Mély, Y. *Biochim. Biophys. Acta* **2011**, *1808*, 424.
(51) Yushchenko, D. A.; Fauerbach, J. A.; Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M. *J. Am. Chem. Soc.* **2010**, *132*, 7860.
(52) Enander, K.; Choulier, L.; Olsson, A. L.; Yushchenko, D. A.; Kanmert, D.; Klymchenko, A. S.; Demchenko, A. P.; Mely, Y.; Altschuh, D. *Bioconjugate Chem.* **2008**, *19*, 1864.
(53) Boudier, C.; Klymchenko, A. S.; Mely, Y.; Follenius-Wund, A. *Photochem. Photobiol. Sci.* **2009**, *8*, 814.
(54) Zamotaiev, O. M.; Postupalenko, V. Y.; Shvadchak, V. V.; Pivovarenko, V. G.; Klymchenko, A. S.; Mély, Y. *Bioconjugate Chem.* **2011**, *22*, 101.
(55) Shvadchak, V. V.; Klymchenko, A. S.; de Rocquigny, H.; Mely, Y. *Nucleic Acids Res.* **2009**, *37*, 885.
(56) Klymchenko, A. S.; Demchenko, A. P. *New J. Chem.* **2004**, *28*, 687.
(57) Strandjord, A. J. G.; Barbara, P. F. *J. Chem. Phys.* **1985**, *89*, 2355.
(58) Pivovarenko, V. G.; Wróblewska, A.; Błazejowski, J. *J. Mol. Struct.* **2004**, *708*, 175.
(59) Eastman, J. W. *Photochem. Photobiol.* **1967**, *6*, 55.
(60) Choulier, L.; Shvadchak, V. V.; Naidoo, A.; Klymchenko, A. S.; Mély, Y.; Altschuh, D. *Anal. Biochem.* **2010**, *401*, 188.
(61) Doroshenko, A. O.; Sychevskaya, L. B.; Grygorovych, A. V.; Pivovarenko, V. G. *J. Fluoresc.* **2002**, *12*, 455.
(62) Shynkar, V. V.; Mely, Y.; Duportail, G.; Piemont, E.; Klymchenko, A. S. *J. Phys. Chem. A* **2003**, *107*, 9522.
(63) Blandamer, M. J.; Engberts, J. B. F. N.; Gleeson, P. T.; Reis, J. C. R. *Chem. Soc. Rev.* **2005**, *34*, 440.
(64) Bakhshiev, N. G., Ed.; *Solvatochromism. Problems and Methods*; LGU: Leningrad, Russia, 1989.
(65) Salari, H.; Khodadadi-Moghaddam, M.; Harifi-Mood, A. R.; Gholami, M. R. *J. Phys. Chem. B* **2010**, *114*, 9586.
(66) Boudier, C.; Storchak, R.; Sharma, K. K.; Didier, P.; Follenius-Wund, A.; Muller, S.; Darlix, J. L.; Mély, Y. *J. Mol. Biol.* **2010**, *400*, 487.
(67) Darlix, J.-L.; Godet, J.; Ivanyi-Nagy, R.; Fossé, P.; Mauffret, O.; Mély, Y. *J. Mol. Biol.* **2011**, *410*, 565.

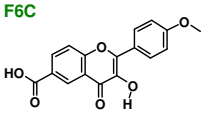
Supporting Information

Quantification of Local Hydration at the Surface of Biomolecules Using Dual-Fluorescence Labels

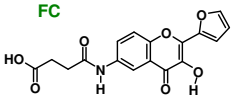
Vasyl G. Pivovarenko*[a,b] Oleksandr M. Zamotaiev,[a] Volodymyr V. Shvadchak,[b] Viktoriia Y. Postupalenko,[b] Andrey S. Klymchenko,[b] Yves Mély[b]

Labels:

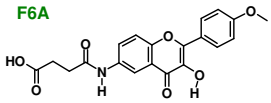
F6C



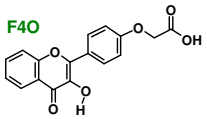
FC



F6A



F4O



Peptides:

Tat(44-61): **GISYGRKKRRQRRRPPQG**

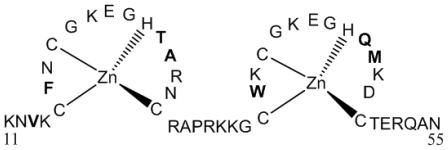
$(Gly)_5$: **GGGGG**

Peptide (4-17): **KRTAMFQDPQERPRC**

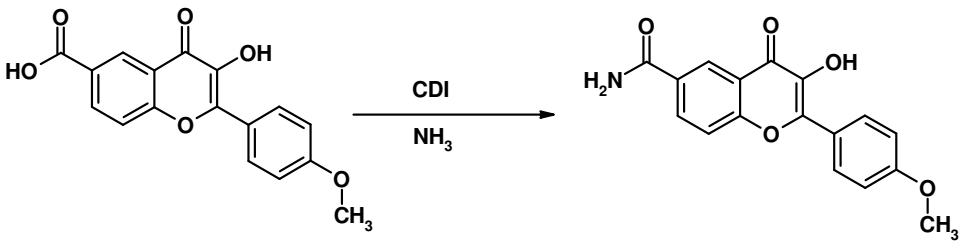
Peptide (7-17): **AMFQDPQERPRC**

Peptide (9-17): **FQDPQERPRC**

NC(11-55):



1. Synthesis of 3-hydroxy-2-(4-methoxyphenyl)-4-oxo-4H-6-chromenecarboxamide (F6Ca).



A mixture of 70 mg (0.182 mmol) of 3-hydroxy-2-(4-methoxyphenyl)-4-oxo-4H-chromene-6-carboxylic acid [S1] and 55 mg (1.5 eq.) of carbonyl-diimidazole in 5 ml of dry DMF was stirred during 10 minutes at rt under Ar-atmosphere. Then the mixture was heated to 50 ºC and after 15 minutes it was cooled to rt. Afterwards, dry NH_3 was bubbled through the solution during 10 minutes. Obtained yellow solution was left for 30 h. Then, it was diluted with 10 ml of water and acidified to pH 5 with 5 % HCl. White precipitate was filtered, washed with water 3 times and dried.

Yield 52 mg (74 %). The obtained F6C probe was of 98% purity according to HPLC. ^{1}H-NMR (400MHz, DMSO-d_6): δ, p.p.m., 3.87 (s, 3H, O**CH$_3$**), 7.05 (d, 2H, J= 8.8 Hz, Ar**H**), 7.31 (br. s., 1H, N**H**), 7.70 (d, 1H, J= 8.8 Hz, Ar**H**), 8.20 (m, 4H, Ar**H**), 8.67 (br. s., 1H, N**H**), 9.35 (br. s., 1H, **OH**), *m/z* (M+H$^+$) calculated for $C_{17}H_{14}NO_5^+$: 312.087; found: 312.087.

2. Spectroscopic properties of the labels.

Table S1. Spectroscopic properties of FC, F6A, F6C and F4O labels.

Media	C_M	Label	λ_{Abs}, nm	λ_{N^*}, nm	λ_{T^*}, nm	N*/T*	QY, %
Water	55.56	FC	357	431	514	3.18	2.1
		F6A	351	443	514	15	10.6
		F6C	355	442	517	18	8.1
		F4O	353	437	516	26	6.6
MeOH	24.75	FC	349	423	533	0.78	5.8
		F6A	349	429	533	1.07	6.1
		F6C	354	428	535	0.83	4.9
		F4O	356	432	529	4.55	4.5
EtOH	17.13	FC	351	420	535	0,43	6.3
		F6A	351	424	535	0,46	6.5
		F6C	356	419	533	0.23	5.1
		F4O	359	430	535	1.99	3.6
DMF	12.92	FC	344	421	539	0.29	7.1
		F6A	345	411	539	0,35	4.8
		F6C	354	418	544	0.13	5.1
		F4O	356	432	540	0.70	2.7
MF	17.12	FC	353	427	537	0.96	7.6
		F6A	351	429	535	0.88	8.1
		F6C	356	426	536	0.52	6.8
		F4O	357	430	536	1.81	4.9
Acetonitrile	18.90	FC	344	413	533	0.075	2.8
		F6A	344	419	531	0.07	5.0
		F6C	352	422	538	0.084	4.5
		F4O	347	409	529	0.05	4.0
THF	12.22	FC	339	408	537	0.11	5.5
		F6A	338	409	538	0.05	6.3
		F6C	355	411	544	0.04	9.5
		F4O	350	406	537	0.10	6.7
Dioxane	11.67	FC	340	411	537	0.21	6.2
		F6A	339	412	537	0,07	6.4
		F6C	354	408	541	0.084	10.7
		F4O	350	405	534	0.08	8.9

C_M – molar concentration of neat solvent, λ_{ABS} – position of absorption maxima, λ_{N^*} and λ_{T^*} - position of fluorescence maxima of N* and T* forms, respectively. N*/T* - ratio of integral intensities of the two emission bands. QY- fluorescence quantum yield.

Table S2. Spectroscopic properties of F6Ca probe.

Solvent	Molar conc.	ε	Log P	α	β	λ_{abs}	λ_{N^*}	λ_{T^*}	N*/T*	QY
Dioxane	11.67	2.21	-0.31	0	0.65	354	427	543	0.084	10.7
THF	12.22	7.43	0.46	0	0.48	355	411	544	0.040	9.5
Acetonitrile	18.90	35.7	-0.34	0.19	0.32	352	422	538	0.050	4.5
DMF	12.92	37.2	-0.87	0	0.74	354	418	544	0.130	5.1
DMSO	14.01	46.8	-1.35	0	0.76	355	422	546	0.211	7.7
t-PentOH	9.132	-	0.89	-	0.63	356	418	531	0.166	6.4
2-PrOH	13.13	19,3	0.05	0.76	0.56	356	422	536	0.213	4.8
EtOH	17.13	24.9	-0.31	0.83	0.50	356	419	533	0.231	5.0
MeOH	24.75	32.6	-0.77	0.93	0.44	354	428	535	0.826	4.9
Ethyleneglycol	17.93	40.2	-1.36	0.9	0.52	358	438	537	1.46	6.2
Trifluoroethanol	13.72	26.7	0.41	1.51	0.25	363	442	505	4.8	9.8
MF	17.12	182	-0.97	-	0.54	356	426	536	0.519	6.8
FA	25.18	109	-0.82	0.71	0.60	359	432	529	1.15	9.3
Water	55.56	78.4	-1.38	1.17	0.18	355	442	514	18	7.8

ε – dielectric permeability index, Log P – logarithm of partition coefficient in water – 1-octanol system, α,β - HB-donating and HB-accepting coefficients, taken from [S2]. λ_{Abs} – position of absorption maximum, λ_{N^*} and λ_{T^*} - position of fluorescence maxima of N* and T* forms, respectively. N*/T* – integral intensity ratio of the two emission bands (calculated after mathematical deconvolution of the spectra), QY- fluorescence quantum yield.

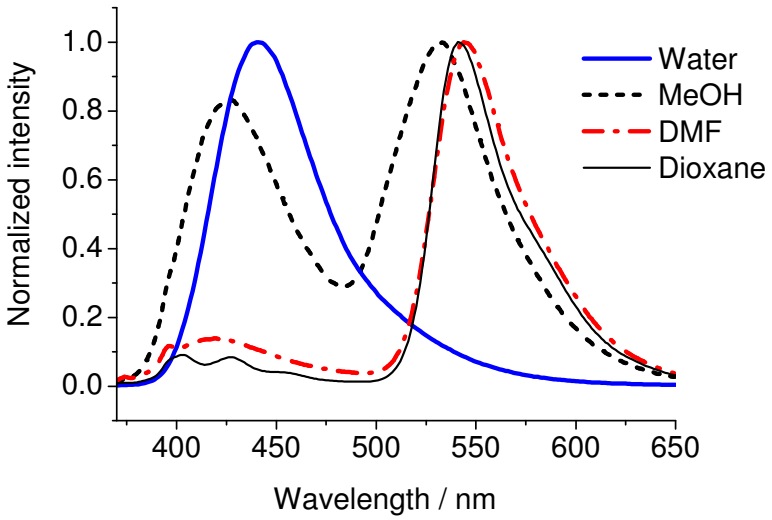


Figure S1. Normalized fluorescence spectra of F6Ca probe in different solvents. The N*/T* ratio of F6Ca in low-polar dioxane (ε=2.21) is similar to that in polar DMF (ε=37.2,

N*/T*=0.13), while in polar HB donor methanol (ε=32.6), the N*/T* ratio is equal to 0.83 (Table S2). This shows that H-bonding with the environment, but not its polarity, predominantly influences the N*/T* fluorescence parameters.

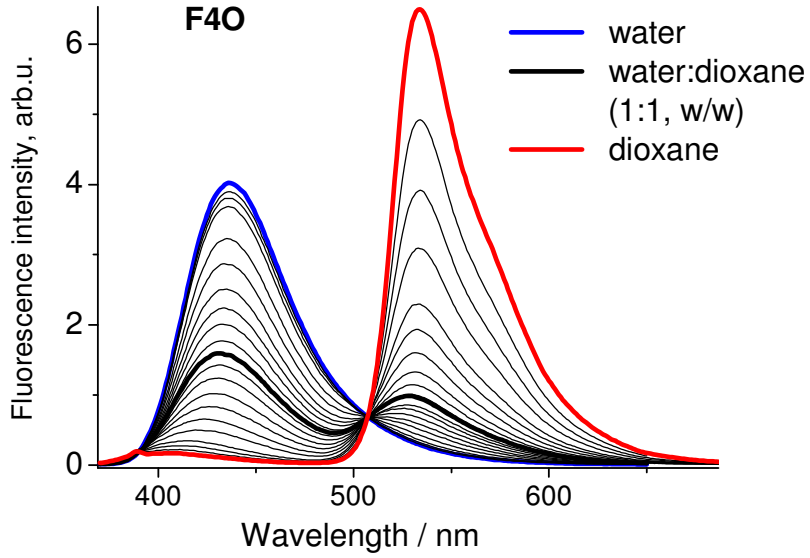


Figure S2. Fluorescence spectra of the F4O probe in water-dioxane binary mixture of different compositions. Probe concentration is 3 μM in all cases.

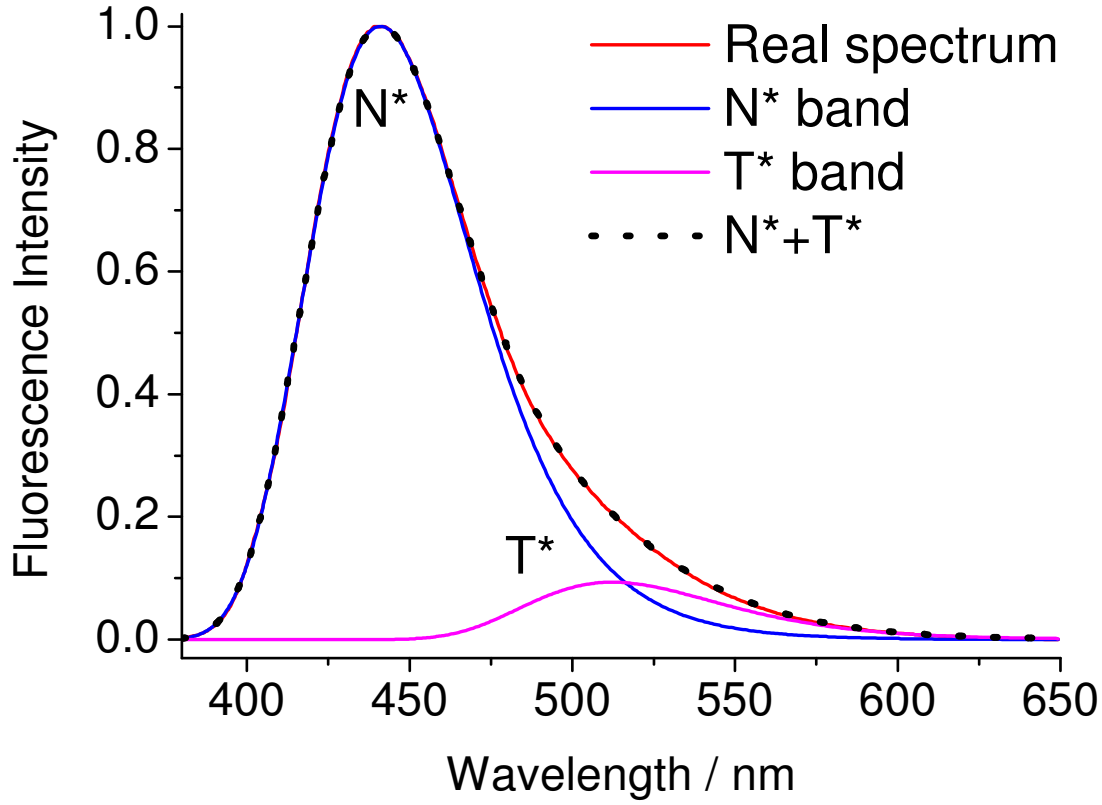


Figure S3. Example of deconvolution of F6Ca probe spectrum in N* and T* bands**.**

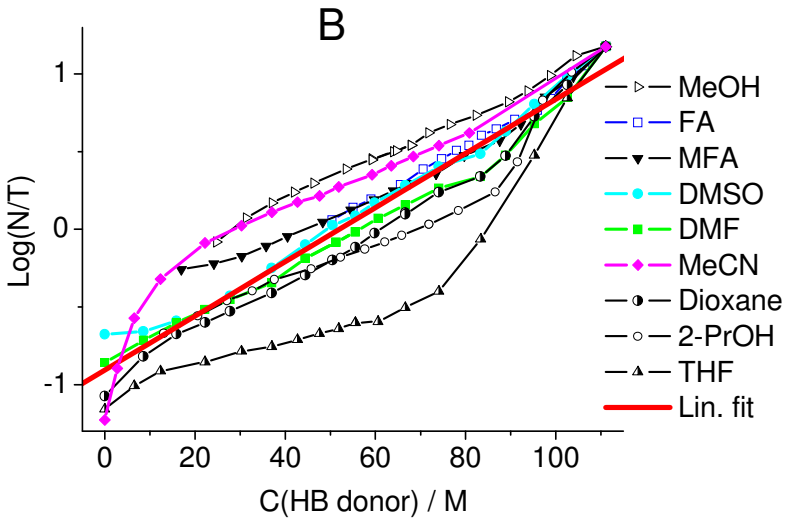


Figure S4. Dependence of Log(N*/T*) on the concentration of hydrogen bond donor in water-solvent mixtures for the F6Ca probe. Since water and formamide possess two HB donor groups, we use $C_{HB\ donor} = 2C_{solvent}$ for these solvents.

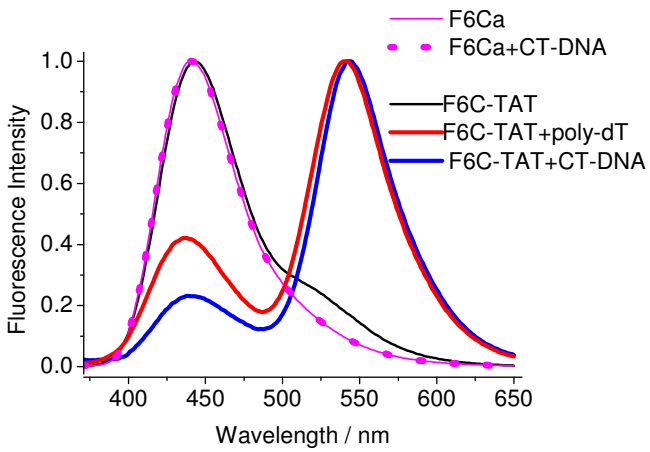


Figure S5. Effect of nucleic acids on the normalized fluorescence spectra of F6Ca probe and F6C conjugated to Tat(44-61). The emission spectrum of F6Ca was recorded in the absence or in the presence of CT-DNA. Tat(44-61) peptide labeled with F6C was recorded in the absence or in the presence of with poly(dT) or CT-DNA. All spectra were obtained in 10 mM phosphate buffer, pH 7.0. The complexes with nucleic acids were obtained with a ratio of one F6C–Tat conjugate per 10 nucleic bases (ss-DNA) or base pairs (ds-DNA), ensuring a complete peptide binding [S3]. Excitation wavelength is 355 nm in all cases.

[S1] Pfister, J.R.; Wymann, W. E.; Schuler, M.E.; Roszkowski, A.P. Inhibition of Histamine-Induced Gastric Secretion by Flavone-6-carboxylic Acids. J. Med.Chem. 1980,23, 335-338.

[S2] 53. M.H. Abraham (1993) Hydrogen bonding. XXXI: Construction of a scale of solute effective or summation hydrogen-bond basicity. J. Phys. Org. Chem., 6: 660-684.

[S3]Kuciak, M., Gabus, C., Ivanyi-Nagy, R., Semrad, K., Storchak, R., Chaloin, O., Muller, S., Mely, Y., and Darlix, J. L. (2008) The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities in vitro. *Nucleic Acids Res. 36*, 3389-3400.

2.1.3. Fluorescent ratiometric nucleoside analogues for monitoring peptide – oligonucleotide interactions

Another way for studying peptide-oligonucleotide interactions can be achieved by oligonucleotide labeling. Due to the fact that the intrinsic fluorescence of the natural nucleotides is far too low, oligonucleotides (ODNs) should be labeled with extrinsic fluorophores. These labels should exhibit high absorption coefficient and quantum yield, absorb in a wavelength range where nucleic acids and proteins are transparent, show large Stokes shift and be sensitive to changes of the local environment.

In this context, due to the severe limitations of the commercially available fluorescent nucleosides, a novel class of fluorescent nucleosides where the natural base is substituted by a 2-aryl-3-hydroxychromones (3HC) was developed (**Publication 4**). These compounds appear as attractive analogues of nucleic bases because i) their size corresponds well to the size of the two complementary (A-T or G-C) bases and ii) as it was shown in our previous studies, 3HC dye when conjugated to spermine can intercalate within the base pairs of dsDNA, giving a strong variation of its dual emission and an increase in its fluorescence intensity (Klymchenko, Shvadchak et al. 2008).

New fluorescent nucleoside analogues were synthesized in collaboration with the team of A. Burger (Nice). Nucleoside 1a (Figure 2.7) bearing the 2-thienyl-3-hydroxychromone nucleobase was prepared using sequential aryl-aldol condensation/cycloetherification or a Friedel-Crafts glycosylation as key steps. The synthesis of the triazolyl derivative 1b (Figure 2.7) was achieved using a convergent 1,3-dipolar cycloaddition strategy.

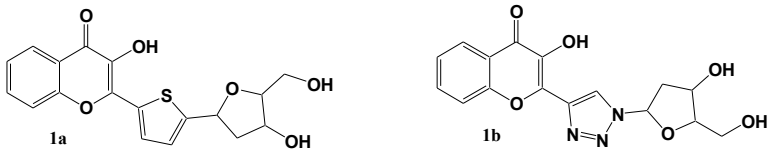


Figure 2.7. Structures of 2-aryl-3HC nucleosides 1a and 1b.

The photophysical properties of compounds 1a and 1b were investigated in solvents of different polarity. In all studied solvents, these compounds present a dual emission with well-resolved short and long-wavelength bands which were assigned to the emission of the N* and T* forms, respectively (Figure 2.8). The N*/T* intensity ratio of 1a nucleoside increases gradually with increasing Et(30) values from 0.13 in acetonitrile up to 1.72 in water (Figure 2.8A). This correlation of the N*/T* ratio values with Et(30) suggests that the ESIPT reaction and thus, the formation of the T* form in 1a is hampered in polar protic solvents, as a result of the formation of intermolecular H-bonds that weaken or even disrupt the intramolecular H-bond, needed for the ESIPT reaction.

For derivative 1b bearing the triazole electron-deficient system, the N*/T* ratio in all solvents is much lower than for 1a with a predominant emission of the T* form due to a faster ESIPT reaction, as previously reported for 2-aryl 3HC dyes with low electron donor ability (Figure 2.8B). Moreover, compound 1b exhibits a relatively low fluorescence quantum yield in most solvents. From the comparison of these two nucleosides, it is clear that 1a is much more advantageous than 1b in terms of spectroscopic properties and sensitivity to solvent polarity. The

strong variation of the N*/T* ratio of 1a, especially in polar solvents makes it attractive for applications in DNA research for monitoring microenvironment changes and DNA dynamics at the probe site due to the polarity differences between the interior and the surface of the DNA helix.

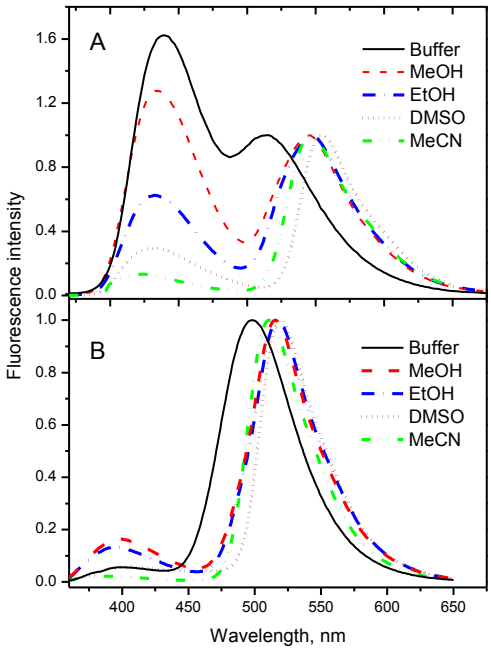


Figure 2.8. Normalized fluorescence spectra of 1a (A) and 1b (B) nucleosides in different solvents.

Publication 4

Efficient synthesis of ratiometric fluorescent nucleosides featuring 3-hydroxychromone nucleobases

Tetrahedron 65 (2009) 7809–7816

Contents lists available at ScienceDirect

Tetrahedron

journal homepage: www.elsevier.com/locate/tet

Efficient Synthesis of Ratiometric Fluorescent Nucleosides Featuring 3-Hydroxychromone Nucleobases

Marie Spadafora [a], Victoria Y. Postupalenko [b], Volodymyr V. Shvadchak [b], Andrey S. Klymchenko [b], Yves Mély [b], Alain Burger [a,*], Rachid Benhida [a,*]

[a] *Laboratoire de Chimie des Molécules Bioactives et des Arômes, UMR 6001 CNRS, Institut de Chimie de Nice, Université de Nice-Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France*

[b] *Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie 74, Route du Rhin, BP 24, 67401 Illkirch cedex, France*

ARTICLE INFO

Article history:
Received 16 June 2009
Received in revised form 2 July 2009
Accepted 4 July 2009
Available online 9 July 2009

ABSTRACT

The synthesis of a novel class of fluorescent nucleosides featuring 2-aryl-3-hydroxychromones (3-HC) as base analogues is described. Nucleoside **1a** bearing the 2-thienyl-3-HC nucleobase was prepared using sequential aryl–aldol condensation/cycloetherification or a Friedel–Crafts glycosylation as key steps. The synthesis of the triazolyl derivative **1b** was achieved using a convergent 1,3-dipolar cycloaddition strategy. Fluorescence studies show that 3-HC-thienyl-nucleoside **1a** displays high sensitivity of its dual emission to polarity changes and therefore is highly promising for nucleic acid labelling.

1. Introduction

Due to its exquisite sensitivity, fluorescence is one of the most sensitive techniques for investigating biological systems. In the context of nucleic acids, the common strategy consists to substitute a natural nucleoside by a fluorescent analogue because the canonical bases are poorly or non-fluorescent. Most of the fluorescent analogues used for nucleic acid labelling are single band emitter probes.[1] Among them, 2′-deoxyribosyl-2-aminopurine (2-AP), a mimic of 2′-deoxyadenosine, is the most popular fluorescent base.[2] However, 2-AP displays severe limitations related to the dramatic drop of its quantum yield when incorporated in oligonucleotides.[3] Moreover, when protein binding to nucleic acids does not induce a transition between stacked and unstacked conformations, the interactions are difficult to monitor due to the limited environment sensitivity of 2-AP. Recently, 8-vinyl-deoxyadenosine (8-vdA), a new fluorescent analogue of 2′-deoxyadenosine was reported by us.[4] Although 8-vdA displays improved sensitivity compared to 2-AP, its domain of application is similar. Other fluorescent nucleoside analogues have been developed but, when incorporated into ODNs, they are quenched, destabilizing or of limited sensitivity to environmental change. As a consequence, there is a strong demand for new fluorescent nucleoside analogues with improved spectroscopic properties and this is the purpose of intense research.[1]

On the other hand, 3-hydroxychromones (3-HC) have been shown to be powerful fluorescence probes for a large range of applications in model membranes, biomembranes and proteins.[5] Due to an excited state intramolecular proton transfer (ESIPT, Fig. 1),[6] these fluorophores exhibit two excited state forms: the initially excited normal (N*) and the tautomeric (T*) forms.[7] Since the ESIPT reaction in these dyes is strongly sensitive to polarity, [8] H-bonds[8b,9] and electric fields,[10] the ratio of the two emission bands could be used to sensitively monitor environmental changes. Moreover, the positions of the absorption and emission bands as well as the fluorescence intensity ratio in 3-HC dyes could be used to further characterize the properties of the probe environment.[8b,11] 3-HC fluorophores bearing a small heterocycle in 2-position are attractive analogues of the nucleic bases

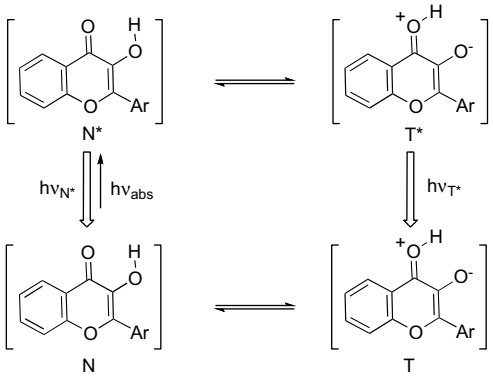


Figure 1. ESIPT reaction in 3-hydroxychromones.

* Corresponding authors. Tel.: +33 4 92 07 61 74; fax: +33 4 92 07 61 51.
E-mail addresses: benhida@unice.fr (R. Benhida), burger@unice.fr (A. Burger).

0040-4020/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2009.07.021

because their size corresponds well to the size of the two complementary (A–T or G–C) bases. Moreover, according to our recent studies, the 2-(2-furanyl)-3-hydroxychromone dye when conjugated to spermine is able to intercalate within the base pairs of double-stranded DNA, giving a strong variation of its dual emission and an increase in its fluorescence intensity.[12]

To explore the potentiality of 3-HC as fluorescent probes for DNA labelling, we report here the first synthesis and preliminary spectroscopic characterization of nucleoside analogues where a 3-HC substitutes a natural base. The structures of the targeted 3-HC-nucleosides **1a** and **1b** are shown in Figure 2. They were designed (i) to keep a 2′-deoxy structure and a β-anomeric configuration for incorporation into DNA oligonucleotides and (ii) to have an electron donating or an electron withdrawing group connected between the sugar and chromone moieties in order to modulate the photophysical properties of the nucleoside analogues.

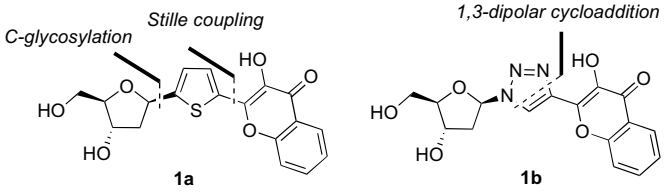


Figure 2. Targeted 2-aryl-3-HC-nucleosides.

2. Results and discussions

For the synthesis of **1a** and **1b** we used two convergent strategies based on C-glycosylation–Stille type coupling (for **1a**), and azide-alkyne 1,3-dipolar cycloaddition (for **1b**) as key steps (Fig. 2).

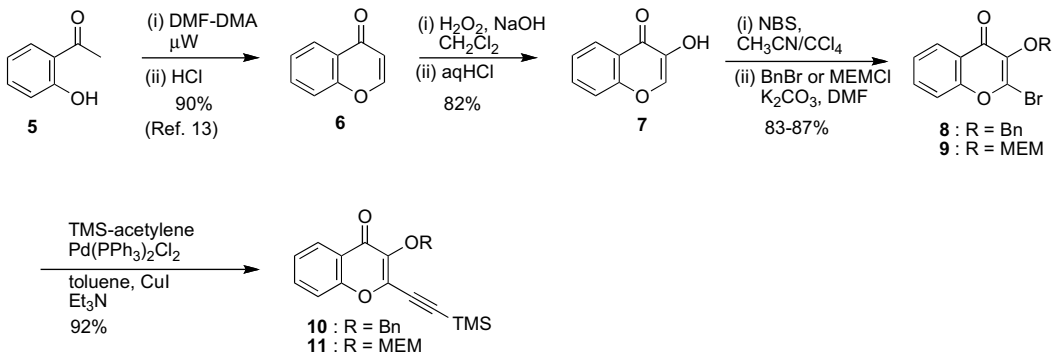
The preparation of 3-HC key intermediates **8**–**11** required for the synthesis of **1a** and **1b** is described in Scheme 1. Starting from 2-hydroxy-acetophenone **5**, condensation with *N*,*N*-dimethylformamide dimethylacetal followed by HCl-mediated in situ cyclization provided the chromone **6**.[13] Epoxidation of **6** using H_2O_2/NaOH in CH_2Cl_2 followed by acid-mediated ring-opening afforded **7** in 82% overall yield. The use of protic solvents (MeOH, EtOH) or other oxidizing agents such as *m*-CPBA and 2-butanone peroxide only gave low yields.[14] Bromination of **7** followed by successive protection with benzyl or MEM group gave **8** and **9**, respectively (83–87%, two steps). Sonogashira coupling carried out on **8** and **9** using TMS-acetylene provided the 3-HC **10** and **11**, respectively (90–93%). We found that the cleavage of the TMS group of compounds **10** and **11** resulted in unstable terminal alkynes. Therefore, the TMS-protected alkynes **10** and **11** were deprotected in situ and used without purification.

With these key intermediates in hand, the thienyl-derived 3-HC-C-nucleoside **1a** bearing the β-C–C linkage was prepared starting from the aldehyde **2** (Path A)[15] or the acetyl-deoxyribose (**3a** or **3b**, Path B) as illustrated in Scheme 2.

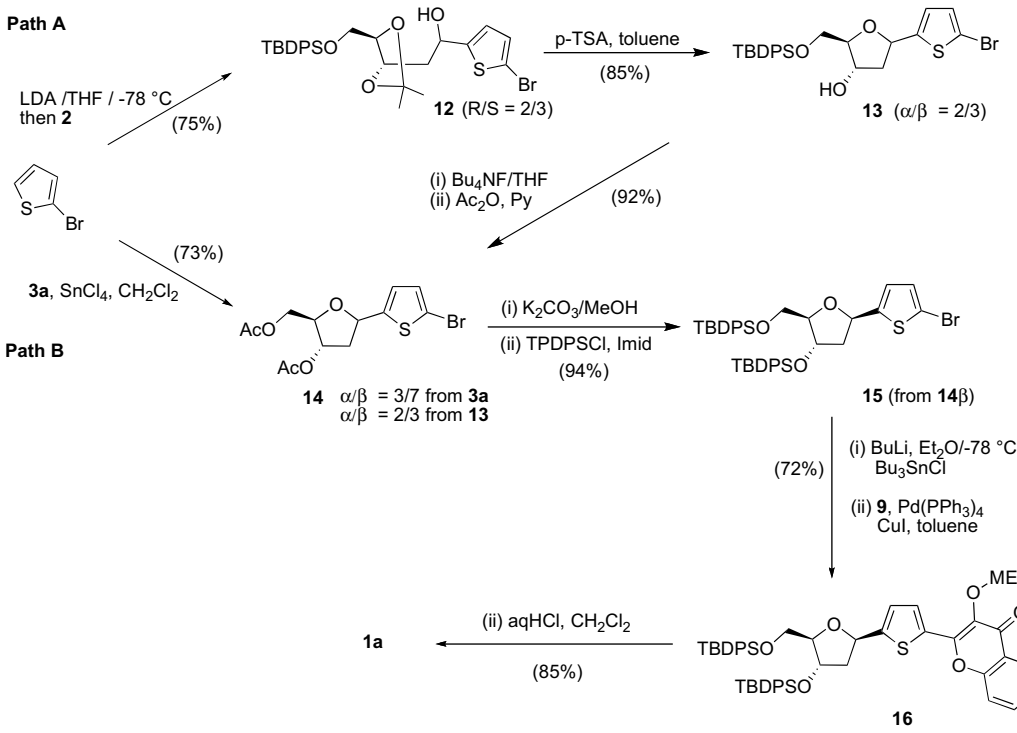
Thus, regioselective *ortho*-lithiation of 2-bromothiophene using LDA in THF at −78 °C followed by addition of the protected aldehyde **2** led to the alcohol **12** in 75% yield as a mixture of inseparable *R*/*S* isomers (40:60). C-Nucleoside **13** (α/β mixture) was then obtained from **12** (*R*/*S*) following isopropylidene cleavage and subsequent C4′–C1′ cycloetherification (85%). However, the separation of α/β anomers required conversion of **13** to its acetyl-protected analogue **14**. Interestingly and according to path B, the direct glycosylation leading to **14** was cleanly achieved in one step using the catalytic Friedel–Crafts reaction.[16] Indeed, treatment of 2-bromothiophene and acetyl-ribose **3a** (or **3b**) with $SnCl_4$ in CH_2Cl_2 provided good yield and diastereoselectivity in favour of the desired β-anomer **14**β (73% yield, β/α=70:30). Furthermore, the anomeric configuration was clearly evidenced by the observed NOE correlations between H1′ and H4′ for **14**β, and between H1′–H3′ and H1′–H5' for **14**α (Scheme 2). The synthesis of **1a** was then pursued from **14**β following subsequent acetyl cleavage/TBDPS protection to provide **15** (94%).[16f] Halogen–metal exchange/stannylation followed by Pd-catalyzed Stille type coupling between **15** and 2-bromo-3-HC **9** led to the protected C-nucleoside **16** (72%), which was quantitatively converted, after MEM and TBDPS cleavage, to the 2-thienyl-3-HC C-nucleoside **1a**.

The synthesis of compound **1b** was achieved straightforwardly. As shown in Scheme 3, after in situ cleavage of the TMS group, the 1,3-dipolar cycloaddition between **10** and azido-sugar **3**[17a] was carried out both in solution using AcOH as a co-catalyst and on SiO_2 under solvent-free and microwave activation, according to our recent published work.[17] Both procedures provided the cycloadduct **17** in high yields (89–95%). Finally, methanolysis of **17** followed by catalytic hydrogenolysis afforded the triazolyl-3-HC nucleoside **1b** in 86% yield (two steps). Following the same strategy, but using the 3-HC **11** rather than **10**, the deprotection of the MEM group (HCl or TFA) failed since a concomitant cleavage of the glycosidic bond was observed.

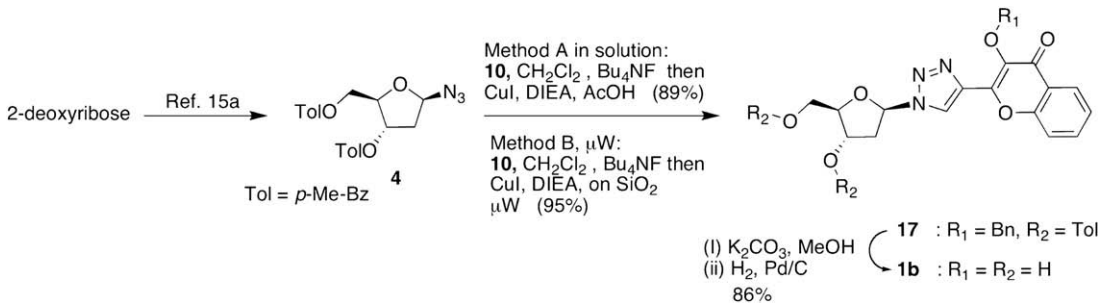
The photophysical properties of compounds **1a** and **1b** were investigated in solvents of different polarity. To scale the polarity, we used the $E_T(30)$ index. This empirical parameter accounts for the dielectric constant of the solvent and its H-bond donor ability,[18] both of which influence strongly the fluorescence properties of the 3-HC dyes.[8b,11] The results are given in Table 1 and Figures 3 and 4. In the studied solvents, compound **1a** shows an absorption band centred around 361–369 nm, whereas **1b** exhibits a significantly blue shifted absorption having two distinct maxima at 299–304 and 332–338 nm (Fig. 3). The position of the absorption maximum slightly increases with $E_T(30)$ (Table 1). In all studied solvents, compound **1a**, featuring the electron-rich thienyl heterocycle, presents a well-resolved dual emission (Fig. 4A), where the short- and the long-wavelength bands could be assigned to the emission



Scheme 1. Synthesis of protected 3-HC **10** and **11**.



Scheme 2. Synthesis of 3-HC C-nucleoside **1a**.



Scheme 3. Synthesis of 3-HC nucleoside **1b**.

Table 1
Spectroscopic data of nucleosides **1a** and **1b** in different solvents[a]

Nucl.	Solv.	$E_T(30)$	λ_{abs}	ε	λ_{N^*}	λ_{T^*}	N^*/T^*	QY
1a	Buffer	63.1	367	13,000	440	515	1.72	0.046
	MeOH	55.4	368	12,500	434	545	1.34	0.075
	EtOH	51.9	369	12,300	433	548	0.68	0.077
	BuOH	49.7	369	—	431	546	0.52	0.091
	DMSO	45.1	368	—	431	554	0.30	0.197
	MeCN	45.6	362	—	421	546	0.13	0.179
	Acetone	42.2	362	—	421	552	0.22	0.053
	CH_2Cl_2	40.7	363	—	418	543	0.04	0.141
	EtOAc	38.1	361	—	415	551	0.09	0.100
1b	Buffer	63.1	338	11,500	401	500	0.06	0.022
	MeOH	55.4	336	12,000	398	516	0.17	0.024
	EtOH	51.9	335	12,000	397	517	0.13	0.027
	BuOH	49.7	333	—	393	518	0.16	0.033
	DMSO	45.1	341	—	390	520	0.06	0.052
	MeCN	45.6	332	—	387	511	0.02	0.189
	Acetone	42.2	334	—	387	515	0.07	0.037
	CH_2Cl_2	40.7	333	—	—	506	<0.01	0.183
	EtOAc	38.1	332	—	387	512	0.03	0.074

[a] λ_{abs} is the position of the absorption maxima, λ_{N^*} and λ_{T^*} are the positions of the fluorescence maxima of the N^* and T^* bands, respectively; N^*/T^* is the intensity ratio of the two emission bands at their maxima; QY is the fluorescence quantum yield calculated using quinine sulfate (QY=0.577 in 0.5 M H_2SO_4) as a reference; ε ($M^{-1}cm^{-1}$) is the extinction coefficient at the maximum of absorption; 10 mM phosphate buffer (pH 6.5) was used.

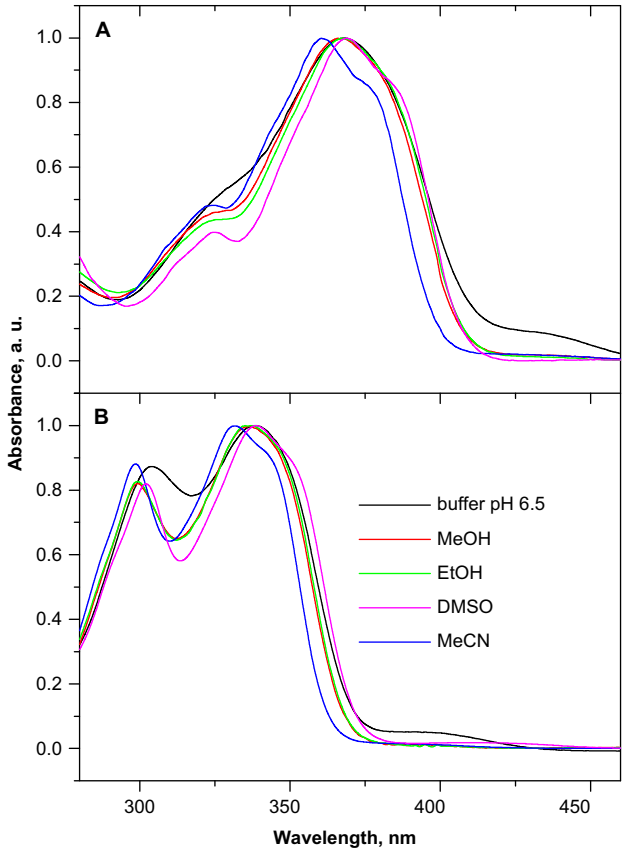


Figure 3. Normalized absorption spectra of nucleosides **1a** (A) and **1b** (B) in different solvents.

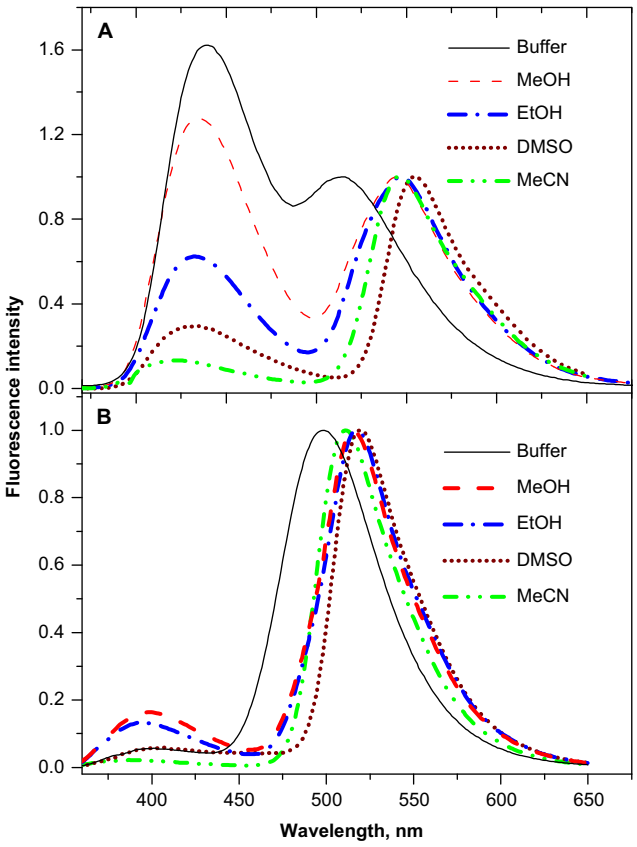


Figure 4. Fluorescence spectra of **1a** (A) and **1b** (B) in different solvents. The spectra were normalized at the long-wavelength maximum. Excitation wavelength was 360 and 340 nm for **1a** and **1b**, respectively.

of the N* and T* species, respectively. The N*/T* intensity ratio of this nucleoside increases gradually with increasing $E_T(30)$ values from 0.04 to 0.09 in dichloromethane and ethyl acetate up to 1.72 in water (Table 1 and Fig. 4A). Figure 5 shows that the N*/T* ratio correlates well with the $E_T(30)$ values. This correlation suggests that the ESIPT reaction and thus, the formation of the T* form in **1a** is hampered especially in polar protic solvents, as for its analogue 2-(2-furyl)-3-hydroxychromone.[11a] The hampering of the ESIPT reaction by protic solvents in 3-HC is connected with the formation of intermolecular H-bonds that weaken or even disrupt the intramolecular H-bond, needed for the ESIPT reaction.[8b,9] Moreover, the increase in $E_T(30)$ leads to a large blue shift of the T* band, especially in water (Table 1), due to its high H-bond donor ability.[19] In contrast, the N* band maximum shifts to the red on $E_T(30)$ increase, similarly to the absorption band. The fluorescence quantum yield of **1a** varies considerably with solvent, showing the highest value in polar aprotic DMSO (0.197) and the lowest one in polar protic water (0.046).

For derivative **1b** bearing the triazole electron-deficient system, the emission maxima are blue shifted in respect to **1a** (Table 1, Fig. 4B). The N*/T* band ratio in all the solvents is much lower than for **1a**, suggesting a faster ESIPT reaction that results in a predominant emission of the T* form. This observation is in line with the fast ESIPT previously reported for 3-HC dyes bearing 2-aryl group with low electron donor ability.[20] Remarkably, the N*/T* ratio is very low in buffer (0.06), probably due to specific solvation of the triazole ring. Moreover, compound **1b** exhibits a relatively low fluorescence quantum yield in most solvents, except in dichloromethane and acetonitrile (Table 1).

From the comparison of these two nucleosides, it is clear that **1a** is much more advantageous than **1b** in terms of spectroscopic properties and sensitivity to solvent polarity. The strong variation of the N*/T* ratio of **1a**, especially in polar solvents makes it attractive for applications in DNA research, where the local environment is expected to be relatively polar.

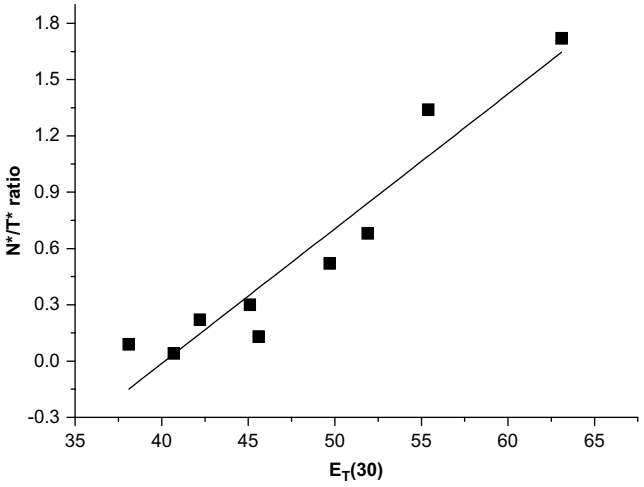


Figure 5. Correlation between the N*/T* intensity ratio of **1a** and the $E_T(30)$ polarity index of the different solvents used in this study (see Table 1). The line represents the linear fit to the experimental data.

3. Conclusion

In conclusion, we developed an efficient synthesis of unprecedented environment-sensitive ratiometric fluorescent

nucleosides bearing 2-aryl-3-HC nucleobases. The synthesis of **1a** involved as key steps an aryl–aldol condensation followed by regiocontrolled C4′–C1' cycloetherification or Friedel–Crafts type glycosylation. In the case of **1b** we used the azide-alkyne 1,3-dipolar cycloaddition, conveniently achieved under the cooperative effects of microwave activation and Cu(I)/SiO_2 catalysis. Fluorescence studies showed that **1a** and to a lesser extent **1b** retain the dual emission highly sensitive to the environment of the 3-HC fluorophore. Thus, a decrease of solvent polarity induces a significant change of the N*/T* ratio in **1a**, together with red and blue shifts of the T* and N* bands, respectively.

Due to the large polarity differences between the interior and the surface of the DNA helix, incorporation of **1a** in DNA should allow monitoring the microenvironmental changes and DNA dynamics at the probe site. Moreover, the N* and T* bands of **1a** at 440 and 515 nm, respectively, are well separated from each other and from the absorption/emission of the natural nucleobases and amino acids. Therefore, **1a** can be selectively excited and studied even in the presence of other nucleic acids and proteins. Incorporation of such ratiometric nucleosides into synthetic oligonucleotides is under way to explore their applications as nucleic acid labels.

4. Experimental section

4.1. General

All reactions were run under nitrogen atmosphere in dried glassware. Solvents were dried and distilled by standard procedures. Toluene was purchased from commercial sources and dried over 4 Å molecular sieves before use. Reagents were purchased and used without further purification. All reactions were monitored by thin layer chromatography (TLC) plates (0.2 mm, silica gel 60 with fluorescent indicator UV_{254}). Flash chromatography was performed using silica gel (60, 0.040–0.063 mm). ^{1}H NMR and ^{13}C NMR were recorded on 200 and 500 instruments (200 and 500 MHz for ^{1}H, 50 and 125 MHz for ^{13}C). Chemical shifts (δ) were reported in parts per million and the coupling constants were reported in hertz (Hz). Analytic High Performance Liquid Chromatography (HPLC) was recorded using a RP-C18 column (300A, 5 μm particle size). Absorption and fluorescence spectra were recorded on Cary 4 spectrophotometer (Varian) and FluoroMax 3.0 spectrofluorometer (Jobin Yvon, Horiba), respectively. Fluorescence quantum yields were determined by taking quinine sulfate in 0.5 M sulfuric acid (quantum yield, QY = 0.577)[21] as a reference.

4.2. Synthesis of the ethynyl-3-hydroxy-chromone 10 and 11

4.2.1. (E)-3-(Dimethylamino)-1-(2-hydroxyphenyl)-prop-2-en-1-one (enamide intermediate)

A mixture of 2′-hydroxyacetophenone (5 mmol, 0.6 mL) and *N*,*N*-dimethylformamide-dimethylacetal (1 equiv, 0.66 mL) was irradiated under microwave for 15 s (300 W max, T=115 °C). The resulting mixture was cooled at room temperature and crystallized in pentane to give the enamine intermediate (red crystals, 955 mg, 100%). Mp (ether)=130–131 °C (lit. 132–134, Ref. 1). R_f=0.2 (cyclohexane/ethyl acetate: 60:40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=2.95 (s, 3H, NMe), 3.17 (s, 3H, NMe), 5.76 (d, 1H, J=12.1 Hz), 6.81 (dt, 1H, J=1.0 and 7.1 Hz), 6.92 (dd, 1H, J=0.9 and 8.4 Hz), 7.30 (dt, 1H, J=1.5 and 8.4 Hz), 7.69 (dd, 1H, J=1.4 and 8.0 Hz), 7.87 (d, 1H, J=12.0 Hz), 13.99 (s, 1H, OH). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=37.5, 45.5, 90.1, 118.1, 118.3, 120.4, 128.3, 134.0, 154.9, 163.0, 191.6. MS (ESI, MeOH) m/z: 213.70 $[M+Na]^+$.

*4.2.2. Chromone **6***

To a solution of the enamide previously obtained (1 g, 5.24 mmol) in methylene chloride (40 mL) was added concentrated HCl (4 mL). The resulting mixture was refluxed for 1 h. After cooling, the mixture was extracted with methylene chloride (3×40 mL). The combined organic layers were washed with saturated $NaHCO_3$ solution, then with brine, dried over $MgSO_4$, filtered and concentrated to afford chromone **6** (red crystals, 700 mg, 90%). Mp (ether)=52–54 °C (lit. 52 °C, Ref. 1). R_f=0.45 (cyclohexane/ethyl acetate: 60:40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=6.31 (d, 1H, J=6.0 Hz), 7.33–7.44 (m, 2H), 7.63 (ddd, 1H, J=1.8, 7.1 and 8.4 Hz), 7.83 (d, 1H, J=6.0 Hz), 8.18 (dd, 1H, J=1.5 and 7.9 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=113.1, 118.3, 125.0, 125.3, 125.9, 133.8, 155.4, 156.6, 177.7. MS (ESI, MeOH) m/z: 168.9 $[M+Na]^+$.

4.2.3. 2,3-Epoxy-chromone (intermediate)

To a solution of chromone **6** (740 mg, 5 mmol) in methylene chloride (7.5 mL), was slowly added hydrogen peroxide (2 equiv, 35% solution, 1.4 mL) and NaOH (1.5 equiv, 407 mg) at 0 °C. The mixture was stirred 3 h at 0 °C, quenched with water and extracted with methylene chloride (3×20 mL). The combined organic layers were dried (Na_2SO_4) and evaporated under reduced pressure to give a crude amorphous solid (668 mg, 82%), which was used in the next step without purification. Mp (ether)=64–66 °C. R_f=0.8 (cyclohexane/ethyl acetate: 60:40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=3.69 (d, 1H, J=2.4 Hz), 5.66 (d, 1H, J=2.5 Hz), 7.05 (dd, 1H, J=0.7 and 8.5 Hz), 7.14 (td, 1H, J=1.0 and 8.0 Hz), 7.55 (ddd, 1H, J=1.8, 7.2 and 8.5 Hz), 7.85 (dd, 1H, J=1.8 and 8 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=55.4, 77.3, 118.1, 119.9, 123.4, 127.2, 136.4, 155.5, 188.2. MS (ESI, MeOH) m/z: 160.7 $[M-H]^-$.

*4.2.4. 3-Hydroxy-chromone **7***

To the epoxide previously obtained (570 mg, 3.5 mmol) was added concd HCl (20 mL) and the resulting mixture was heated at 70 °C for 1 h. After cooling, water was added (20 mL) and the mixture was extracted with methylene chloride (3×30 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ solution, dried over $MgSO_4$, filtered and concentrated to afford 3-hydroxychromone **7** (brown powder, 568 mg, 100%). Mp (ether)=179–181 °C. R_f=0.5 (cyclohexane/ethyl acetate: 50:50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=6.33 (s, 1H), 7.41 (ddd, 1H, J=1.1, 7.0 and 8.0 Hz), 7.50 (dd, 1H, J=0.6 and 8.6 Hz), 7.69 (ddd, 1H, J=1.8 and 7.0 and 8.6 Hz), 8.01 (s, 1H), 8.26 (dd, 1H, J=1.3 and 8 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=117.3, 123.4, 124.3, 132.3, 137.2, 140.5, 147.9, 155.1, 172.2. MS (ESI, MeOH) m/z: 184.9 $[M+Na]^+$. HRMS (ESI) calcd for $C_9H_7O_3$ $[M+H]^+$, 163.0395; found, 163.0389.

4.2.5. 2-Bromo-3-hydroxy-chromone (intermediate)

To a solution of 3-HC **7** (540 mg, 3.34 mmol, 0.65 equiv) in acetonitrile (10 mL) was added NBS (1 equiv, 940 mg) and 2,2′-azobis(2-methylpropionitrile) (vazo® 67, 0.1 equiv, 100 mg), and the mixture was refluxed for 8 h (0.1 equiv of vazo® 67 was added each 2 h). The mixture was cooled and the solvent evaporated in vacuo. The crude product was purified on silica gel (cyclohexane/ethyl acetate: 90:10) to give the corresponding 2-Br-3-HC derivative (colourless powder, 730 mg, 90%). Mp (ether)=178–180 °C. R_f=0.57 (cyclohexane/ethyl acetate: 50:50). ^{1}H NMR (MeOD, 200 MHz) δ=7.37 (ddd, 1H, J=1.0, 7.1 and 8.1 Hz), 7.47 (dd, 1H, J=0.6 and 8.6 Hz), 7.66 (ddd, 1H, J=1.6, 7.0 and 8.6 Hz), 8.06 (dd, 1H, J=1.7 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=118.1, 121.0, 125.4, 126.0, 129.8, 134.0, 140.3, 156.7, 171.3. MS (ESI, MeOH) m/z: 238.9–240.9 $[M-H]^-$. HRMS (ESI) calcd for $C_9H_4BrO_3$ $[M-H]^-$, 238.9349; found, 238.9353.

*4.2.6. 2-Bromo-3-benzyloxy-chromone **8***

To a solution of 2-bromo-3-HC previously obtained (222 mg, 0.92 mmol) in DMF (3 mL) was added K_2CO_3 (2 equiv, 254 mg) and benzylbromide (2 equiv, 0.22 mL). The mixture was stirred overnight under N_2 atmosphere, then quenched by addition of water

and extracted with methylene chloride. The organic layer was dried ($MgSO_4$), concentrated and the residue was purified by silica gel column chromatography using 10% of ethyl acetate in cyclohexane to afford compound **8** (yellow oil, 275 mg, 90%). R_f=0.37 (cyclohexane/ethyl acetate: 90:10). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=5.26 (s, 2H, CH_2), 7.34–7.55 (m, 7H), 7.65 (ddd, 1H, J=1.8, 7.1 and 8.6 Hz), 8.25 (dd, 1H, J=1.8 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=74.6, 117.8, 124.1, 125.7, 126.4, 128.5, 129.1, 133.9, 136.4, 142.4, 156.2, 172.9. MS (ESI, MeOH) m/z: 352.8–354.8 $[M+Na]^+$. HRMS (ESI) calcd for $C_{16}H_{12}O_3Br$ $[M+H]^+$, 330.9969; found, 330.9975.

4.2.7. 2-Bromo-3-(2-methoxyethoxy)methoxy-chromone **9**

To a solution of 2-bromo-3-HC previously obtained (566 mg, 2.33 mmol) in DMF (6 mL) was successively added K_2CO_3 (2 equiv, 644 mg) and MEMCl (2 equiv, 0.53 mL). The mixture was stirred overnight under N_2 atmosphere, then quenched by addition of water and extracted three times with methylene chloride. The organic layer was dried ($MgSO_4$), concentrated and the residue was purified by silica gel column chromatography using 30% of ethyl acetate in cyclohexane to give compound **9** (white resin, 725 mg, 95%). R_f=0.47 (cyclohexane/ethyl acetate: 50:50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=3.37 (s, 3H, CH_3), 3.57–3.62 (m, 2H, CH_2), 4.00–4.06 (m, 2H, CH_2), 5.39 (s, 2H, CH_2), 7.38–7.49 (m, 2H), 7.68 (ddd, 1H, J=1.8, 7.1 and 8.6 Hz), 8.20 (dd, 1H, J=1.5 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=59.2, 69.5, 71.8, 96.5, 117.7, 124.0, 125.7, 126.4, 134.0, 142.2, 172.6. MS (ESI, MeOH) m/z: 351–353 $[M+Na]^+$. HRMS (ESI) calcd for $C_{13}H_{14}O_5Br$ $[M+H]^+$, 329.0024; found, 329.0020.

4.2.8. 3-Benzyloxy-2-(trimethysilyl-ethynyl)-chromone **10**

Note that the cleavage of TMS group of pure compounds **10** and **11** was observed. In addition, the corresponding terminal alkynes of **10** and **11** were unstable and underwent slow degradation in solution. Therefore, these compounds were stored at low temperature under N_2 atmosphere and deprotected and used in situ in the next coupling step.

To a solution of compound **8** (245 mg, 0.74 mmol), trimethylsilyl-acetylene (0.615 mL, 6 equiv), TEA (2 mL, 13.86 mmol, 20 equiv) in toluene (12 mL) was added $PdCl_2(PPh_3)_2$ (52 mg, 0.1 equiv), CuI (26 mg, 0.2 equiv) under N_2 atmosphere. The mixture was stirred at 120 °C for 1 h, filtered through Celite and washed several times with ethyl acetate. The filtrate was evaporated and the obtained residue was purified by flash chromatography on silica gel, using 5% of ethyl acetate in cyclohexane, to afford **10** as yellow oil (235 mg, 92%). Compound **11** was prepared in similar manner. Analysis of compound **10** and the corresponding terminal alkyne (partial TMS-cleavage in solution). R_f=0.26 (cyclohexane/ethyl acetate: 90:10). ^{1}H NMR of a mixture of **10** and its acetylenic derivative ($CDCl_3$, 200 MHz) δ=0.31 (s, 5H, TMS), 3.79 (s, 0.3H, H-alkyne), 5.32 and 5.34 (2s, 2H, CH_2), 7.32–7.54 (m, 7H), 7.61–7.70 (m, 1H), 8.20 (dd, 1H, J=1.0 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=0.0, 74.8, 89.8, 93.9, 109.8, 118.1, 124.7, 124.9, 125.1, 125.9, 126.0, 128.3, 128.4, 128.5, 128.8, 129.0, 130.6, 130.8, 132.1, 132.3, 133.8, 134.0, 134.3, 134.6, 136.6, 136.9, 155.6, 174.3. MS (ESI, MeOH) m/z: 370.8 $[M+Na]^+$.

4.3. Synthesis of the thienyl 3-hydroxychromone nucleoside 1a

Path A. Aryl–aldol condensation.

4.3.1. 1-(5-Bromothiophen-2-yl)-2-((4R,5R)-5-((tert-butyldiphenylsilyloxy)methyl)-2,2-dimethyl-1,3-dioxolan-4-yl)ethanol **12**

To a solution of 2-bromothiophene (4 mmol), in THF (30 mL) was slowly added LDA (6 mmol) at −78 °C. After 45 min, the aldehyde **2** was added dropwise (1.5 mmol in 2 mL of THF). After completion of the reaction (TLC monitoring), the mixture was quenched with saturated aq NH_4Cl solution, extracted with methylene chloride, dried over $MgSO_4$, filtered and concentrated in vacuo. The obtained residue was purified by flash chromatography on silica gel (eluting with cyclohexane/ethyl acetate: 90:10) to give compound **12** (yellow oil, *R*/*S*=40:60, 75%). R_f=0.53 (cyclohexane/ethyl acetate: 70:30). Major isomer: ^{1}H NMR ($CDCl_3$, 200 MHz) δ=1.08 (s, 9H, *t*-Bu), 1.35 (s, 3H, CH_3), 1.41 (s, 3H, CH_3), 2.14–2.20 (m, 2H, H-2′), 3.41 (br d, 1H, J=6.3 Hz, OH), 3.70–3.74 (m, 2H, 2H-5′), 4.26 (dd, 1H, J=5.9 and 12.8 Hz, H-4′), 4.53 (dd, 1H, J=6.0 and 13.1 Hz, H-3′), 5.12 (dd, 1H, J=5.5 and 11.1 Hz, H-1′), 6.68 (dd, J=0.9 and 3.7 Hz, H-thiophene), 6.91–6.93 (d, J=3.7 Hz, H-thiophene), 7.39–7.43 (m, 6H), 7.63–7.69 (m, 4H). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=19.3, 25.6, 27.0, 28.1, 29.8, 30.3, 37.6, 39.0, 62.4, 68.1, 70.4, 74.2, 76.8, 77.5, 108.4, 111.1, 123.3, 127.9, 129.6, 130.0, 133.0, 133.1, 135.6, 135.7, 150.8. MS (ESI, MeOH) m/z: 596.7–598.7 $[M+Na]^+$.

4.3.2. α and β-1-(5-Bromothiophen-2-yl)-3-hydroxy-5-O-tert-butyldiphenylsilyl-2-deoxy-D-ribofuranose **13**

To a solution of **12** (*R*/*S*=40:60) (1.51 mmol) in toluene (40 mL) was added *p*-toluenesulfonic acid (0.2 equiv, 0.3 mmol). The mixture was stirred at 50 °C for 1 h then quenched with a saturated solution of sodium hydrogencarbonate and extracted with methylene chloride. The combined organic layers were dried ($MgSO_4$) and evaporated under reduced pressure to give a crude oil. Silica gel column chromatography purification (cyclohexane/ethyl acetate: 90:10 to 50:50) afforded **13** as a yellow oil (740 mg, 85%, α/β=40:60). R_f=0.35 (cyclohexane/ethyl acetate: 70:30). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=1.02 (s, 9H, *t*-Bu), 2.00–2.32 (m, 2H, 2H-2′), 3.63–4.08 (m, 3H, 2H-5′and H-4′), 4.53–4.60 (m, 1H, H-3′), 5.26–5.34 (m, 1H, H-1′), 6.72 (dd, J=0.8 and 3.7 Hz, H-thiophene), 6.87 (d, J=3.7 Hz, H-thiophene), 7.35–4.10 (m, 6H), 7.67–7.80 (m, 4H). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=19.4, 27.0, 43.0, 43.7, 64.7, 64.9, 74.4, 74.6, 76.2, 85.6, 87.4, 111.7 (C–Br), 124.5, 124.7, 127.9, 127.9, 127.9, 129.4, 129.6, 129.9, 130.0, 133.2, 134.9, 135.6, 135.7, 135.8, 147.1. MS (ESI, MeOH) m/z: 539.2–541.2 $[M+Na]^+$. HRMS (ESI) calcd for $C_{25}H_{28}BrO_3SSi$ $[M-H]^-$, 515.0717; found, 515.0723.

Path B. Friedel–Crafts glycosylation.

4.3.3. β-1-(5-Bromothiophen-2-yl)-3,5-di-O-acetyl-2-deoxy-D-ribofuranose **14β**

To a stirred solution of acetyl-deoxyribose **3a** or **3b** (8.6 mmol) in CH_2Cl_2 (50 mL) and 2-bromothiophene (2 equiv) was added dropwise $SnCl_4$ (1 equiv) at 0 °C. The mixture was stirred 30 min, then quenched with saturated aqueous $NaHCO_3$ solution and extracted with methylene chloride (3×50 mL). The combined organic layers were dried over $MgSO_4$, concentrated and the crude product was purified on silica gel chromatography (cyclohexane/ethyl acetate: 95:5 to 80:20) to give **14** (2.78 g, α/β=30:70, 73% combined yield). Compound **14-β**: R_f=0.55 (cyclohexane/ethyl acetate: 50:50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=2.09 (s, 3H, CH_3), 2.10 (s, 3H, CH_3), 2.17 (ddd, 1H, J=13.8, 10.7 and 6.0 Hz, $H_{2'}$), 2.36 (dd, 1H, J=13.8, and 5.8 Hz, $H_{2'}$), 4.18 (m, 2H, $H_{4'}$ and $H_{5'}$), 4.32 (dd, 1H, J=13.2 and 5.2 Hz, $H_{5'}$), 5.22 (d, J=6.0 Hz, 1H, $H_{3'}$), 5.26 (dd, 1H, J=10.7 and 5.2 Hz, $H_{1'}$), 6.75 (dd, 1H, J=3.7 and 0.6 Hz, H-thiophene), 6.89 (d, 1H, J=3.7 Hz, H-thiophene). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=21.0, 21.2, 41.5, 64.3, 76.5, 76.7, 82.8, 112.3, 125.3, 129.5, 145.7, 170.6, 170.8. MS (ESI, MeOH) m/z: 384.9–386.9 $[M+Na]^+$, 400.8–402.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{13}H_{14}BrO_5S$ $[M-H]^-$, 360.9750; found, 360.9740.

4.3.4. β-1-(5-Bromothiophen-2-yl)-2-deoxy-D-ribofuranose (diol intermediate)

To a solution of **14-β** (835 mg, 2.3 mmol) in MeOH (12 mL) was added K_2CO_3 (3 equiv, 953 mg) and the mixture was stirred at room

temperature for 1 h. The solution was evaporated in vacuo and the residue was purified by flash chromatography (methylene chloride/methanol: 90:10) to give the diol as a foam (630 mg, 98%). The spectral data of this product are in accordance with those recently reported by M. Hocek et al. (Ref. 2) (see ^{1}H and ^{13}C spectral data). HRMS (ESI) calcd for $C_9H_{10}BrO_3S$ $[M-H]^-$, 276.9539; found, 276.9528.

4.3.5. β-1-(5-Bromothiophen-2-yl)-3,5-di-O-tert-butyldiphenylsilyl-2-deoxy-D-ribofuranose **15**

To a solution of diol previously obtained (610 mg, 2.19 mmol) in dry DMF (11 mL) were successively added imidazole (3.5 equiv) and TBDPSCl (3.5 equiv, 1.97 mL) under N_2 atmosphere. After stirring for 24 h, the reaction mixture was quenched with a saturated solution of NH_4Cl and then extracted three times with methylene chloride. The combined organic layers were dried over $MgSO_4$, filtered and evaporated. The residue was purified by silica gel column (cyclohexane/ethyl acetate: 90:10) to afford **15** as a colourless oil (1.57 g, 95%). R_f=0.8 (cyclohexane/ethyl acetate: 80:20). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=0.85 (s, 9H, *t*-Bu), 1.01 (s, 9H, *t*-Bu), 1.81 (ddd, 1H, J=12.6, 10.9 and 5.2 Hz, $H_{2'}$), 1.81 (dd, 1H, J=11.4 and 5.1 Hz, $H_{2''}$), 3.24 (dd, 1H, J=11.0 and 4.0 Hz, $H_{5'}$), 3.40 (dd, 1H, J=11.0 and 4.0 Hz, $H_{5''}$), 4.00 (dt, 1H, J=3.7 and 1.1 Hz, $H_{4'}$), 4.32 (br d, 1H, J=5.0 Hz, $H_{3'}$), 4.96 (dd, 1H, J=10.9 and 5.1 Hz, 1H, $H_{1'}$), 6.64 (dd, 1H, J=3.7 and 0.6 Hz, H-thiophene), 6.77 (d, 1H, J=3.7, H-thiophene), 7.115–7.65 (m, 20H, Ar). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=19.2, 26.9, 27.1, 44.7, 64.3, 75.6, 88.5, 127.7, 127.7, 127.8, 127.9, 129.3, 129.8, 130.0, 133.2, 133.7, 133.8, 135.9, 147.2. MS (ESI, MeOH) *m*/*z*: 777.4–779.4 $[M+Na]^+$. HRMS (ESI) calcd for $C_{41}H_{46}BrO_3SSi_2$ $[M-H]^-$, 753.1895; found, 753.1884.

4.3.6. β-1-(5-Tributylstannylthiophen-2-yl)-3,5-di-O-tert-butyldiphenylsilyl-2-deoxy-D-ribofuranose (intermediate)

To a solution of **15** (2.08 mmol) in dry ether (8 mL) was added dropwise *n*-BuLi (1.1 equiv, 1.6 M solution) at 0 °C, under N_2 atmosphere. After 1 h, the mixture was cooled to −78 °C and tributyltin chloride (1.1 equiv) was slowly added and the mixture was stirred overnight. The reaction mixture was quenched with saturated NH_4Cl solution and extracted with ether. The combined organic layers were dried over $MgSO_4$, filtered and concentrated in vacuo. The obtained crude product was used in the next step without further purification.

4.3.7. β-1-(5-(3-Hydroxy-chromone-2)-thiophen-2-yl)-3,5-di-O-tert-butyldiphenylsilyl-2-deoxy-D-ribofuranose **16**

To a stirred solution of 2-bromo-3-HC **9** (284 mg, 0.86 mmol) and tin derivative (2 mmol) in toluene (10 mL) under N_2 atmosphere were successively added $Pd(PPh_3)_4$ (50 mg) and CuI (16 mg). The mixture was stirred at 120 °C for 2 h, filtered through Celite and concentrated in vacuo. The crude residue was then purified by flash chromatography on silica gel (eluting with cyclohexane/ethyl acetate: 90:10) to afford **16** as a yellow oil (573 mg, 72%). R_f=0.30 (cyclohexane/ethyl acetate: 70:30). ^{1}H NMR (acetone-d_6, 200 MHz) δ=0.87 (s, 9H, *t*-Bu), 1.02 (s, 9H, *t*-Bu), 1.97 (m, 1H, $2H_{2'}$), 2.29 (dd, 1H, J=12.6 and 5.0 Hz, $H_{2''}$), 3.02 (s, 3H, CH_3), 3.19 (t, 2H, J=4.6 Hz, CH_2 MEM), 3.30 (dd, 1H, J=11.0 and 3.7 Hz, $H_{5'}$), 3.47 (dd, 1H, J=11.1 and 3.8 Hz, $H_{5''}$), 3.62 (t, 2H, J=4.6 Hz, CH_2 MEM), 4.05 (dt, 1H, J=3.5 and 1.3 Hz, $H_{4'}$), 4.58 (d, 1H, J=5.1 Hz, $H_{3'}$), 5.37 (s, 2H, OCH_2O), 5.47 (dd, 1H, J=10.4 and 5.0 Hz, $H_{1''}$), 7.08 (d, 1H, J=3.9, H-thiophene), 7.15–7.70 (m, 23H, Ar), 7.85 (d, 1H, J=3.9 Hz, H-thiophene), 8.02 (dd, 1H, J=7.8 and 1.4 Hz, H-chromone). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=20.0, 20.1, 27.6, 27.8, 46.2, 59.2, 65.5, 71.0, 72.7, 76.9, 77.7, 89.8, 97.0, 119.1, 125.3, 126.0, 126.0, 126.5, 129.0, 129.1, 129.2, 131.0, 131.1, 131.3, 131.5, 132.0, 134.2, 134.3, 134.8, 134.9, 136.4, 136.7, 136.8, 137.0, 151.7, 154.8, 158.1, 174.0. MS (ESI, MeOH) *m*/*z*: 947.4 $[M+Na]^+$.

4.3.8. β-1-(5-(3-Hydroxy-chromone-2)-thiophen-2-yl)-2-deoxy-D-ribofuranose **1a**

To a solution of MEM-protected nucleoside **16** previously obtained (140 mg, 0.15 mmol) in dioxane (3 mL) was added HCl 6 N (1.5 mL). The mixture was stirred at room temperature overnight then neutralized by saturated aqueous $NaHCO_3$ solution and extracted with ethyl acetate (3×10 mL). The organic layers were evaporated in vacuo and the residue was purified by silica gel column chromatography (methylene chloride/methanol: 95:5 to 90:10) to give the free 3-HC-nucleoside **1a** as a pale resin (46 mg, 85%). R_f=0.26 (9:1 CH_2Cl_2/MeOH). ^{1}H NMR (DMSO-d_6, 500 MHz) δ=1.85 (ddd, 1H, J=2.1, 11.3 and 13.7 Hz, $H_{2'}$), 2.12 (ddd, 1H, J=2.1 and 3.8 and 13.7 Hz, $H_{2''}$), 3.30–3.62 (m, 1H, $H_{5'}$), 3.65–3.67 (m, 2H, $H_{5''}$and $H_{4'}$), 3.99 (br m, 1H, $H_{3'}$), 4.97 (dd, 1H, J=2.1 Hz and 11.3 Hz, $H_{1'}$), 7.15 (d, 1H, J=3.9 Hz, H-thiophene), 7.47 (t, 1H, J=7.4 Hz, H-6), 7.71 (d, 1H, J=8.3 Hz, H-8), 7.81 (m, 1H, H-7), 7.82 (d, 1H, J=3.9 Hz, H-thiophene), 8.11 (dd, 1H, J=1.2 and 7.4 Hz, H-5). ^{13}C NMR (DMSO-d_6, 125 MHz) δ=42.4, 67.8, 68.2, 68.7, 71.0, 120.0, 123.7, 126.0, 126.5, 126.7, 129.9, 132.8, 135.5, 138.5, 145.2, 152.8, 156.1, 173.9. MS (ESI, MeOH) *m*/*z*: 358.9 $[M-H]^-$. HRMS (ESI) calcd for $C_{18}H_{17}O_6S$ $[M+H]^+$, 361.0745; found, 361.0741. IR (KBr) ν: 3432, 1651 cm^{-1}.

4.4. Synthesis of the triazolyl-3-hydroxy-chromone 1b

4.4.1. β-1-(4-(3-Benzyloxy-chromone-2)-triazol-1-yl)-3,5-di-O-tolouyl-2-deoxy-D-ribofuranose **17**

Method A. To a stirred solution of azido-sugar **4** (1 mmol), alkyne **10** (1.1 equiv) and *n*-Bu_4NF (1.1 equiv) in methylene chloride (5 mL) were successively added CuI (164 mg, 2 equiv), DIEA (0.37 mL, 5 equiv) and acetic acid (1 equiv). The mixture was stirred 4 h at room temperature, filtered and the solvent removed in vacuo. The crude product was purified by flash chromatography (cyclohexane/ethyl acetate: 80:20) to give **17** (598 mg, 89%).

Method B. A mixture of azido-sugar **4** (1 mmol), alkyne **10** (1.1 equiv), *n*-Bu_4NF (1.1 equiv), CuI (164 mg, 2 equiv), DIEA (0.37 mL, 5 equiv) was adsorbed on silica gel (1 g) using methylene chloride. After evaporation, the resulting yellow powder was placed into a microwave and irradiated for 2 min. The mixture was eluted twice with ethyl acetate and the solvent evaporated under reduced pressure to give a crude product, which was subjected to a simple filtration over silica gel (cyclohexane/ethyl acetate: 80:20) to give **17** (638 mg, 95%). Mp (methylene chloride/ether)=183–185 °C. R_f=0.56 (cyclohexane/ethyl acetate: 50:50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ=2.25 (s, 3H, CH_3), 2.44 (s, 3H, CH_3), 2.82–2.92 (m, 1H, $H_{2'}$), 3.11–3.24 (m, 1H, $H_{2''}$), 4.50 (d, 2H, J=4.4 Hz, $2H_{5'}$), 4.63–4.67 (m, 1H, $H_{4'}$), 5.21 and 5.32 (2d, 2H, J=11.0 Hz, CH_2), 5.69–5.74 (m, 1H, $H_{3'}$), 6.42 (t, 1H, J=5.7 Hz, $H_{1'}$), 7.13 (d, 2H, J=8.0 Hz), 7.30–7.47 (m, 8H), 7.75 (m, 2H), 7.90 (d, 2H, J=8.2 Hz), 7.95 (d, 2H, J=8.2 Hz), 8.30 (d, 2H, J=8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ=21.7, 21.9, 38.4, 63.7, 74.4, 74.6, 83.9, 89.2, 118.7, 124.5, 125.0, 125.8, 126.5, 128.7, 129.0, 129.3, 129.4, 129.8, 129.9, 133.8, 136.8, 139.2, 144.2, 144.7, 155.3, 165.9, 166.2, 174.4. MS (ESI, MeOH) *m*/*z*=709.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{39}H_{33}N_3O_8$ $[M+H]^+$, 672.2345; found, 672.2341.

4.4.2. β-1-(4-(3-Benzyloxy-chromone-2)-triazol-1-yl)-2-deoxy-D-ribofuranose (intermediate)

By the same procedure as described above for the synthesis of **14**, compound **17** (100 mg, 0.15 mmol) was deprotected using K_2CO_3 in MeOH to afford the pure product (62 mg, 95%). Mp (methylene chloride/ether)=94–96 °C. R_f=0.24 (9:1 CH_2Cl_2/MeOH). ^{1}H NMR (MeOD, 200 MHz) δ=2.47–2.78 (m, 2H, $2H_{2'}$), 3.60 (dd, 1H, J=4.6 and 11.9 Hz, $1H_{5'}$), 3.70 (dd, 1H, J=3.8 and 11.9 Hz, $1H_{5''}$), 4.03 (dd, 1H, J=4.2 and 8.2 Hz, $H_{4'}$), 4.48–4.56 (dd, 1H, J=5.5 and 10.0 Hz, $H_{3'}$), 5.28 (s, 2H, CH_2), 6.42 (t, 1H, J=5.9 Hz, $H_{1'}$), 7.20–7.35 (m, 4H), 7.44 (td, 1H, J=1.1 and 8.1 Hz), 7.64 (d, 1H, J=8.4 Hz), 7.75 (ddd, 1H,

J=1.5, 7.0 and 8.4 Hz), 7.85 (d, 1H, J=8.2 Hz), 8.15 (dd, 1H, J=1.3 and 8.1 Hz), 8.71 (s, 1H, H-5). ^{13}C NMR (MeOD, 50 MHz) δ=42.0, 56.0, 63.0, 72.1, 75.0, 89.9, 90.7, 119.5, 125.1, 126.2, 126.4, 127.3, 129.4, 129.7, 130.0, 130.5, 130.7, 135.4, 137.6, 139.2, 139.5, 156.5, 176.1. MS (ESI, MeOH) m/z=457.9 $[M+Na]^+$, 473.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{23}H_{22}N_3O_6$ $[M+H]^+$, 436.1508; found, 436.1509.

4.4.3. β-1-(4-(3-Hydroxy-chromone-2)-triazol-1-yl)-2-deoxy-D-ribofuranose **1b**

To a degassed solution of the above triazolyl-compound intermediate (0.3 mmol) in THF (10 mL) was added Pd/C (10% molar). Hydrogenolysis was then conducted in a high pressure autoclave (Parr apparatus, 3 bar) for 10 h. The catalyst was filtered and washed with THF. The filtrate was concentrated in vacuo and the obtained residue was purified by flash chromatography to afford **1b** as a white powder (34 mg, 90%). R_f=0.32 (8:2 CH_2Cl_2/MeOH). ^{1}H NMR (DMSO-d_6, 500 MHz) δ=2.42–2.47 (m, 1H, $H_{2'}$), 2.73 (ddd, 1H, J=5.8 and 13.5 Hz, $H_{2'}$), 3.45 (dd, 1H, J=4.8 and 11.7 Hz, $H_{5'}$), 3.55 (dd, 1H, J=4.3 and 11.7 Hz, $H_{5'}$), 3.91 (dd, 1H, J=4.8 and 8.8 Hz, $H_{4'}$), 4.43–4.44 (m, 1H, $H_{3'}$), 4.87 (br s, 1H, $OH_{5'}$), 5.36 (br s, 1H, $OH_{3'}$), 6.52 (t, 1H, J=5.9 Hz, $H_{1'}$), 7.48 (ddd, 1H, J=1.0, 7.0 and 8.0 Hz), 7.73 (d, 1H, J=8.2 Hz), 7.81 (ddd, 1H, J=1.6, 7.0 and 8.6 Hz), 8.13 (dd, 1H, J=1.6 and 8.0 Hz), 8.81 (s, 1H, H-5), 10.00 (s, 1H, OH). ^{13}C NMR (DMSO-d_6, 125 MHz) δ=40.0, 61.5, 70.4, 88.4, 88.5, 118.4, 122.0, 124.7, 124.9, 125.3, 133.8, 137.9, 138.0, 140.3, 154.5, 172.1. MS (ESI, MeOH) m/z: 343.7 $[M-H]^-$, 368.0 $[M+Na]^+$. HRMS (ESI) calcd for $C_{16}H_{16}N_3O_6$ $[M+H]^+$, 346.1039; found, 346.1042. IR (KBr) ν: 3392, 1621 cm^{-1}.

Acknowledgements

The authors thank J-M. Guigonis IFR 50 for HRMS spectra, ANR, CNRS and Ministère de l'Education Nationale et de la Recherche for financial support and for the doctoral fellowships of M.S. (MENRT) and V.Y.P. (ANR). V.V.S. is supported by an Eiffel doctoral fellowship.

Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.tet.2009.07.021.

References and notes

1. For recent reviews see: (a) Tor, I. *Tetrahedron* **2007**, *63*, 3415 (special issue, see contribution of other authors); (b) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265; (c) Rist, M. J.; Marino, J. P. *Curr. Org. Chem.* **2002**, *6*, 775.
2. For recent articles see: (a) Grigorenko, N. A.; Leumann, C. J. *Chem.—Eur. J.* **2009**, *15*, 639; (b) Lee, B. J.; Barch, M.; Castner, E. W.; Völker, J.; Breslauer, K. J. *Biochemistry* **2007**, *46*, 10756; (c) Ramreddy, T.; Rao, B. J.; Krishnamoorthy, G. *J. Phys. Chem. B* **2007**, *111*, 5757; (d) DeLucia, A. M.; Grindley, N. D. F.; Joyce, C. M. *Biochemistry* **2007**, *46*, 10790; (e) Lenz, T.; Bonnist, E. Y. M.; Pljevaljić, G.; Neely, R. K.; Dryden, D. T.; Scheidig, A. J.; Jones, A. C.; Weinhold, E. *J. Am. Chem. Soc.* **2007**, *129*, 6240; (f) Reha, D.; Hocek, M.; Hobza, P. *Chem. Eur. J.* **2006**, *12*, 3587; (g) Dupradeau, F.-Y.; Case, D. A.; Yu, C.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2005**, *127*, 15612. See also Ref. 1.
3. (a) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graslund, A.; McLaughlin, L. W. *Biochemistry* **1989**, *28*, 9095; (b) Guest, C. R.; Hochstrasser, R. A.; Dupuy, C. G.; Allen, D. J.; Benkovic, S. J.; Millar, D. P. *Biochemistry* **1991**, *30*, 8759.
4. (a) Ben Gaied, N.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mely, Y. *Nucleic Acids Res.* **2005**, *33*, 1031; (b) Kenfack, C.; Burger, A.; Mely, Y. *J. Phys. Chem. B* **2006**, *110*, 26327; (c) Kenfack, C.; Piemont, E.; Ben Gaied, N.; Burger, A.; Mély, Y. *J. Phys. Chem. B* **2008**, *112*, 9736.
5. (a) Shynkar, V. V.; Klymchenko, A. S.; Kunzelmann, C.; Duportail, G.; Muller, C. D.; Demchenko, A. P.; Freyssinet, J. M.; Mely, Y. *J. Am. Chem. Soc.* **2007**, *129*, 2187; (b) Shynkar, V. V.; Klymchenko, A. S.; Duportail, G.; Demchenko, A. P.; Mely, Y. *Biochim. Biophys. Acta* **2005**, *1712*, 128; (c) Klymchenko, A. S.; Duportail, G.; Mely, Y.; Demchenko, A. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11219; (d) Enander, K.; Choulier, L.; Olsson, A. L.; Yushchenko, D. A.; Kanmert, D.; Klymchenko, A. S.; Demchenko, A. P.; Mely, Y.; Altschuh, D. *Bioconjugate Chem.* **2008**, *19*, 1864.
6. Formosinho, J. S.; Arnaut, G. L. J. *Photochem. Photobiol. A: Chem.* **1993**, *75*, 21.
7. Sengupta, P. K.; Kasha, M. *Chem. Phys. Lett.* **1979**, *68*, 382.
8. (a) Chou, P. T.; Martinez, M. L.; Clements, J. H. *J. Phys. Chem.* **1993**, *97*, 2618; (b) Klymchenko, A. S.; Demchenko, A. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 461.
9. McMorrow, D.; Kasha, M. *J. Phys. Chem.* **1984**, *88*, 2235.
10. Klymchenko, A. S.; Demchenko, A. P. *J. Am. Chem. Soc.* **2002**, *124*, 12372.
11. Klymchenko, A. S.; Kenfack, C.; Duportail, G.; Mély, Y. *J. Chem. Sci.* **2007**, *119*, 83.
12. Klymchenko, A. S.; Shvadchak, V. V.; Yushchenko, D. A.; Jain, N.; Mély, Y. *J. Phys. Chem. B* **2008**, *112*, 12050.
13. Pleier, A.-K.; Glas, H.; Grosche, M.; Sirsch, P.; Thiel, W. R. *Synthesis* **2001**, *1*, 55.
14. Bernini, R.; Mincione, E.; Coratti, A.; Fabrizi, G.; Battistuzzi, G. *Tetrahedron* **2004**, *60*, 967.
15. (a) Spadafora, M.; Burger, A.; Benhida, R. *Synlett* **2008**, 1225; (b) Peyron, C.; Benhida, R. *Synlett* **2009**, 472.
16. The Friedel–Craft C-glycosylation approach was published simultaneously by some of us (Ref. 16a) and by M. Hocek group (Ref. 16b): (a) Spadafora, M.; Mehiri, M.; Burger, A.; Benhida, R. *Tetrahedron Lett.* **2008**, *49*, 3967; (b) Bárta, J.; Pohl, R.; Klepetárová, B.; Ernsting, N. P.; Hocek, M. *J. Org. Chem.* **2008**, *73*, 3798; (c) Hainke, S.; Arndt, S.; Seitz, O. *Org. Biomol. Chem.* **2005**, *3*, 4233; For other contributions see: (d) Joubert, N.; Pohl, R.; Klepetárová, B.; Hocek, M. *J. Org. Chem.* **2007**, *72*, 6797; (e) Hocek, M.; Pohl, R.; Klepetárová, B. J. *Eur. J. Org. Chem.* **2005**, 4525; (f) Guianvarc'h, D.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S.; Benhida, R. *Org. Lett.* **2002**, *4*, 4209; (g) Guianvarc'h, D.; Fourrey, J.-L.; TranHuuDau, M.-E.; Guérineau, V.; Benhida, R. *J. Org. Chem.* **2002**, *67*, 3724; (h) The spectral data of these analogues are in accordance with those reported by M. Hocek, Ref. 16b (see Supplementary data).
17. (a) Guezguez, R.; Bougrin, K.; EL Akri, K.; Benhida, R. *Tetrahedron Lett.* **2006**, *47*, 4807; (b) EL Akri, K.; Bougrin, K.; Balzarini, J.; Faraj, A.; Benhida, R. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6656; (c) Broggi, J.; Joubert, N.; Díez-González, S.; Berteina-Raboin, S.; Zevaco, T.; Nolan, S. P.; Agrofoglio, L. A. *Tetrahedron* **2009**, *65*, 4525; (d) Broggi, J.; Kumamoto, H.; Berteina-Raboin, S.; Nolan, S. P.; Agrofoglio, L. A. *Eur. J. Org. Chem.* **2009**, 1880; (e) Pradere, U.; Roy, V.; McBrayer, T. M.; Schinazi, R. F.; Agrofoglio, L. A. *Tetrahedron* **2008**, *64*, 9044.
18. Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319.
19. Kamlet, M. J.; Abboud, J. L. M.; Abraham, M. H.; Taft, R. W. *J. Org. Chem.* **1983**, *48*, 2877.
20. Bader, A. N.; Pivovarenko, V. G.; Demchenko, A. P.; Ariese, F.; Gooijer, C. *J. Phys. Chem. B* **2004**, *108*, 10589.
21. Eastman, J. W. *Photochem. Photobiol.* **1967**, *6*, 55.

SUPPORTING INFORMATION

Efficient Synthesis of Ratiometric Fluorescent Nucleosides Featuring 3-Hydroxychromone Nucleobases.

Marie Spadafora,[†] Vitoria Y. Postupalenko,[‡] Volodymyr V. Shvadchak,[‡] Andrey S. Klymchenko,[‡] Yves Mély,[‡] Alain Burger,[†]* and Rachid Benhida[†]*

[†] *Laboratoire de Chimie des Molécules Bioactives et des Arômes, UMR 6001 CNRS, Institut de Chimie de Nice, Université de Nice-Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France.*

[‡] *Département de Pharmacologie et Physico-Chimie des Interactions Cellulaires et Moléculaires, UMR CNRS 7175, Faculté de Pharmacie, Université Louis Pasteur, 74 Route du Rhin, BP 24, 67401 Illkirch cedex, France.*

Corresponding Authors : benhida@unice.fr and burger@unice.fr

Table of Contents

1. Experimental procedures and data for the compounds

2. NMR spectra of compounds

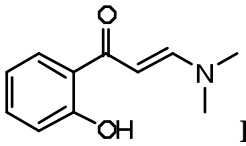
General information

All reactions were run under nitrogen atmosphere in dried glassware. Solvents were dried and distilled by standard procedures. Toluene was purchased from commercial sources and dried over 4Å molecular sieves before use. Reagents were purchased and used without further purification. All reactions were monitored by thin layer chromatography (TLC) plates (0.2 mm, silica gel 60 with fluorescent indicator UV_{254}). Flash chromatography was performed using silica gel (60, 0.063-0.200 mm).

^{1}H NMR and ^{13}C NMR were recorded on 200 and 500 instruments (200 and 500 MHz for ^{1}H, 50 and 125 MHz for ^{13}C). Chemical shifts (δ) were reported in ppm and the coupling constants were reported in Hertz (Hz).

Analytic High Performance Liquid Chromatography (HPLC) was recorded using a RP-C18 column (300A, 5 μm particle size).

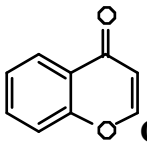
Synthesis of 3-Hydroxychromone precursors :



Enamide intermediate

A mixture of 2'-hydroxyacetophenone (5 mmol, 0.6 mL) and *N*,*N*-dimethylformamide-dimethylacetal (1 equiv, 0.66 mL) was irradiated under microwave for 15 sec (300 W max, T = 115°C). The resulting mixture was cooled at room temperature and crystallized in pentane to give the enamine intermediate (red crystals, 955 mg, 100%).

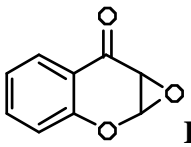
Mp (ether) = 130-131°C (Lit. 132-134, ref. 1). R*f* = 0.2 (cyclohexane/ethyl acetate : 60/40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 2.95 (s, 3H, NMe), 3.17 (s, 3H, NMe), 5.76 (d, 1H, *J* = 12.1 Hz), 6.81 (dt, 1H, *J* = 1.0 and 7.1 Hz), 6.92 (dd, 1H, *J* = 0.9 and 8.4 Hz), 7.30 (dt, 1H, *J* = 1.5 and 8.4 Hz), 7.69 (dd, 1H, *J* = 1.4 and 8.0 Hz), 7.87 (d, 1H, *J* = 12.0 Hz), 13.99 (s, 1H, OH). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 37.5, 45.5, 90.1, 118.1, 118.3, 120.4, 128.3, 134.0, 154.9, 163.0, 191.6. MS (ESI, MeOH) m/z: 213.70 $[M+Na]^+$.



Chromone 6

To a solution of the enamide previously obtained (1g, 5.24 mmol) in methylene chloride (40 mL) was added concentrated HCl (4 mL). The resulting mixture was refluxed for 1h. After cooling, the mixture was extracted with methylene chloride (3x40 mL). The combined organic layers were washed with saturated $NaHCO_3$ solution, then with brine, dried over MgSO4, filtered and concentrated to afford chromone **5** (red crystals, 700 mg, 90%).

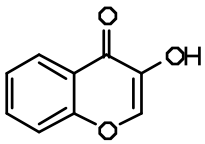
Mp (ether) = 52-54°C (Lit. 52°C, ref. 1). R*f* = 0.45 (cyclohexane/ethyl acetate : 60/40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 6.31 (d, 1H, *J* = 6.0 Hz), 7.33-7.44 (m, 2H), 7.63 (ddd, 1H, *J* = 1.8, 7.1 and 8.4 Hz), 7.83 (d, 1H, *J* = 6.0 Hz), 8.18 (dd, 1H, *J* = 1.5 and 7.9 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 113.1, 118.3, 125.0, 125.3, 125.9, 133.8, 155.4, 156.6, 177.7. MS (ESI, MeOH) m/z: 168.9 $[M+Na]^+$.



Epoxide intermediate

To a solution of chromone **6** (740 mg, 5 mmol) in methylene chloride (7.5 mL), was slowly added hydrogen peroxide (2 equiv, 35% solution, 1.4 mL) and NaOH (1.5 equiv, 407 mg) at 0°C. The mixture was stirred 3h at 0°C, quenched with water and extracted with methylene chloride (3x 20 mL). The combined organic layers were dried (Na_2SO_4) and evaporated under reduced pressure to give a crude amorphous solid (668 mg, 82%), which was used in the next step without purification.

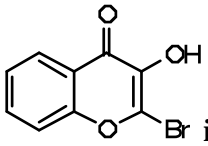
Mp (éther) = 64-66°C (ether). R*f* = 0.8 (cyclohexane/ethyl acetate : 60/40). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 3.69 (d, 1H, *J* = 2.4 Hz), 5.66 (d, 1H, *J* = 2.5 Hz), 7.05 (dd, 1H, *J* = 0.7 and 8.5 Hz), 7.14 (td, 1H, *J* = 1.0 and 8.0 Hz), 7.55 (ddd, 1H, *J* = 1.8, 7.2 and 8.5 Hz), 7.85 (dd, 1H, *J* = 1.8 and 8 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 55.4, 77.3, 118.1, 119.9, 123.4, 127.2, 136.4, 155.5, 188.2. MS (ESI, MeOH) m/z: 160.7 $[M-H]^-$.



7

To the epoxide previously obtained (570 mg, 3.5 mmol) was added conc. HCl (20 mL) and the resulting mixture was heated at 70°C for 1h. After cooling, water was added (20 mL) and the mixture was extracted with methylene chloride (3x 30 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ solution, dried over $MgSO_4$, filtered and concentrated to afford 3-Hydroxychromone **7** (brown powder, 568 mg, 100%).

Mp (ether) = 179-181°C. R*f* = 0.5 (cyclohexane/ethyl acetate : 50/50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 6.33 (s, 1H), 7.41 (ddd, 1H, *J* = 1.1, 7.0 and 8.0 Hz), 7.50 (dd, 1H, *J* = 0.6 and 8.6 Hz), 7.69 (ddd, 1H, *J* = 1.8 and 7.0 and 8.6 Hz), 8.01 (s, 1H), 8.26 (dd, 1H, *J* = 1.3 and 8 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 117.3, 123.4, 124.3, 132.3, 137.2, 140.5, 147.9, 155.1, 172.2. MS (ESI, MeOH) m/z: 184.9 $[M+Na]^+$. HRMS (ESI) calcd for $C_9H_7O_3$ $[M+H]^+$, 163.0395; found, 163.0389.

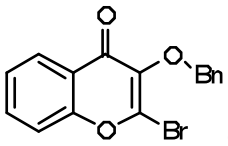


intermediate

To a solution of 3-HC **7** (540 mg, 3.34 mmol, 0.65 equiv) in acetonitrile (10 mL) was added NBS (1 equiv, 940 mg)) and vazo 67 (0.1 equiv, 100 mg), and the mixture was refluxed for 8h (0.1 equiv of vazo 67 was added each 2h). The mixture was cooled and the solvent evaporated

in vacuo. The crude product was purified on silica gel (cyclohexane/ethyl acetate : 90/10) to give the corresponding 2-Br-3HC derivative (colorless powder, 730 mg, 90%).

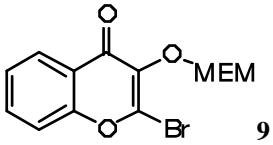
Mp (ether) = 178-180°C. R*f* = 0.57 (cyclohexane/ethyl acetate : 50/50). ^{1}H NMR (MeOD, 200 MHz) δ = 7.37 (ddd, 1H, *J* = 1.0, 7.1 and 8.1 Hz), 7.47 (dd, 1H, *J* = 0.6 and 8.6 Hz), 7.66 (ddd, 1H, *J* = 1.6, 7.0 and 8.6 Hz), 8.06 (dd, 1H, *J* = 1.7 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 118.1, 121.0, 125.4, 126.0, 129.8, 134.0, 140.3, 156.7, 171.3. MS (ESI, MeOH) m/z: 238.9-240.9 $[M-H]^-$. HRMS (ESI) calcd for $C_9H_4BrO_3$ $[M-H]^-$, 238.9349; found, 238.9353.



8

To a solution of 2-bromo-3-HC previously obtained (222 mg, 0.92 mmol) in DMF (3 mL) was added K_2CO_3 (2 equiv, 254 mg) and benzylbromide (2 equiv, 0.22 mL). The mixture was stirred overnight under N_2 atmosphere, then quenched by addition of water and extracted with methylene chloride. The organic layer was dried ($MgSO_4$), concentrated and the residue was purified by silica gel column chromatography using 10% of ethyl acetate in cyclohexane to afford compound **8** (yellow oil, 275 mg, 90%).

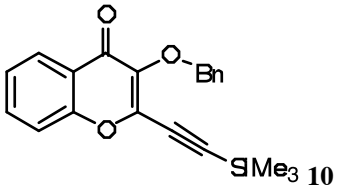
R*f* = 0.37 (cyclohexane/ethyl acetate : 90/10). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 5.26 (s, 2H, CH_2), 7.34-7.55 (m, 7H), 7.65 (ddd, 1H, *J* = 1.8, 7.1 and 8.6 Hz), 8.25 (dd, 1H, *J* = 1.8 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 74.6, 117.8, 124.1, 125.7, 126.4, 128.5, 129.1, 133.9, 136.4, 142.4, 156.2, 172.9. MS (ESI, MeOH) m/z: 352.8 - 354.8 $[M+Na]^+$. HRMS (ESI) calcd for $C_{16}H_{12}O_3Br$ $[M+H]^+$, 330.9969 ; found, 330.9975.



9

To a solution of 2-bromo-3-HC previously obtained (566 mg, 2.33 mmol) in DMF (6 mL) was successively added K_2CO_3 (2 equiv, 644 mg) and MEMCl (2 equiv, 0.53 mL). The mixture was stirred overnight under N_2 atmosphere, then quenched by addition of water and extracted three times with methylene chloride. The organic layer was dried ($MgSO_4$), concentrated and the residue was purified by silica gel column chromatography using 30% of ethyl acetate in cyclohexane to give compound **9** (white resin, 725 mg, 95%).

R*f* = 0.47 (cyclohexane/ethyl acetate : 50/50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 3.37 (s, 3H, CH_3), 3.57-3.62 (m, 2H, CH_2), 4.00-4.06 (m, 2H, CH_2), 5.39 (s, 2H, CH_2), 7.38-7.49 (m, 2H), 7.68 (ddd, 1H, *J* = 1.8, 7.1 and 8.6 Hz), 8.20 (dd, 1H, *J* = 1.5 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 59.2, 69.5, 71.8, 96.5, 117.7, 124.0, 125.7, 126.4, 134.0, 142.2, 172.6. MS (ESI, MeOH) m/z: 351 - 353 $[M+Na]^+$. HRMS (ESI) calcd for $C_{13}H_{14}O_5Br$ $[M+H]^+$, 329.0024 ; found, 329.0020.



10

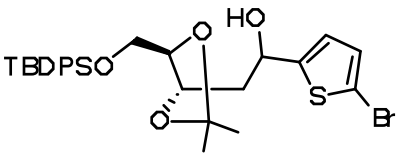
Note that the cleavage of TMS group of pure compounds **10** and **11** was observed. In addition, the corresponding terminal alkynes of **10** and **11** were unstable and underwent slow degradation in solution. Therefore, these compounds were stored at low temperature under N_2 atmosphere and used *in situ* in the next coupling step.

To a solution of compound **8** (245 mg, 0.74 mmol), trimethylsilyl-acetylene (0.615 mL, 6 equiv.), TEA (2 mL, 13.86 mmol, 20 equiv) in toluene (12 mL) was added $PdCl_2(PPh_3)_2$ (52 mg, 0.1 equiv), CuI (26 mg, 0.2 equiv) under N_2 atmosphere. The mixture was stirred at 120°C for 1h, filtered through Celite and washed several times with ethyl acetate. The filtrate was evaporated and the obtained residue was purified by flash chromatography on silica gel, using 5% of ethyl acetate in cyclohexane, to afford **10** as yellow oil (235 mg, 92%). Compound **11** was prepared in similar manner.

Analysis of compound **10** and the corresponding terminal alkyne (partial TMS-cleavage in solution). R*f* = 0.26 (cyclohexane/ethyl acetate : 90/10). ^{1}H NMR of a mixture of **10** and its acetylenic derivative ($CDCl_3$, 200 MHz) δ = 0.31 (s, 5H, TMS), 3.79 (s, 0.3H, H-alkyne), 5.32 and 5.34 (2s, 2H, CH_2), 7.32-7.54 (m, 7H), 7.61-7.70 (m, 1H), 8.20 (dd, 1H, *J* = 1.0 and 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 0.0, 74.8, 89.8, 93.9, 109.8, 118.1, 124.7, 124.9, 125.1, 125.9, 126.0, 128.3, 128.4, 128.5, 128.8, 129.0, 130.6, 130.8, 132.1, 132.3, 133.8, 134.0, 134.3, 134.6, 136.6, 136.9, 155.6, 174.3. MS (ESI, MeOH) m/z: 370.8 $[M+Na]^+$.

Synthesis of the thienyl 3-Hydroxychromone nucleoside 1a

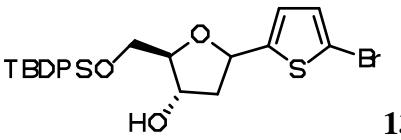
Path A : Aryl-aldol condensation



12

To a solution of 2-bromothiophene (4 mmol), in THF (30 mL) was slowly added LDA (6 mmol) at -78°C. After 45 min, the aldehyde **1** was added dropwise (1.5 mmol in 2 mL of THF). After completion of the reaction (TLC monitoring), the mixture was quenched with saturated aq NH_4Cl solution, extracted with methylene chloride, dried over $MgSO_4$, filtered and concentrated in vacuo. The obtained residue was purified by flash chromatography on silica gel (eluting with cyclohexane/ethyl acetate : 90/10) to give compound **12** (yellow oil, *R/S* = 40/60, 75%).

R*f* = 0.53 (cyclohexane/ethyl acetate : 70/30). Major isomer : ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 1.08 (s, 9H, *t*-Bu), 1.35 (s, 3H, CH_3), 1.41 (s, 3H, CH_3), 2.14-2.20 (m, 2H, H-2'), 3.41 (br d, 1H, J = 6.3 Hz, OH), 3.70-3.74 (m, 2H, 2H-5'), 4.26 (dd, 1H, *J* = 5.9 and 12.8 Hz, H-4'), 4.53 (dd, 1H, *J* = 6.0 and 13.1 Hz, H-3'), 5.12 (dd, 1H, *J* = 5.5 and 11.1 Hz, H-1'), 6.68 (dd, *J* = 0.9 and 3.7 Hz, H-thiophene), 6.91-6.93 (d, *J* = 3.7 Hz, H-thiophene), 7.39-7.43 (m, 6H), 7.63-7.69 (m, 4H). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 19.3, 25.6, 27.0, 28.1, 29.8,30.3, 37.6, 39.0, 62.4, 68.1, 70.4, 74.2, 76.8, 77.5, 108.4, 111.1, 123.3, 127.9, 129.6, 130.0, 133.0, 133.1, 135.6, 135.7, 150.8. MS (ESI, MeOH) m/z: 596.7 - 598.7 $[M+Na]^+$.



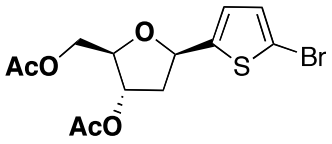
13

To a solution of **12** (*R/S* = 40/60) (1.51 mmol) in toluene (40 mL) was added *p*-toluenesulfonic acid (0.2 equiv, 0.3 mmol). The mixture was stirred at 50°C for 1h then quenched with a saturated solution of sodium hydrogencarbonate and extracted with methylene chloride. The combined organic layers were dried ($MgSO_4$) and evaporated under reduced pressure to give a crude oil. Silica gel column chromatography purification (cyclohexane / ethyl acetate : 90/10 to 50/50) afforded **13** as a yellow oil (740 mg, 85%, α/β = 40/60).

R*f* = 0.35 (cyclohexane/ethyl acetate : 70/30). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 1.02 (s, 9H, *t*-Bu), 2.00-2.32 (m, 2H, 2H-2'), 3.63-4.08 (m, 3H, 2H-5'and H-4'), 4.53-4.60 (m, 1H, H-3'), 5.26-5.34 (m, 1H, H-1'), 6.72 (dd, *J* = 0.8 and 3.7 Hz, H-thiophene), 6.87 (d, *J* = 3.7 Hz, H-thiophene), 7.35-4.10 (m, 6H), 7.67-7.80 (m, 4H). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 19.4, 27.0,

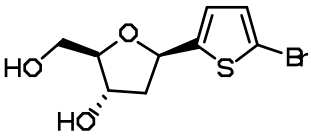
43.0, 43.7, 64.7, 64.9, 74.4, 74.6, 76.2, 85.6, 87.4, 111.7 (C-Br), 124.5, 124.7, 127.9, 127.9, 127.9, 129.4, 129.6, 129.9, 130.0, 133.2, 134.9, 135.6, 135.7, 135.8, 147.1. MS (ESI, MeOH) m/z: 539.2 - 541.2 $[M+Na]^+$. HRMS (ESI) calcd for $C_{25}H_{28}BrO_3SSi$ $[M-H]^-$, 515.0717; found, 515.0723.

Path B : Friedel-Crafts glycosylation



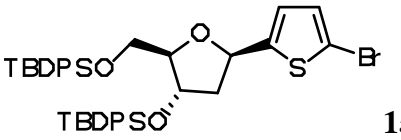
14-β

To a stirred solution of acetyl-deoxyribose **3a** or **3b** (8.6 mmol) in CH_2Cl_2 (50 mL) and 2-bromothiophene (2 equiv) was added dropwise $SnCl_4$ (1 equiv) at 0°C. The mixture was stirred 30 min, then quenched with saturated aqueous $NaHCO_3$ solution and extracted with methylene chloride (3x50 mL). The combined organic layers were dried over $MgSO_4$, concentrated and the crude product was purified on silica gel chromatography (cyclohexane / ethyl acetate : 95/5 to 80/20) to give **14** (2.78 g, α/β = 30/70, 73% combined yield). **14-β** : R*f* = 0.55 (cyclohexane/ethyl acetate : 50/50). ^1H NMR ($CDCl_3$, 200 MHz) δ = 2.09 (s, 3H, CH_3), 2.10 (s, 3H, CH_3), 2.17 (ddd, 1H, *J* = 13.8, 10.7 and 6.0 Hz, $H_{2'}$), 2.36 (dd, 1H, *J* = 13.8, and 5.8 Hz, $H_{2'}$), 4.18 (m, 2H, $H_{4'}$ and $H_{5'}$), 4.32 (dd, 1H, *J* = 13.2 and 5.2 Hz, $H_{5'}$), 5.22 (d, *J* = 6.0 Hz, 1H, $H_{3'}$), 5.26 (dd, 1H, *J* = 10.7 and 5.2 Hz, $H_{1'}$), 6.75 (dd, 1H, *J* = 3.7 and 0.6 Hz, H-thiophene), 6.89 (d, 1H, *J* = 3.7 Hz, H-thiophene). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 21.0, 21.2, 41.5, 64.3, 76.5, 76.7, 82.8, 112.3, 125.3, 129.5, 145.7, 170.6, 170.8. MS (ESI, MeOH) m/z: 384.9-386.9 $[M+Na]^+$, 400.8-402.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{13}H_{14}BrO_5S$ $[M-H]^-$, 360.9750; found, 360.9740.



diol intermediate

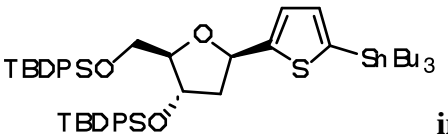
To a solution of **14-β** (835 mg, 2.3 mmol) in MeOH (12 mL) was added K_2CO_3 (3 equiv, 953 mg) and the mixture was stirred at room temperature for 1h. The solution was evaporated in vacuo and the residue was purified by flash chromatography (methylene chloride / methanol : 90/10) to give the diol as a foam (630 mg, 98%). The spectral data of this product are in accordance with those recently reported by M. Hocek and *Coll*. (ref. 2) (see ^1H and ^{13}C spectral data). HRMS (ESI) calcd for $C_9H_{10}BrO_3S$ $[M-H]^-$, 276.9539; found, 276.9528.



15

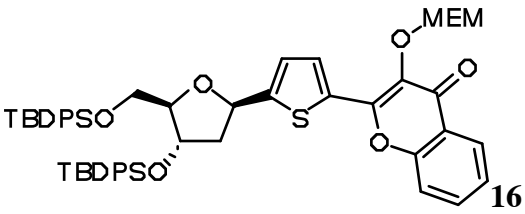
To a solution of diol previously obtained (610 mg, 2.19 mmol) in dry DMF (11 mL) were successively added imidazole (3.5 equiv) and TBDPSCl (3.5 equiv, 1.97 mL) under N_2 atmosphere. After stirring for 24h, the reaction mixture was quenched with a saturated solution of NH_4Cl and then extracted three times with methylene chloride. The combined organic layers were dried over $MgSO_4$, filtered and evaporated. The residue was purified by silica gel column (cyclohexane/ethyl acetate : 90/10) to afford **15** as a colorless oil (1.57g, 95%).

R*f* = 0.8 (cyclohexane/ethyl acetate : 80/20). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 0.85 (s, 9H, *t*-Bu), 1.01 (s, 9H, *t*-Bu), 1.81 (ddd, 1H, *J* = 12.6, 10.9 and 5.2 Hz, $H_{2'}$), 1.81 (dd, 1H, *J* = 11.4 and 5.1 Hz, $H_{2'}$), 3.24 (dd, 1H, *J* = 11.0 and 4.0 Hz, $H_{5'}$), 3.40 (dd, 1H, *J* = 11.0 and 4.0 Hz, $H_{5'}$), 4.00 (dt, 1H, *J* = 3.7 and 1.1 Hz, $H_{4'}$), 4.32 (br d, 1H, *J* = 5.0 Hz, $H_{3'}$), 4.96 (dd, 1H, *J* = 10.9 and 5.1 Hz, 1H, $H_{1'}$), 6.64 (dd, 1H, *J* = 3.7 and 0.6 Hz, H-thiophene), 6.77 (d, 1H, *J* = 3.7, H-thiophene), 7.115-7.65 (m, 20H, Ar). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 19.2, 26.9, 27.1, 44.7, 64.3, 75.6, 88.5, 127.7, 127.7, 127.8, 127.9, 129.3, 129.8, 130.0, 133.2, 133.7, 133.8, 135.9, 147.2. MS (ESI, MeOH) m/z: 777.4 - 779.4 $[M+Na]^+$. HRMS (ESI) calcd for $C_{41}H_{46}BrO_3SSi_2$ $[M-H]^-$, 753.1895; found, 753.1884.



intermediate

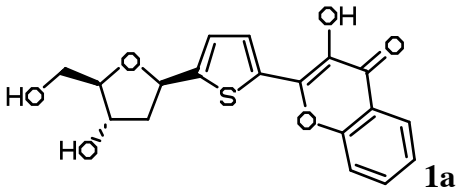
To a solution of **15** (2.08 mmol) in dry ether (8 mL) was added dropwise n-BuLi (1.1 equiv, 1.6M solution) at 0°C, under N_2 atmosphere. After 1h, the mixture was cooled to -78°C and tributyltin chloride (1.1 equiv) was slowly added and the mixture was stirred overnight. The reaction mixture was quenched with saturated NH_4Cl solution and extracted with ether. The combined organic layers were dried over $MgSO_4$, filtered and concentrated in vacuo. The obtained crude product was used in the next step without further purification.



16

To a stirred solution of 2-bromo-3-HC **9** (284 mg, 0.86 mmol) and tin derivative (2 mmol) in toluene (10 mL) under N_2 atmosphere was successively added $Pd(PPh_3)_4$ (50 mg) and CuI (16 mg). The mixture was stirred at 120°C for 2h, filtered through Celite and concentrated in vacuo. The crude residue was then purified by flash chromatography on silica gel (eluting with cyclohexane/ ethyl acetate : 90/10) to afford **16** as a yellow oil (573 mg, 72%).

R*f* = 0.30 (cyclohexane/ethyl acetate : 70/30). ^{1}H NMR (acetone-d_6, 200 MHz) δ = 0.87 (s, 9H, *t*-Bu), 1.02 (s, 9H, *t*-Bu), 1.97 (m, 1H, $2H_{2'}$), 2.29 (dd, 1H, *J* = 12.6 and 5.0 Hz, $H_{2'}$), 3.02 (s, 3H, CH_3), 3.19 (t, 2H, *J* = 4.6 Hz, CH_2 MEM), 3.30 (dd, 1H, *J* = 11.0 and 3.7 Hz, $H_{5'}$), 3.47 (dd, 1H, *J* = 11.1 and 3.8 Hz, $H_{5'}$), 3.62 (t, 2H, *J* = 4.6 Hz, CH_2 MEM), 4.05 (dt, 1H, *J* = 3.5 and 1.3 Hz, $H_{4'}$), 4.58 (d, 1H, *J* = 5.1 Hz, $H_{3'}$), 5.37 (s, 2H, OCH_2O), 5.47 (dd, 1H, *J* = 10.4 and 5.0 Hz, $H_{1'}$), 7.08 (d, 1H, *J* = 3.9, H-thiophene), 7.15-7.70 (m, 23H, Ar), 7.85 (d, 1H, *J* = 3.9, H-thiophene), 8.02 (dd, 1H, *J* = 7.8 and 1.4 Hz, H-chromone). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 20.0, 20.1, 27.6, 27.8, 46.2, 59.2, 65.5, 71.0, 72.7, 76.9, 77.7, 89.8, 97.0, 119.1, 125.3, 126.0, 126.0, 126.5, 129.0, 129.1, 129.2, 131.0, 131.1, 131.3, 131.5, 132.0, 134.2, 134.3, 134.8, 134.9, 136.4, 136.7, 136.8, 137.0, 151.7, 154.8, 158.1, 174.0. MS (ESI, MeOH) m/z: 947.4 $[M+Na]^+$.



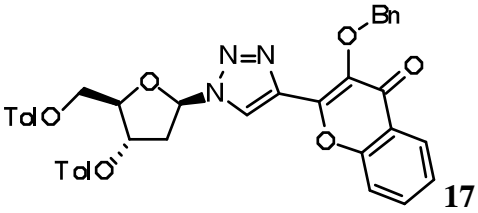
1a

To a solution of MEM-protected nucleoside **16** previously obtained (140 mg, 0.15 mmol) in dioxane (3 mL) was added HCl 6N (1.5 mL). The mixture was stirred at room temperature overnight then neutralized by saturated aqueous $NaHCO_3$ solution and extracted with ethyl acetate (3 x 10 mL). The organic layers were evaporated in vacuo and the residue was purified by silica gel column chromatography (methylene chloride/ methanol : 95/5 to 90/10) to give the free 3-HC-nucleoside **1a** as a pale resin (46 mg, 85%).

R*f* = 0.26 (9:1 CH_2Cl_2/MeOH). ^{1}H NMR (DMSO-d_6, 500 MHz) δ = 1.85 (ddd, 1H, *J* = 2.1, 11.3 and 13.7 Hz, $H_{2'}$), 2.12 (ddd, 1H, *J* = 2.1 and 3.8 and 13.7 Hz, $H_{2'}$), 3.30-3.62 (m, 1H, $H_{5'}$), 3.65-3.67 (m, 2H, $H_{5'}$and $H_{4'}$), 3.99 (br m, 1H, $H_{3'}$), 4.97 (dd, 1H, *J* = 2.1 Hz and 11.3

Hz, $H_{1'}$), 7.15 (d, 1H, *J* = 3.9 Hz, H-thiophene), 7.47 (t, 1H, *J* = 7.4 Hz, H-6), 7.71 (d, 1H, *J* = 8.3 Hz, H-8), 7.81 (m, 1H, H-7), 7.82 (d, 1H, *J*= 3.9 Hz, H-thiophene), 8.11 (dd, 1H, *J* = 1.2 and 7.4 Hz, H-5). ^{13}C NMR (DMSO-d_6, 125 MHz) δ = 42.4, 67.8, 68.2, 68.7, 71.0, 120.0, 123.7, 126.0, 126.5, 126.7, 129.9, 132.8, 135.5, 138.5, 145.2, 152.8, 156.1, 173.9. MS (ESI, MeOH) m/z: 358.9 $[M-H]^-$. HRMS (ESI) calcd for $C_{18}H_{17}O_6S$ $[M+H]^+$, 361.0745; found, 361.0741. IR (KBr) ν: 3432, 1651 cm^{-1}.

Synthesis of the triazolyl-3-Hydroxychromone 1b

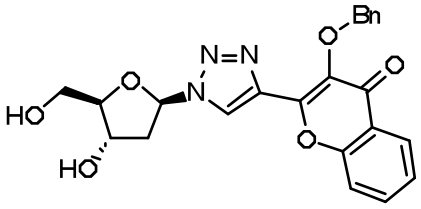


17

Method A : To a stirred solution of azido-sugar **4** (1 mmol), alkyne **10** (1.1 equiv) and nBu_4NF (1.1 equiv.) in methylene chloride (5 mL) was successively added CuI (164 mg, 2 equiv.), DIEA (0.37 mL, 5 equiv) and acetic acid (1 equiv). The mixture was stirred 4h at room temperature, filtered and the solvent removed in vacuo. The crude product was purified by flash chromatography (cyclohexane / ethyl acetate : 80/20) to give **17** (598 mg, 89%)

Method B : A mixture of azido-sugar **4** (1 mmol), alkyne **10** (1.1 equiv), nBu_4NF (1.1 equiv.), CuI (164 mg, 2 equiv.), DIEA (0.37 mL, 5 equiv.) was adsorbed on silica gel (1 g). The resulting yellow powder was placed into a microwave and irradiated for 2 min. The mixture was eluted twice with ethyl acetate and the solvent evaporated under reduced pressure to give a crude product which was subjected to a simple filtration over silica gel (cyclohexane / ethyl acetate: 80/20) to give **17** (638 mg, 95%).

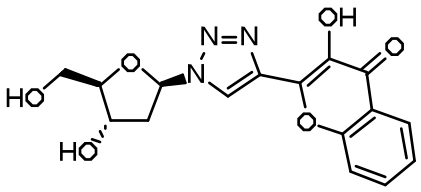
Mp =183-185°C (methylene chloride/ether). R*f* = 0.56 (cyclohexane / ethyl acetate: 50/50). ^{1}H NMR ($CDCl_3$, 200 MHz) δ = 2.25 (s, 3H, CH_3), 2.44 (s, 3H, CH_3), 2.82-2.92 (m, 1H, $H_{2'}$), 3.11-3.24 (m, 1H, $H_{2'}$), 4.50 (d, 2H, *J* = 4.4 Hz, $2H_{5'}$), 4.63-4.67 (m, 1H, $H_{4'}$), 5.21 and 5.32 (2d, 2H, *J* = 11.0 Hz, CH_2), 5.69-5.74 (m, 1H, $H_{3'}$), 6.42 (t, 1H, *J* = 5.7 Hz, $H_{1'}$), 7.13 (d, 2H, *J* = 8.0 Hz), 7.30-7.47 (m, 8H), 7.75 (m, 2H), 7.90 (d, 2H, *J* = 8.2 Hz), 7.95 (d, 2H, *J* = 8.2 Hz), 8.30 (d, 2H, *J* = 8.0 Hz). ^{13}C NMR ($CDCl_3$, 50 MHz) δ = 21.7, 21.9, 38.4, 63.7, 74.4, 74.6, 83.9, 89.2, 118.7, 124.5, 125.0, 125.8, 126.5, 128.7, 129.0, 129.3, 129.4, 129.8, 129.9, 133.8, 136.8, 139.2, 144.2, 144.7, 155.3, 165.9, 166.2, 174.4. MS (ESI, MeOH) m/z = 709.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{39}H_{33}N_3O_8$ $[M+H]^+$, 672.2345; found, 672.2341.



intermediate

By the same procedure as described above, compound **4b** (100 mg, 0.15 mmol) was deprotected using K_2CO_3 in MeOH to afford the pure product (62 mg, 95%)

Mp = 94-96°C (methylene chloride/ether). R*f* = 0.24 (9:1 CH_2Cl_2/MeOH). ^{1}H NMR (MeOD, 200 MHz) δ = 2.47-2.78 (m, 2H, $2H_{2'}$), 3.60 (dd, 1H, *J* = 4.6 and 11.9 Hz, $1H_{5'}$), 3.70 (dd, 1H, *J* = 3.8 and 11.9 Hz, $1H_{5'}$), 4.03 (dd, 1H, *J* = 4.2 and 8.2 Hz, $H_{4'}$), 4.48-4.56 (dd, 1H, *J* = 5.5 and 10.0, $H_{3'}$), 5.28 (s, 2H, CH_2), 6.42 (t, 1H, *J* = 5.9 Hz, $H_{1'}$), 7.20-7.35 (m, 4H), 7.44 (td, 1H, *J* = 1.1 and 8.1 Hz), 7.64 (d, 1H, *J* = 8.4 Hz), 7.75 (ddd, 1H, *J* = 1.5, 7.0 and 8.4 Hz), 7.85 (d, 1H, *J* = 8.2 Hz), 8.15 (dd, 1H, *J* = 1.3 and 8.1 Hz), 8.71 (s, 1H, H-5). ^{13}C NMR (MeOD, 50 MHz) δ = 42.0, 56.0, 63.0, 72.1, 75.0, 89.9, 90.7, 119.5, 125.1, 126.2, 126.4, 127.3, 129.4, 129.7, 130.0, 130.5, 130.7, 135.4, 137.6, 139.2, 139.5, 156.5, 176.1. MS (ESI, MeOH) m/z = 457.9 $[M+Na]^+$, 473.8 $[M+K]^+$. HRMS (ESI) calcd for $C_{23}H_{22}N_3O_6$ $[M+H]^+$, 436.1508; found, 436.1509.



1b

To a degassed solution of triazolyl-compound **5b** (0.3 mmol) in THF (10 mL) was added Pd/C (10% molar). Hydrogenolysis was then conducted in a high pressure autoclave (Parr apparatus, 3 bar) for 10h. The catalyst was filtered and washed with THF. The filtrate was concentrated in vacuo and the obtained residue was purified by flash chromatography to afford **1b** as a white powder (34 mg, 90 %).

R*f* = 0.32 (8:2 CH_2Cl_2/MeOH). ^{1}H NMR (DMSO-d_6, 500 MHz) δ = 2.42-2.47 (m, 1H, $H_{2'}$), 2.73 (ddd, 1H, *J*= 5.8 and 13.5Hz, $H_{2'}$), 3.45 (dd, 1H, *J* = 4.8 and 11.7 Hz, $H_{5'}$), 3.55 (dd, 1H, *J* = 4.3 and 11.7 Hz, $H_{5'}$), 3.91 (dd, 1H, *J* = 4.8 and 8.8, $H_{4'}$), 4.43-4.44 (m, 1H, $H_{3'}$), 4.87 (br s, 1H, $OH_{5'}$), 5.36 (br s, 1H, $OH_{3'}$), 6.52 (t, 1H, *J* = 5.9 Hz, $H_{1'}$), 7.48 (ddd, 1H, *J* = 1.0, 7.0 and 8.0 Hz), 7.73 (d, 1H, *J* = 8.2 Hz), 7.81 (ddd, 1H, *J* = 1.6, 7.0 and 8.6 Hz), 8.13 (dd, 1H, *J* = 1.6 and 8.0 Hz), 8.81 (s, 1H, H-5), 10.00 (s, 1H, OH). ^{13}C NMR (DMSO-d_6, 125 MHz) δ = 40.0, 61.5, 70.4, 88.4, 88.5, 118.4, 122.0, 124.7, 124.9, 125.3, 133.8, 137.9, 138.0, 140.3,

154.5, 172.1. MS (ESI, MeOH) m/z: 343.7 $[M-H]^-$, 368.0 $[M+Na]^+$. HRMS (ESI) calcd for $C_{16}H_{16}N_3O_6$ $[M+H]^+$, 346.1039; found, 346.1042. IR (KBr) ν: 3392, 1621 cm^{-1}.

References

1. Pleier, A-K.; Glas, H.; Grosche, M.; Sirsch, P.; Thiel, W.R. *Synthesis*. **2001**, *1*, 55
2. Bárta, J.; Pohl, R.; Klepetárová, Ernsting, N. P.; Hocek, M. *J. Org. Chem*. **2008**, *73*, 3798.

Based on our previous results on the synthesis of fluorescent nucleoside analogues and their photophysical characteristics (Spadafora, Postupalenko et al. 2009), the 3HC-thienylchromone derivative which displays the most promising photophysical properties in protic solvents was selected for incorporation into ODNs, since the surrounding ODN environment is relatively polar (**Publication 5**). Our strategy for DNA labeling was based on the design of 3HC connected to deoxyribose via an aryl moiety for incorporation into ODNs. 3HC fluorophores bearing a small heterocycle in 2-position are attractive analogues of nucleic bases because they are flat and their size corresponds well to the size of the two complementary A-T and G-C base pairs.

First, we synthesized the 3HC-thienyldeoxyribose (Figure 2.9) adapting the approach previously described for the free thienylchromone (Spadafora, Postupalenko et al. 2009). This was efficiently performed using phosphoramidite chemistry and solid phase. As a next step, 15- to 16-mer ODNs containing the thienylchromone at the positions labeled **M** were synthesized on solid phase. These 15- and 16-mer sequences, d(CGT TTT X**M**X TTT TGC) and 5'-d(**M** CGT TTT TAT TTT TGC) were chosen as model sequences for the photophysical characterization of the nucleoside analogues (Figure 2.9). The purity and mass of the labeled ODNs as well as the integrity of the chromone were confirmed by RP HPLC, Mass and UV/Vis analysis. Next, UV melting curves showed that the incorporation of the nucleoside analogues marginally affects the stability of the duplex and that the 3HC derivative can substitute any natural nucleobase. Furthermore, CD studies show that the 3HC base does not affect the duplex conformation. These results indicate that the 3HC-thienyldeoxyribose possesses the remarkable characteristics of a universal base.

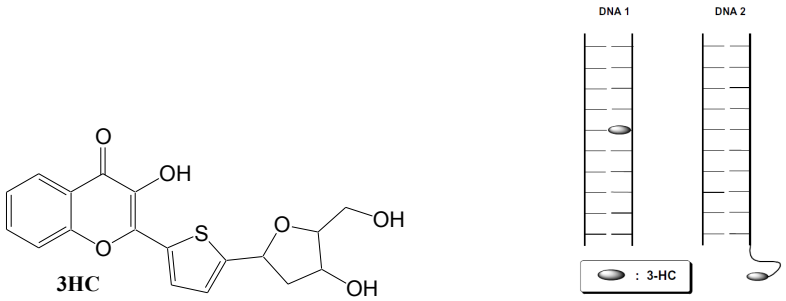


Figure 2.9. Structure of the 3HC-thienyldeoxyribose (3HC). Scheme of the two different labeled positions.

Furthermore, we characterized the UV and fluorescence spectroscopic properties of labeled single-stranded (ss) ODNs and double-stranded (ds) ODNs. In both single- and double-stranded states, labeled ODNs showed dual emission, characteristic of 3HC dyes, where the short- and long-wavelength bands can be assigned to N* and T* forms, respectively. Comparison of the emission spectra of the free nucleoside and the labeled ODNs in buffer shows a large decrease of the N*/T* intensity ratio of the labeled ODNs in respect to the free dye in buffer (Figure 2.10), which can be explained by the decrease of the polarity and H-bonding donor ability of the probe microenvironment in the ODNs. By comparison to the free chromone, increased quantum yields, up to 8 fold, were measured for the labeled ODNs. The quantum yield increase is more pronounced in double strands than in single strands. This remarkable result distinguishes the 3HC nucleoside from most of the available fluorescent base analogues that are severely quenched when incorporated in ODNs.

Moreover, the photostability of the 3HC dye under continuous irradiation was examined by measuring the fluorescence intensity of the samples as a function of time. The photostability was found to be dramatically improved when incorporated in ODNs in comparison to the free dye and Prodan, taken as a reference.

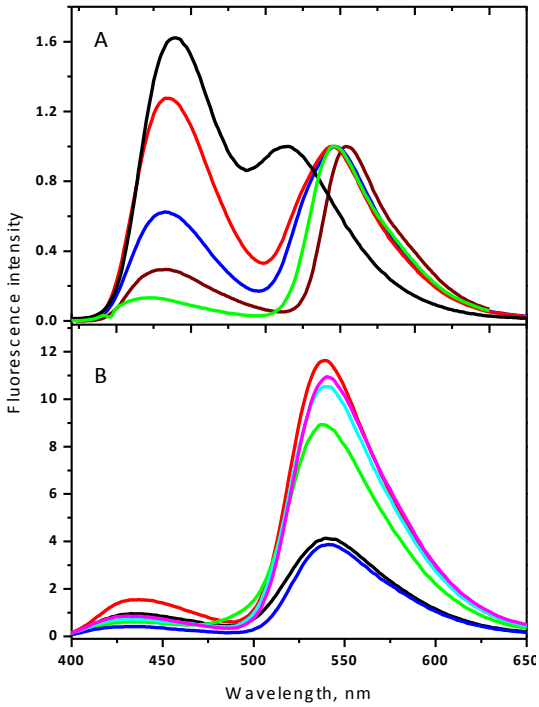


Figure 2.10. Fluorescence spectra: (A) free 3HC-thienylchromone in 10 mM phosphate buffer pH 6.5 (black), methanol (red), ethanol (blue), DMSO (brown), CH_3CN (green) (Spadafora, Postupalenko et al. 2009). The spectra were normalized at the T* band. (B) Single- and double-stranded labeled ODN in 10 mM cacodylate, 150 mM NaCl, buffer pH 6.5 (ODN sequences: d(CGT TTT T**M**T TTT TGC) (1) and d(GCA AAA A**Y**A AAA ACG) (2), **Y**= T, A, G, C, abasic site). Black: ssODN (1); red: dsODN (Y=T); green: dsODN (Y=A); deep blue: dsODN (Y=G); light blue: dsODN (Y=C); purple: dsODN (Y= abasic site).

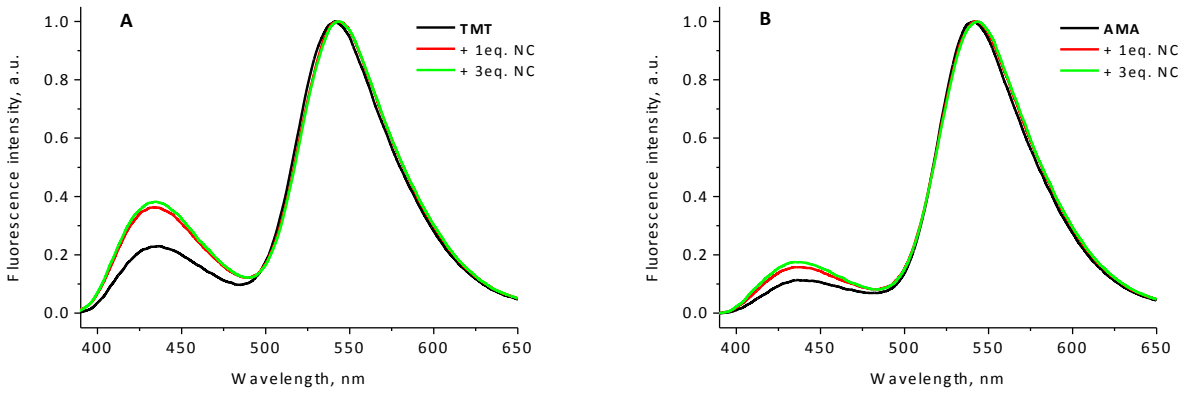
To explore the applicability of the 3HC-thienylchromone to detect, quantify and characterize peptide-ODN interactions, we investigated the interaction of labeled 15- and 16-mer ODNs (TMT, AMA, GMG, CMC and MCG) with the nucleocapsid protein (NC) from HIV-1. NC binds to a large range of nucleic acid sequences through sequence-specific (Clever, Sassetti et al. 1995; de Guzman, Wu et al. 1998; Vuilleumier, Bombarda et al. 1999; Amarasinghe, De Guzman et al. 2000; Shubsda, Paoletti et al. 2002; Bourbigot, Ramalanjaona et al. 2008) and non-specific (Mely, de Rocquigny et al. 1995; Fisher, Rein et al. 1998; Vuilleumier, Bombarda et al. 1999; Beltz, Clauss et al. 2005) interactions involving electrostatic and hydrophobic contributions. Specific interaction of NC with ODNs is mainly mediated through the folded finger motifs, with a key role being played by the hydrophobic platform at the top of the fingers (de Guzman, Wu et al. 1998; Bourbigot, Ramalanjaona et al. 2008).

Addition of NC to the labeled TMT ODN was found to significantly change the ratio of the two emission bands, indicating that the probe could sense the interaction (Figure 2.11A). The

increase in the N*/T* band ratio from 0.23 (ssODN alone) to 0.38 (ssODN/NC complex) was accompanied by an increase of the fluorescence quantum yield (Table 1) suggesting that the probe is shifted to a less hydrophobic environment probably as a consequence of an interaction with NC amino acids. In contrast, for AMA, we observed only a slight increase in N*/T* band ratio and quantum yield (Figure 2.11B, Table 1). The difference in response between AMA and TMT sequences, suggests that in combination with T neighbors, the 3HC label constitutes a stronger binding site than with A neighbors. Upon addition of NC to the GMG labeled ODN, an increase in the quantum yield was observed (Table 1), probably, due to the stacking of the NC aromatic amino acids with the G bases which prevents their stacking with the chromone. Noticeably, the N*/T* ratio of GMG in buffer was much higher than for other sequences (N*/T* = 0.66, Figure 2.11C), suggesting that the label was more exposed to water in this case. After binding of GMG to NC, the N*/T* ratio was decreased, suggesting that the probe was shifted to a less hydrophilic environment due to the screening of the label from the bulk water by the bound peptide. CMC labeled ODN also increased its N*/T* band ratio and quantum yield upon binding to NC peptide, similarly to TMT one (Figure 2.11D, Table 1). In contrast, a non-folded mutant of NC (NC-SSHS) that binds ssODNs only through electrostatic interactions with the ODN phosphate groups, was unable to induce any change in the emission spectra of the labeled ODN sequences (data not shown).

With MCG, we observed as with TMT that binding to NC gives a saturating signal already at 1eq. of protein (based on quantum yield, Table 1). Since it is not possible to have the strongest binding site both in the middle and at the end of the ODN sequence, suggesting that 3HC may introduce artificially a strong binding site. The N*/T* band ratio was not affected upon binding of the peptide (Figure 2.11E). No change was observed in the absorption spectra of labeled ODNs after binding to peptide (Table 1). Moreover, NC appeared to not bind to double stranded sequences since both the N*/T* ratio and the quantum yield were not unaffected (data not shown).

These data suggest that the probe allows site-specific monitoring of NC/ODN interactions, though it may increase the affinity of the site in which it is included.



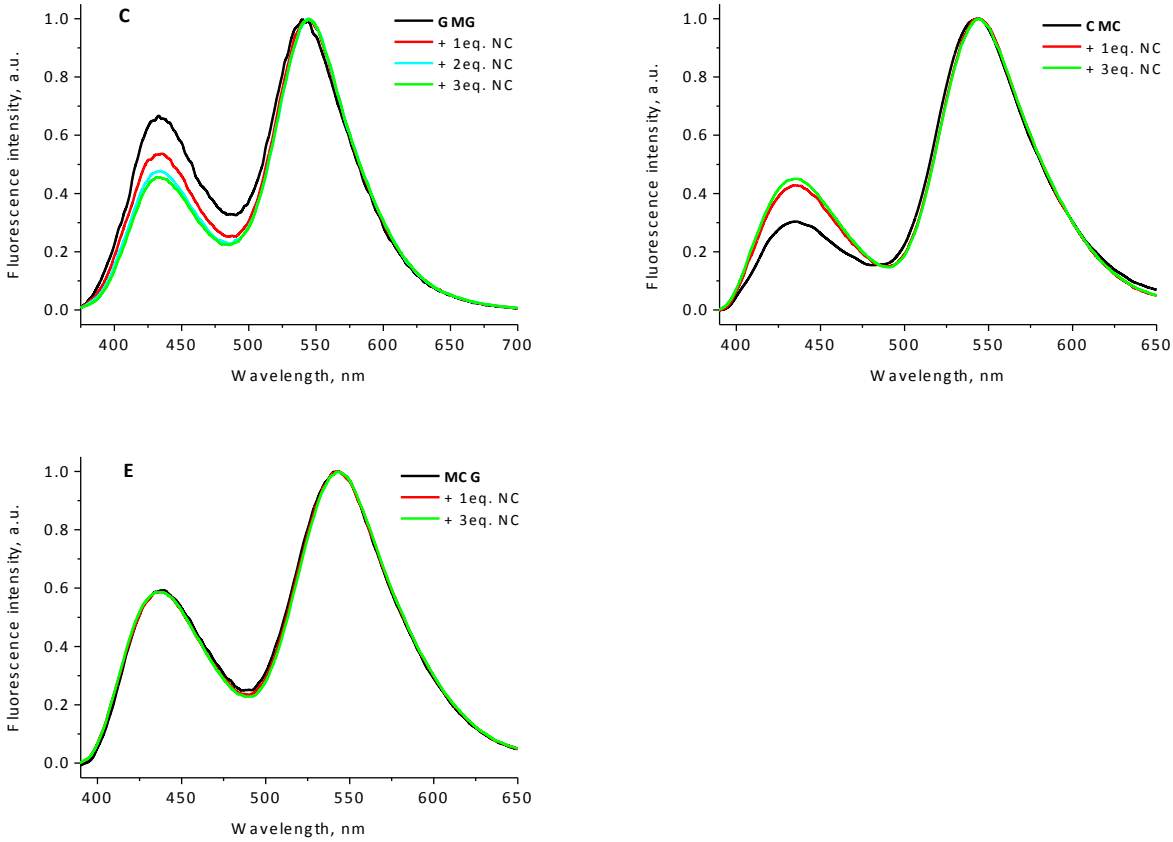


Figure 2.11. Fluorescence spectra of labeled oligonucleotides TMT (A), AMA (B), GMG (C), CMC (D) and MCG (E) in the absence and presence of the 1-3eq. NC peptide. The spectra were normalized at the maximum of the long wavelength T* band. The experiments were performed in 10mM cacodylate buffer containing 150 mM NaCl, pH 7.0 at 0.5 μM concentration of oligonucleotides. The concentration for GMG was 2 μM and experiments were performed in a quartz micro-cuvette (50 μL volume) due to limited amount of this ODN. The excitation wavelength was 360 nm.

Table 1. Spectroscopic characteristics of the labeled ODNs and their complexes with NC

ODN	Peptide	λ_{Abs}	λ_{N^*}	λ_{T^*}	N*/T*	QY, %
TMT		373	435	540	0.23	14
	+ 1eq. NC	373	434	542	0.36	18
	+ 3eq. NC	374	433	542	0.39	22
AMA		376	440	541	0.11	17
	+ 1eq. NC	376	438	541	0.16	18
	+ 3eq. NC	376	438	541	0.18	20
GMG		374	433	541	0.66	2
	+ 1eq. NC	374	433	543	0.54	4
	+ 2eq. NC	373	433	544	0.48	5
	+ 3eq. NC	373	433	544	0.45	6
CMC		374	434	543	0.29	4
	+ 1eq. NC	374	435	544	0.43	14
	+ 3eq. NC	373	435	544	0.45	18
MCG		374	438	542	0.59	8
	+ 1eq. NC	373	437	543	0.59	13
	+ 3eq. NC	373	437	543	0.59	13

λ_{Abs} – position of absorption maximum, λ_{N^*} and λ_{T^*} – positions of fluorescence maxima of N* and T* forms, respectively. I_{N^*}/I_{T^*} – ratio of the intensities of the two emission bands at their peak maxima. QY – fluorescence quantum yield. Concentration of ODNs was 0.5 μM. The concentration for GMG was 2 μM and experiments were performed in a quartz micro-cuvette (50 μL volume) due to limited amount of this ODN. 10 mM cacodylate buffer, 150 mM NaCl, pH 7.0 was used. Excitation wavelength was 360 nm. As a reference for calculation of quantum yields, quinine sulfate in H_2SO_4 (QY = 57.7%) was used.

To further investigate peptide-oligonucleotide interactions by using the 3HC-thienylchromone nucleoside analogue, the ΔP(-)PBS sequence was chosen for binding studies with the NC protein. ΔP(-)PBS is the cDNA copy of the primer binding site deleted of its 5' overhang. ΔP(-)PBS was shown to bind NC with high affinity and 1:1 stoichiometry (Bourbigot, Ramalanjaona et al. 2008). NMR studies show that NC preferentially interacts at the 5′ end of the loop (5-CTG-7) of ΔP(−)PBS (Bourbigot, Ramalanjaona et al. 2008). This interaction is mediated through the hydrophobic plateau (Val13, Phe16, Thr24, Ala25, Trp37, and Met46) on the zinc finger domain (Mély, Jullian et al. 1994; Morellet, de Rocquigny et al. 1994; Lee, Guzman et al. 1998). The Phe16 and Trp37 residues insert between T6 and G7 bases, allowing a strong stacking of the Trp37 aromatic ring with the G7 base. The interaction with the aromatic residues likely

constitutes the determinant factor of the interaction (Mori, Dietrich et al. 2010) and directs nucleotides T6 and G7 toward the exterior of the loop.

Several positions in the ΔP(-)PBS sequence were selected for the substitution by the 3HC-thienylchromone nucleoside analogue: thymine at position 2 (2-TMC), 8 (8-GMT) and 9 (9-GMC). The labeled nucleotide was also added as an additional nucleotide at the 5' end (0-MGT). The labeled ODNs were synthesized on solid state. Their purity and mass, as well as the integrity of the chromone were confirmed by RP HPLC, mass and UV/Vis analysis.

Table 2. Sequences of synthesized labeled ΔP(-)PBS

0-MGT	5'-**M**GT-CCC-TGT-TCG-GGC-3'
2-GMC	5'- G**M**-CCC-TGT-TCG-GGC-3'
8-GMT	5'- GT-CCC-TG**M**-TCG-GGC-3'
9-TMC	5'- GT-CCC-TGT-**M**CG-GGC-3'

Labeled ΔP(-)PBS sequences showed a very low fluorescence for all labeled positions, probably, due to the presence of G neighbors in the sequence that quench the dye fluorescence (quantum yield is 0.5-1.8%). For all sequences, well-resolved dual emission was observed in the emission spectra, with N* blue-shifted band centered at 435-437 nm and T* red-shifted one centered at 532-548 nm (Figure 2.12, Table 3). There was also a strong contribution of an anionic form in the emission (Figure 2.12) and absorption (data not shown) spectra for 8-GMT which was also observed for 2-GMC but to a smaller extend (Figure 2.12). As expected, comparison of the N*/T* ratios indicates that the 3HC dye was more exposed to the bulk water at position 2 (2-GMC) and at the 5' end (0-MGT) as compared to the positions in the loop. (Figure 2.12, Table 3).

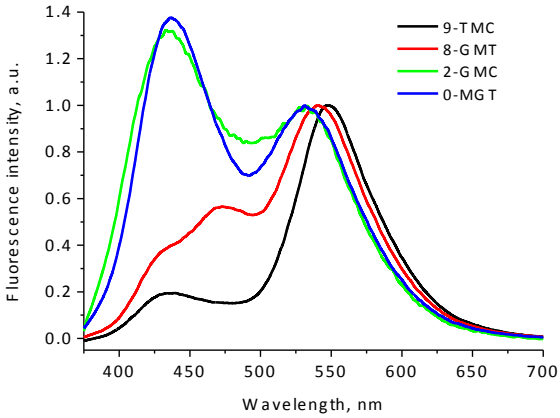
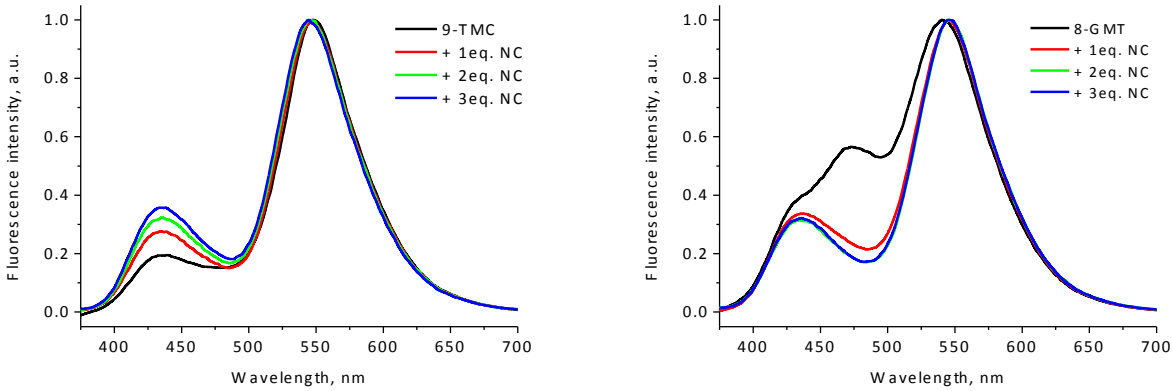


Figure 2.12. Fluorescence spectra of the labeled oligonucleotides 9-TMC, 8-GMT, 2-GMC and 0-MGT in buffer (10 mM cacodylate, 150 mM NaCl, pH 7.0). Concentration of labeled oligonucleotides was 5 μM. Excitation wavelength was 360 nm. Spectra were normalized at the T* band.

After addition of NC, a significant increase of the N*/T* ratio of the two emission bands (from 0.2 to 0.36) was observed in case of 9-TMC (Figure 2.13, Table 3). Monitoring the N*/T* intensity ratio as a function of increasing protein concentrations evidenced that the ratio changes reached a plateau value (Figure 2.13). Upon binding of NC with 9-TMC, an increase of the fluorescence quantum yield was also observed, suggesting that the probe is shifted to a more hydrophilic environment probably as a consequence of the insertion of the NC amino acids within the bases and the stretching of the loop, in full line with the NMR structure of the NC/PBS complex (Figure 2.14) (Bourbigot, Ramalanjaona et al. 2008). Binding of NC induced large changes in the 8-GMT ODN spectra, with the disappearance of the anionic form, resulting in N*/T* ratio similar to that of 9-TMC (Figure 2.13). For the 2-GMC ODN, NC shifted the 3HC label to a less hydrophilic environment, as shown by the decrease of the N*/T* ratio (Figure 2.13, Table 3). For 0-MGT, addition of NC provided more limited changes (Figure 2.13, Table 3). For all labeled ODNs, the quantum yield was slightly increased upon NC binding (Table 3). We observed changes of N*/T* band ratio in all labeled ΔP(-)PBS sequences upon addition of NC peptide, in line with the binding of at least two peptides to these ODNs. The strongest binding site was likely at position 8, since the spectroscopic changes reached a plateau at a 1:1 stoichiometry. This is in line with the strong binding site evidenced by NMR in this region (Bourbigot, Ramalanjaona et al. 2008). In further line with these NMR data, additional, but lower binding sites were suggested at the stem terminus and at the 3'end of the loop.

In sharp contrast, a non-folded mutant of NC that binds ΔP(−)PBS only through electrostatic interactions with the ΔP(−)PBS phosphate groups, was unable to induce any change in the emission spectra of the labeled ΔP(−)PBS sequences (data not shown).



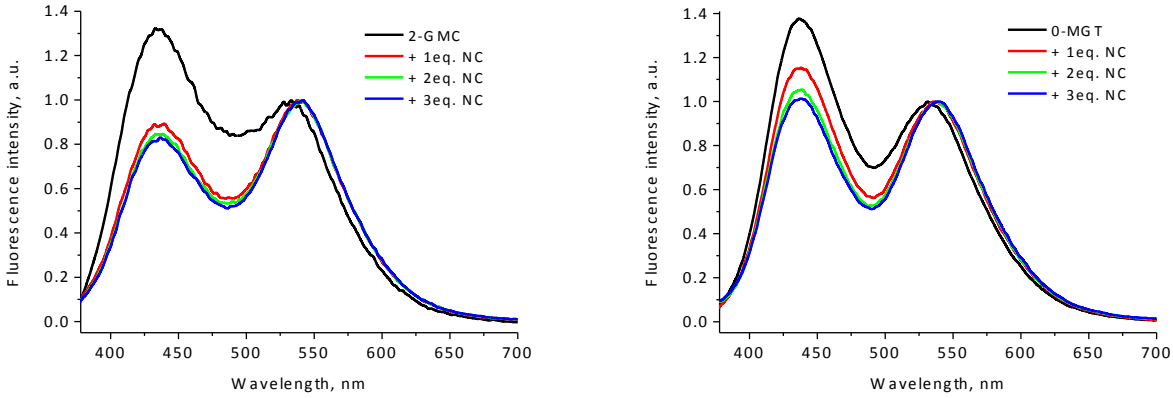


Figure 2.13. Normalized fluorescence spectra of labeled ΔP(−)PBS in the absence and presence of NC peptide (1-3 eq.). Experiments were performed in buffer (10 mM cacodylate, 150 mM NaCl, pH 7.0). The concentration of ODNs was 5 μM and experiments were performed in a quartz micro-cuvette (50 μL volume) due to small amount of labeled ΔP(−)PBS available.

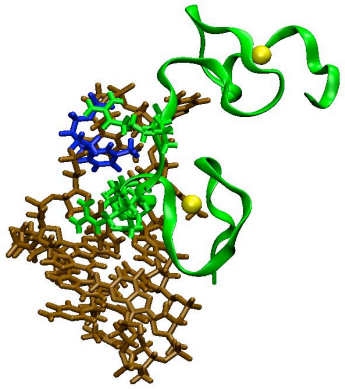


Figure 2.14. 3D structure of the ΔP(−)PBS/NC complex (Bourbigot, Ramalanjaona et al. 2008). The ΔP(−)PBS is in brown, while the protein is in green. Yellow spheres correspond to zinc atoms. The blue nucleoside base corresponds to the thymine at 9 position substituted by the 3HC-thienylchromone nucleoside in one of the derivatives.

Table 3. Spectroscopic characteristics of the labeled ΔP(−)PBS and their complexes with NC

ODN	Peptide	λ_{Abs}	λ_{N*}	λ_{T*}	I_{N*}/I_{T*}	QY, %
9-TMC		376	436	548	0.20	1.8
	+ 1eq. NC	375	435	546	0.28	2.2
	+ 2eq. NC	375	435	546	0.32	2.5
	+ 3eq. NC	375	435	545	0.36	2.8
8-GMT		375	a	541	a	1.6
	+ 1eq. NC	375	437	546	0.34	2.0
	+ 2eq. NC	373	436	546	0.32	2.6
	+ 3eq. NC	372	436	546	0.32	3.0
2-GMC		375	435	532	1.32	0.5
	+ 1eq. NC	374	436	540	0.89	0.7
	+ 2eq. NC	374	436	540	0.85	0.7
	+ 3eq. NC	374	436	540	0.83	0.9
0-MGT		371	437	532	1.38	1.1
	+ 1eq. NC	372	437	538	1.15	1.3
	+ 2eq. NC	372	437	538	1.05	1.5
	+ 3eq. NC	372	437	539	1.01	1.7

λ_{Abs} – position of absorption maxima, λ_{N*} and λ_{T*} - position of fluorescence maxima of N* and T* forms, respectively. I_{N*}/I_{T*} - ratio of the intensities of the two emission bands at their peak maxima. QY- fluorescence quantum yield. a – Position of N* band and N*/T* ratio could not be determined in this case. Concentration of the oligonucleotides was 5 μM. 10 mM cacodylate buffer, 150 mM NaCl, pH 7.0 was used. Excitation wavelength was 360 nm. As a reference for calculation of quantum yields, quinine sulfate in H_2SO_4 (QY = 57.7%) was used.

In this work, a new environmental sensitive fluorescent nucleoside analogue bearing a 3HC-thienylchromone as a base surrogate was introduced in various ODNs. We showed that this nucleoside analogue can be used as a universal base for ODN labeling. Moreover, labeling of short ODNs at selected positions provided site-specific multi-channel information on changes in the environment induced by the interaction with proteins of interest.

Publication 5

Synthesis and characterization of oligonucleotides incorporating a 3-hydroxychromone as a ratiometric fluorescent base surrogate

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

A Universal Nucleoside with Strong Two-Band Switchable Fluorescence and Sensitivity to the Environment for Investigating DNA Interactions

Dmytro Dziuba,[†] Viktoriia Y. Postupalenko,[‡] Marie Spadafora,[†] Andrey S. Klymchenko,[‡] Vincent Guérineau,[§] Yves Mély,[‡] Rachid Benhida,*[,†] and Alain Burger*[,†]

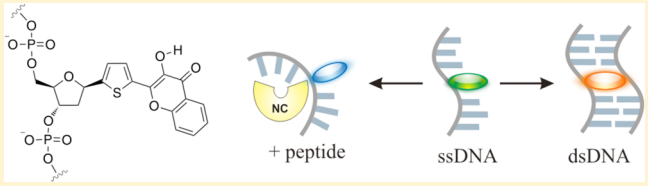
[†]Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice cedex 2, France

[‡]Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France

[§]Centre de recherche de Gif, Institut de Chimie des Substances Naturelles CNRS, Avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France

S *Supporting Information*

ABSTRACT: With the aim of developing a new tool to investigate DNA interactions, a nucleoside analogue incorporating a 3-hydroxychromone (3HC) fluorophore as a nucleobase mimic was synthesized and incorporated into oligonucleotide chains. In comparison with existing fluorescent nucleoside analogues, this dye features exceptional environmental sensitivity switching between two well-resolved fluorescence bands. In labeled DNA, this nucleoside analogue does not alter the duplex conformation and exhibits a high fluorescence quantum yield. This probe is up to 50-fold brighter than 2-aminopurine, the fluorescent nucleoside standard. Moreover, the dual emission is highly sensitive to the polarity of the environment; thus, a strong shielding effect of the flanking bases from water was observed. With this nucleoside, the effect of a viral chaperone protein on DNA base stacking was site-selectively monitored.



INTRODUCTION

Designing DNA-based fluorescent structures with site-specific responses to intermolecular interactions is a great challenge. The majority of environmentally sensitive fluorescent nucleosides are single-band emitters that respond to environmental changes by changes in the fluorescence intensity or a shift in the emission maximum.[1] These nucleosides suffer from limitations such as quenching by neighboring nucleobases or poor sensitivity to perturbations in the DNA structure. For example, 2-aminopurine (2-AP), the most popular nucleotide mimic,[1b] is used as an intensiometric probe but exhibits low sensitivity to the environment and low quantum yields in DNA duplexes.[2] These drawbacks have stimulated further research combining top achievements in photophysics, spectroscopy, and synthetic chemistry.[1b] It is particularly important that the spectrum of the artificial nucleobase be highly sensitive to intermolecular interactions. Ideally, a single dye should be employed to avoid double labeling that is needed for fluorescence resonance energy transfer (FRET), excimers, or J-aggregates.[3] In addition, the effect of the nucleotide mimic on the double-helical structure of native DNA must be minimal.

Therefore, we addressed the unique photophysical properties of 3-hydroxychromone (3HC) dyes, which exhibit polarity- and hydration-sensitive switching between well-resolved highly intense fluorescence bands in the visible range. As a result of excited-state intramolecular proton transfer (ESIPT), these fluorophores exhibit two excited states: the initially excited normal form (N*) and the tautomeric form (T*); each form generates one well-resolved emission band (Figure 1).[4]

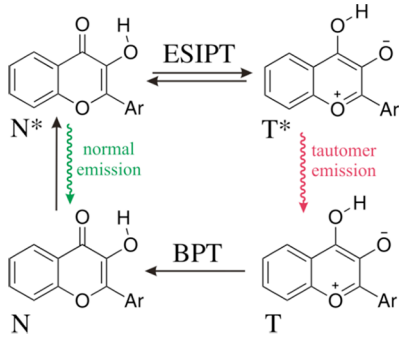


Figure 1. ESIPT reaction in 3-hydroxychromones. BPT denotes back proton transfer, and N* and T* represent the normal and tautomeric emissive forms, respectively.

The dual emission of 3HCs is highly sensitive to the environment because an increase in the donating ability of the hydrogen bond and the dielectric constant of the solvent inhibit

Received: March 30, 2012
Published: May 16, 2012

   dx.doi.org/10.1021/ja3030388 | *J. Am. Chem. Soc.* 2012, 134, 10209−10213

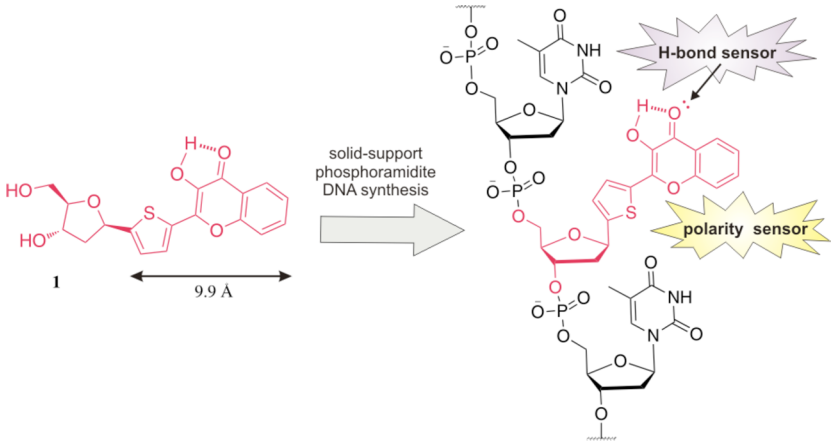


Figure 2. Hydrogen-bonding- and dielectric-constant-sensitive nucleoside **1** bearing 2-thienyl-3-hydroxychromone as a base surrogate and its incorporation into DNA.

Table 1. Thermal Denaturation of Double-Stranded ODN Samples

entry	dsDNA[a]	T_m (°C)	ΔT_m (°C)	entry	dsDNA[a]	T_m (°C)	ΔT_m (°C)	entry	dsDNA[a]	T_m (°C)	ΔT_m (°C)
1	**AAA** + ***TTT***	50.6	0	12	**TAT** + ***ATA***	50.8	0	22	**CAC** + ***GTG***	55.1	0
2	**AAA** + ***TAT***	39.2	−11.4	13	**TAT** + ***AAA***	42.1	−8.7	23	**CAC** + ***GAG***	45.8	−9.3
3	**AAA** + ***TGT***	40.7	−9.9	14	**TAT** + ***AGA***	42.8	−8	24	**CAC** + ***GGG***	50.3	−8.8
4	**AAA** + ***TCT***	39.8	−10.8	15	**TAT** + ***ACA***	40.9	−9.9	25	**CAC** + ***GCG***	47.2	−7.9
5	**AAA** + ***TAbT***	33.7	−16.9	16	**TAT** + ***AAbA***	33.1	−17.7	26	**CAC** + ***GAbG***	40.5	−14.6
6	**AMA** + **TTT**	41.4	−9.2	17	**TMT** + ***ATA***	45.8	−5	27	**CMC** + ***GTG***	48.9	−6.2
7	**AMA** + ***TAT***	41.2	−9.4	18	**TMT** + ***AAA***	46.3	−4.5	28	**CMC** + ***GAG***	49.9	−5.2
8	**AMA** + ***TGT***	41.4	−9.2	19	**TMT** + ***AGA***	45.3	−5.5	29	**CMC** + ***GGG***	50.0	−5.1
9	**AMA** + ***TCT***	43.1	−7.5	20	**TMT** + ***ACA***	46.5	−4.3	30	**CMC** + ***GCG***	51.5	−3.6
10	**AMA** + ***TAbT***	41.0	−9.6	21	**TMT** + ***AAbA***	46.6	−4.2	31	**CMC** + ***GAbG***	51.3	−3.8
11	**M-TAT** + ***ATA***	55.5	+4.7								

[a]In pH 7 buffer (10 mM cacodylate, 150 mM NaCl).

the ESIPT reaction and thus decrease the relative intensity of the T* band.[5] A critical advantage of ratiometric dyes over conventional single-band dyes is that this ratio depends only on the microenvironment of the fluorophore and not on the local concentration of the dye or the instrument settings. In addition to the ratio of the two emission bands, information could also be extracted from the positions of the absorption and two emission maxima.[6] 3HCs have been shown to be powerful fluorescent tools to probe lipid bilayers, cell membranes, proteins, and peptides,[7] but the synthesis of DNA labeled with ratiometric 3HC derivatives and their application for sensing have never been described. Our strategy relies on a thienyl-3HC-modified deoxyribose **1** that allows their incorporation into oligonucleotides (ODNs) (Figure 2).

We previously described the synthesis and photophysical properties of such nucleoside analogues.[8] Derivative **1** displays the most promising photophysical properties as a polarity- and hydration-sensitive label for incorporation into ODNs. The thienyl-3HC group of **1** is an attractive analogue of nucleic bases because it is a flat molecule and its size corresponds well to the size of the two complementary AT and GC base pairs. We preferred to graft the dye to deoxyribose at the thienyl substituent rather than to the benzene ring of the chromone. This grafting should favor the exposure of the 4-carbonyl moiety to local water in singly labeled ODNs (Figure 2). Annealing with complementary ODNs or formation of a protein−DNA complex should reduce the local hydration and polarity of the dye and thus change the ratio of the intensities of the two emission bands of **1**. Herein we report the synthesis of ODNs incorporating **1**, the photophysical properties of the labeled ODNs, and the ability of **1** to sense the DNA microenvironment and DNA−protein interactions site-selectively.

RESULTS AND DISCUSSION

The syntheses of the phosphoramidite of **1** and the labeled 15- and 16-mer ODNs is described in the Supporting Information. The 15-mer d(CGT TTT **XMX** TTT TGC) and 16-mer 5′-d(**M** CGT TTT **TAT** TTT TGC) sequences containing thienylchromone **1** at the positions labeled **M** were chosen as model sequences to characterize the photophysical properties of **1** in ODNs. In the 15-mers (**AMA**, **TMT**, **CMC** and **GMG**), **1** was incorporated in the middle of the sequence where the influence of the nucleoside analogue on the thermal stability should be the most pronounced. Moreover, the sequences differed by the residues flanking the modified nucleotide to compare the properties of label **1** in different ODN contexts. In the 16-mer (**M-TAT**), the ODN contained a dangling **1** at the 5′ end.[9] Control experiments were conducted with reference wild-type strands of composition d(CGT TTT **XAX** TTT TGC), where **X** = **A**, **C**, **G** or **T**. Each ODN was then annealed with its complementary ODN (noted in italics) containing a natural base or an abasic site located at the position opposite dye **1**, resulting in duplexes with different compositions (see Table S1 in the Supporting Information). The sequences of the complementary strands were d(GCA AAA ***YXY*** AAA ACG), in

dx.doi.org/10.1021/ja3030388 | J. Am. Chem. Soc. 2012, 134, 10209−10213

which ***Y*** is ***A***, ***C***, ***G***, or ***T*** and ***X*** is any of the natural bases or an abasic site (***Ab***).

As a first step to characterize the ODNs containing thienylchromone **1**, the thermodynamic stabilities of their duplexes were compared with the stabilities of the corresponding unlabeled duplexes by monitoring the temperature-induced absorbance changes at 260 nm (Table 1 and Figure S4 in the Supporting Information). Thermal denaturation studies showed that incorporation of nucleoside analogue **1** opposite a natural base or an abasic site led to duplexes with comparable thermal stabilities (entries 6−10, 17−21, and 27−31). Although less stable than the corresponding fully canonical duplexes, those containing **1** are more stable than the nonlabeled duplexes with mismatches (Table 1). For example, the duplexes containing the natural pair (**TAT** + ***ATA***), the mismatched pair (**TAT** + ***AAA***), and dye **1** (**TMT** + ***ATA***) melted with T_m values of 50.8, 42.1, and 45.8 °C, respectively (entries 12, 13, and 17). The data indicate that the dye stacks more strongly than the natural bases, likely as a result of a larger area of contact with the surrounding bases. A similar effect was observed in previous studies with labeled ODNs containing aromatic groups (e.g., pyrene, biphenyl, bipyridine, isoquinolone) instead of the natural base.[10] Thus, since thienyl-3HC **1** is nondiscriminating toward natural bases and less destabilizing than a mismatch, thienyl-3HC shows characteristics of a universal base.[11] Relative to the canonical duplexes, **TMT** + ***AXA*** and **CMC** + ***GXG*** gave smaller differences in T_m (ΔT_m) than the **AMA** + ***TXT*** duplexes (Table 1). The ΔT_m values can be grouped in two sets. One is in the range of 3.6−6.2 °C, and the other has values of 7.5−9.4 °C. The more favorable ΔT_m values for the **TMT** + ***AXA*** and **CMC** + ***GXG*** duplexes might be due to an increased overlap between the chromone moiety and the flanking purines of the complementary strand, as proposed for ODNs incorporating hydroxyphenylbenzoxazole.[12] Furthermore, in comparison to the natural base adenine, **1** enhanced the T_m of duplexes containing an abasic site by ∼7 °C for **AMA** + ***TAbT*** and 11−13 °C for **TMT** + ***AAbA*** and **CMC** + ***GAbG*** (Table 1). The stacking ability of dye **1** was further confirmed by the 4.7 °C increase in T_m observed for the **M-TAT** + ***ATA*** duplex containing the dangling thienyl-3HC (Table 1).[9]

In a second step, the influence of **1** on the secondary structure of the labeled duplexes was studied by circular dichroism (CD) spectroscopy (Figure S5 in the Supporting Information). The CD spectra were similar to those of the unlabeled duplexes, indicating that the dye does not affect the B-helix conformation of the duplexes. Altogether, the thermal denaturation and CD data support the idea that thienyl-3HC **1** can replace any natural nucleobase or base pair, giving stable duplexes with the B conformation.

In a third step, we characterized the UV absorbance and fluorescence properties of **1** in the 15- and 16-mer ODNs, either in their single-stranded (ss) or double-stranded (ds) states (Table 2 and Figure 3). In buffer, labeled ssODNs and dsODNs showed absorption bands centered at 373−376 and 372−381 nm, respectively. The absorption maxima of the labeled ssODNs and dsODNs were red-shifted by 5−14 nm relative to that of the free dye in buffer, suggesting stacking of the dye with its flanking bases (Table 2 and Figure S6 in the Supporting Information). A red shift of the absorption maximum is commonly observed with intercalating dyes (e.g., acridine and phenanthridinium).[13] In ssODNs, the effect of flanking bases on the 3HC absorption maximum was larger with purines than pyrimidines, suggesting a stronger stacking with purines because of their larger size.

Table 2. Spectroscopic Data for Nucleoside 1 and Labeled ODNs

entry	sample[a]	λ_{abs} (nm)[b]	λ_{N^*} (nm)[c]	λ_{T^*} (nm)[d]	I_{N^*}/I_{T^*}[e]	QY (%)[f]
1	**1**[g]	367	440	515	1.72	4.6
2	**AMA**	376	440	541	0.11	17
3	**AMA** + ***TTT***	376	441	540	0.17	24
4	**AMA** + ***TAT***	376	439	542	0.09	20
5	**AMA** + ***TGT***	375	439	542	0.10	16
6	**AMA** + ***TCT***	376	438	543	0.09	25
7	**AMA** + ***TAbT***	377	437	542	0.07	30
8	**TMT**	373	435	540	0.23	14
9	**TMT** + ***ATA***	373	437	540	0.14	38
10	**TMT** + ***AAA***	376	434	540	0.07	28
11	**TMT** + ***AGA***	373	433	541	0.09	13
12	**TMT** + ***ACA***	373	433	540	0.07	30
13	**TMT** + ***AAbA***	373	433	541	0.08	35
14	**CMC**	373	434	543	0.28	4
15	**CMC** + ***GTG***	375	434	540	0.41	1
16	**CMC** + ***GAG***	381	429	519	0.10	2
17	**CMC** + ***GGG***	373	433	540	0.43	1
18	**CMC** + ***GCG***	375	434	540	0.45	1
19	**CMC** + ***GAbG***	375	433	540	0.39	1
20	**GMG**	374	433	541	0.66	2
21	**M-TAT**	374	438	542	0.59	8
22	**M-TAT** + ***ATA***	372	437	532	1.05	1

[a]In pH 7 buffer (10 mM cacodylate, 150 mM NaCl). [b]λ_{abs} is the position of the absorption maximum. [c]λ_{N^*} is the position of the fluorescence maximum of the N* band. [d]λ_{T^*} is the position of the fluorescence maximum of the T* band. [e]I_{N^*}/I_{T^*} is the ratio of the intensities of the two emission bands at their maxima. [f]QY is the fluorescence quantum yield calculated using quinine sulfate (QY = 0.577 in 0.5 M H_2SO_4) as a reference. [g]Data taken from ref 8, where a pH 6.5 buffer (10 mM phosphate) was used.

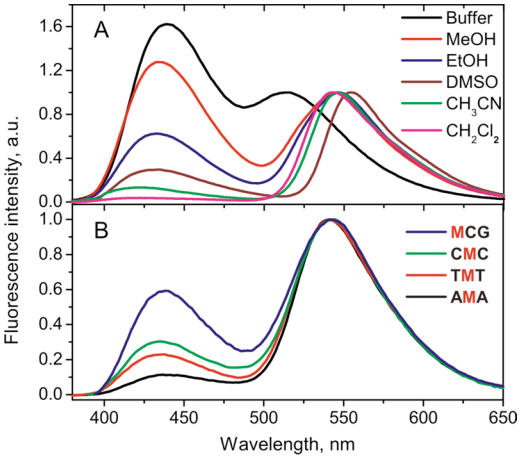


Figure 3. Normalized fluorescence spectra of (A) free thienyl-3HC **1** in different solvents[8] and (B) labeled ssODNs in pH 7 buffer (10 mM cacodylate, 150 mM NaCl).

The labeled ODNs showed well-resolved two-band emission, with the short- and long-wavelength maxima centered at 429−441 and 519−543 nm, respectively (Table 2). These values are very close to those obtained for free **1** in various organic solvents (Figure 3). Thus, the short- and long-wavelength

dx.doi.org/10.1021/ja3030388 | J. Am. Chem. Soc. 2012, 134, 10209−10213

bands were assigned to the emissions of the N* and T* forms, respectively, indicating that the 3HC dye still undergoes an ESIPT reaction in ODNs and therefore maintains its fundamental properties. The Stokes shifts for the first and second emission bands were ~3870 and ~8100 cm^{-1}, respectively. The large Stokes shift obtained for the second emission band is typical for ESIPT probes.

Interestingly, the fluorescence quantum yield (QY) of the dye in single- and double-stranded ODNs varied from 1 to 38% depending on the surrounding nucleobases (Table 2). In ODNs with A or T flanking the dye, the quantum yields of the ss and ds forms (entries 2−13) were 3−8-fold larger than that of the free dye **1** in buffer.

Moreover, with the exception of G opposite **1** (Table 2, entries 5 and 11), the quantum yields were higher in duplexes than in single strands and superior to the highest value determined for free **1** in aprotic solvents (QY = 20% in dimethyl sulfoxide).[8] These observations are consistent with the dye being shielded in the duplex, with reduced quenching by the solvent. In addition, reduced flexibility and/or rotation around the single bond between the chromone and thienyl moieties might also contribute to the higher quantum yield, as proposed for thiazole orange, green fluorescent protein (GFP)-like, and uracil-containing pyridine fluorophores.[14] The quantum yields in ODNs with C or G flanking the dye were lower (entries 14−22). The decrease in quantum yield in duplexes with G opposite **1** confirmed the propensity of G to quench the dye (Table 2, entries 5, 11, and 17). Quenching of fluorescent nucleoside analogues in ODNs is common.[1b] For example, G, T, and C efficiently quench pyrene,[15] while 2-AP is quenched by all four natural nucleobases.[2] Importantly, the fluorescence quantum yield of **1** is 2−25-fold larger than that of 2-AP in corresponding ODN sequences.[16] In addition, the molar absorptivity of **1** (13 000 M^{-1} cm^{-1}) is about twice that of 2-AP (7200 M^{-1} cm^{-1}), which makes it up to 50-fold brighter.[8,16]

In comparison to free dye **1** in buffer, the labeled ODNs showed a strong decrease in the N*/T* intensity ratio (I_{N*}/I_{T*}) together with blue and red shifts of the N* and T* bands, respectively (Table 2 and Figure 3). For most of the labeled ODNs, I_{N*}/I_{T*} was between those for the free dye in dichloromethane and acetonitrile (Figure 3), indicating that the environment of the labeling site is mainly aprotic and of medium polarity.[8] Here we compared only ODNs showing sufficient fluorescence quantum yields (≥4%; entries 2−14). In ssODNs, I_{N*}/I_{T*} varied strongly with the nature of the neighboring bases. **TMT** and **CMC** showed larger I_{N*}/I_{T*} than **AMA**, likely as a consequence of more efficient stacking of dye **1** with purines and thus more efficient shielding from water. These results are consistent with the UV spectroscopy data, where the more pronounced red shift of the 3HC absorption maximum was observed with flanking adenines in **AMA** (see above). In dsODNs, I_{N*}/I_{T*} rather uniformly showed lower values, with the exception of **T** opposite **1** in both the **AMA** and **TMT** duplexes (entries 3 and 9). It is clear that in dsDNA, the environment of **1** is highly dehydrated, which explains the low I_{N*}/I_{T*} values. The anomalous effect of having **T** opposite **1** could be explained by H-bonding between this base and the 4-carbonyl of **1**, which would hamper the ESIPT and increase I_{N*}/I_{T*}. Finally the highest N*/T* intensity ratio was observed for **M-TAT** ssDNA, where dye **1**, located at the 5′ end, is more exposed to water. Thus, both the fluorescence quantum yield and I_{N*}/I_{T*} for thienyl-3HC dye **1** show high sensitivity to the neighboring bases, which is a unique feature with respect to the existing fluorescent bases.[1b]

Furthermore, the photostabilities of nucleoside **1** and its ssODN in buffer were comparable to that of prodan in ethanol, whereas for the corresponding dsODN the photostability was remarkably higher (Figure S7 in the Supporting Information).

Finally, to validate the application of the new fluorescent base as a probe for sensing biomolecular interactions, we studied the interaction of the labeled ODNs with the HIV-1 nucleocapsid protein (NC). This nucleic acid chaperone protein can locally destabilize ssODNs or hairpins through interactions with the hydrophobic amino acids of its folded zinc finger motifs.[17] Addition of NC to ssODNs (**AMA**, **TMT**, and **CMC**) significantly increased their N*/T* intensity ratios, showing a 1:1 interaction stoichiometry (Figure 4 and Figure S8 in the Supporting Information).

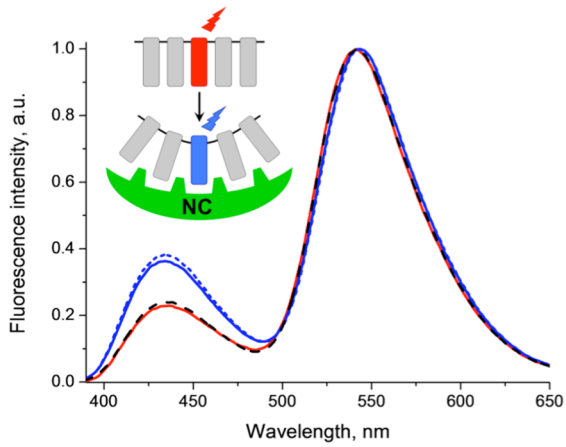


Figure 4. Fluorescence spectra of the **TMT** ODN in the absence (red) or presence of 1 equiv (blue solid) or 3 equiv (blue dotted) of NC or 3 equiv of NC-SSHS (black dashed) peptides. The spectra were normalized at the T* band. The experiments were performed in pH 7 buffer (10 mM cacodylate, 150 mM NaCl). Binding of the NC peptide to the labeled ODN leads to an increase in I_{N*}/I_{T*} due to the insertion of the hydrophobic amino acids of its folded finger motifs between the bases.

As already described for several other NC−ODN complexes, these observations suggest that intercalation of the zinc finger residues between the bases of these labeled ssODNs, at the level of **1**, decreases the base stacking and increases the exposure of the bases to water.[18] In contrast, no change in the probe signal was observed upon interaction of NC with the labeled dsODNs (data not shown) because of the limited destabilizing effect of NC on stable ds regions.[19] Moreover, a nonfolded mutant of NC (NC-SSHS) that binds ssODNs only through electrostatic interactions with the ODN phosphate groups[18a,20] induced no change in the emission spectra of the labeled ODN sequences (Figure 4). These data confirmed that, in contrast to NC, the NC-SSHS mutant does not affect the base stacking in ssODNs. Thus, our new fluorescent base allows site-selective monitoring of subtle conformational changes produced by interacting proteins.

CONCLUSION

ODNs incorporating the two-band fluorescent thienyl-3HC **1** were successfully synthesized. Dye **1** can replace any natural nucleobase or base pair to form duplexes with the B

dx.doi.org/10.1021/ja3030388 | J. Am. Chem. Soc. 2012, 134, 10209−10213

conformation having nearly the same stability as the native nonlabeled duplex. Thus, **1** possesses the characteristics of a universal base. In comparison to the "gold standard" 2-AP, the new base is up to 50-fold brighter, and its absorption is red-shifted by 60 nm. Consequently, the absorption of the new base does not overlap with the intrinsic absorptions of nucleic acids and proteins. The new fluorescent base demonstrates all the advantages of wavelength ratiometric detection, including a large separation between the emissions of the two bands and a dramatic variation of their relative contributions. Moreover, the high sensitivity of the dual emission of **1** to the polarity of the environment induces a strong shielding effect of the flanking bases from water. Finally, the new base was successfully applied to monitor local conformational changes of ssODNs upon interaction with the viral nucleocapsid protein. Therefore, the new base appears to be a new universal tool for DNA research. Further chemical modifications of 3HC nucleobases will provide new possibilities to tune and improve their properties.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental procedures, analytical data, and spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

rachid.benhida@unice.fr; burger@unice.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the ANR (07-BLAN-0287) and the FRM (DCM20111223038) for financial support and grants for D.D. and V.Y.P. We thank Prof. Alexander Demchenko and Dr. Jennifer Wytko for their fruitful comments and critical reading of the manuscript.

■ REFERENCES

(1) For reviews, see: (a) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265−4274. (b) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579−2619. (c) Srivatsan, S. G.; Sawant, A. A. *Pure Appl. Chem.* **2011**, *83*, 213−232. (d) Dai, N.; Kool, E. T. *Chem. Soc. Rev.* **2011**, *40*, 5756−5770.

(2) (a) Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969**, *244*, 1228−1237. (b) Rachofsky, E. L.; Osman, R.; Ross, J. B. *Biochemistry* **2001**, *40*, 946−956.

(3) Dual emission and ratiometric measurement are commonly obtained with doubly labeled ODNs (e.g., FRET, excimer).[1d] Few examples with singly labeled ODNs are known. For ODNs labeled with pH-sensitive or charge-transfer dyes, see: (a) Kashida, H.; Sano, K.; Hara, Y.; Asanuma, H. *Bioconjugate Chem.* **2009**, *20*, 258−265. (b) Okamoto, A.; Tainaka, K.; Nishiza, K.; Saito, I. *J. Am. Chem. Soc.* **2005**, *127*, 13128−13129.

(4) (a) Demchenko, A. P. *FEBS Lett.* **2006**, *580*, 2951−2957. (b) Demchenko, A. P.; Klymchenko, A. S.; Pivovarenko, V. G.; Ercelen, S.; Duportail, G.; Mély, Y. *J. Fluoresc.* **2003**, *13*, 291−295.

(5) (a) Das, R.; Klymchenko, A. S.; Duportail, G.; Mély, Y. *Photochem. Photobiol. Sci.* **2009**, *8*, 1583−1589. (b) Chou, P. T.; Martinez, M. L.; Clements, J. H. *J. Phys. Chem.* **1993**, *97*, 2618−2622. (c) Hsieh, C.-C.; Jiang, C.-M.; Chou, P.-T. *Acc. Chem. Res.* **2010**, *43*, 1364−1374.

(6) (a) Demchenko, A. P. *Trends Biotechnol.* **2005**, *23*, 456−460. (b) Demchenko, A. *J. Fluoresc.* **2010**, *20*, 1099−1128.

(7) (a) Demchenko, A. P.; Mely, Y.; Duportail, G.; Klymchenko, A. S. *Biophys. J.* **2009**, *96*, 3461−3470. (b) Shynkar, V. V.; Klymchenko, A. S.; Kunzelmann, C.; Duportail, G.; Muller, C. D.; Demchenko, A. P.; Freyssinet, J.-M.; Mely, Y. *J. Am. Chem. Soc.* **2007**, *129*, 2187−2193. (c) Shvadchak, V. V.; Klymchenko, A. S.; de Rocquigny, H.; Mély, Y. *Nucleic Acids Res.* **2009**, *37*, No. e25. (d) Yushchenko, D. A.; Fauerbach, J. A.; Thirunavukkuarasu, S.; Jares-Erijman, E. A.; Jovin, T. M. *J. Am. Chem. Soc.* **2010**, *132*, 7860−7861.

(8) Spadafora, M.; Postupalenko, V. Y.; Shvadchak, V. V.; Klymchenko, A. S.; Mély, Y.; Burger, A.; Benhida, R. *Tetrahedron* **2009**, *65*, 7809−7816.

(9) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **1996**, *118*, 8182−8183.

(10) (a) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191−6192. (b) Berger, M.; Ogawa, A. K.; McMinn, D. L.; Wu, Y. Q.; Schultz, P. G.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2000**, *39*, 2940−2942. (c) Singh, I.; Hecker, W.; Prasad, A. K.; Parmar, V. S.; Seitz, O. *Chem. Commun.* **2002**, 500−501. (d) Brotschi, C.; Mathis, G.; Leumann, C. J. *Chem.—Eur. J.* **2005**, *11*, 1911−1923.

(11) Loakes, D. *Nucleic Acids Res.* **2001**, *29*, 2437−2447.

(12) Ogawa, A. K.; Abou-Zied, O. K.; Tsui, V.; Jimenez, R.; Case, D. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 9917−9920.

(13) (a) Asseline, U.; Delarue, M.; Lancelot, G.; Toulmé, F.; Thuong, N. T.; Montenay-Garestier, T.; Hélène, C. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3297−3301. (b) Amann, N.; Huber, R.; Wagenknecht, H.-A. *Angew. Chem., Int. Ed.* **2004**, *43*, 1845−1847.

(14) (a) Köhler, O.; Jarikote, D. V.; Singh, I.; Parmar, V. S.; Weinhold, E.; Seitz, O. *Pure Appl. Chem.* **2005**, *77*, 327−338. (b) Wenge, U.; Wagenknecht, H.-A. *Synthesis* **2011**, 502−508. (c) Sinkeldam, R. W.; Marcus, P.; Uchenik, D.; Tor, Y. *ChemPhysChem* **2011**, *12*, 2260−2265.

(15) Wilson, J. N.; Cho, Y.; Tan, S.; Cuppoletti, A.; Kool, E. T. *ChemBioChem* **2008**, *9*, 279−285.

(16) (a) Ben Gaied, N.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mély, Y. *Nucleic Acids Res.* **2005**, *33*, 1031−1039. (b) Kenfack, C. A.; Piemont, E.; Ben Gaied, N.; Burger, A.; Mély, Y. *J. Phys. Chem. B* **2008**, *112*, 9736−9745.

(17) Godet, J.; Mély, Y. *RNA Biol.* **2010**, *7*, 687−699.

(18) (a) Godet, J.; Ramalanjaona, N.; Sharma, K. K.; Richert, L.; de Rocquigny, H.; Darlix, J.-L.; Duportail, G.; Mély, Y. *Nucleic Acids Res.* **2011**, *39*, 6633−6645. (b) Bourbigot, S.; Ramalanjaona, N.; Boudier, C.; Salgado, G. F. J.; Roques, B. P.; Mély, Y.; Bouaziz, S.; Morellet, N. *J. Mol. Biol.* **2008**, *383*, 1112−1128.

(19) (a) Beltz, H.; Azoulay, J.; Bernacchi, S.; Clamme, J. P.; Ficheux, D.; Roques, B.; Darlix, J. L.; Mély, Y. *J. Mol. Biol.* **2003**, *328*, 95−108. (b) Egele, C.; Schaub, E.; Ramalanjaona, N.; Piemont, E.; Ficheux, D.; Roques, B.; Darlix, J. L.; Mély, Y. *J. Mol. Biol.* **2004**, *342*, 453−466. (c) Heilman-Miller, S. L.; Wu, T.; Levin, J. G. *J. Biol. Chem.* **2004**, *279*, 44154−44165.

(20) Avilov, S. V.; Piemont, E.; Shvadchak, V.; de Rocquigny, H.; Mély, Y. *Nucleic Acids Res.* **2008**, *36*, 885−896.

Supporting information

A universal nucleoside with strong two-band switchable fluorescence and sensitivity to environment for investigating DNA interactions

Dmytro Dziuba, Viktoriia Y. Postupalenko, Marie Spadafora, Andrey S. Klymchenko, Vincent Guérineau, Yves Mély, Rachid Benhida, Alain Burger

Table of content

1.0 Abbreviations.

3HC – 3-Hydroxychromone; Ab – abasic site; Cbz – Carboxybenzyl; CD – Circular dichroism; CH – Cyclohexane; DCM – dichloromethane; DIEA – N,N-Diisopropylethylamine; DMF – dimethylformamide; DMTr – 4,4'-Dimethoxytrityl; EA – Ethyl acetate; ESI – Electrospray ionization; FC – Flash chromatography; HPLC – High-performance liquid chromatography; HRMS – High resolution mass spectra; Imid – Imidazole; MALDI – matrix-assisted laser desorption/ionization; MS – mass spectra; ODN – oligondeoxynucleotide; NC – nucleocapsid protein from HIV-1; Py – Pyridine; RT – retention time, TBDPSCl – tert-butyldiphenylsilyl chloride; TEA – Triethylamine; THF – tetrahydrofuran; TLC – Thin layer chromatography; TOF - time-of-flight.

2.0 Experimental section.

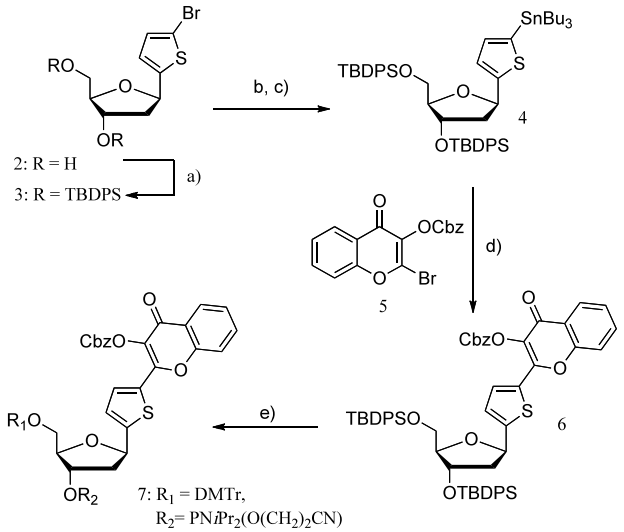
2.1. Preparation of 3HC containing oligodeoxynucleotides (ODNs)

2.2.1. General

Reagents and solvents were purchased from Sigma-Aldrich and AlfaAesar and were used without further purification. Reactions involving water-sensitive reagents were carried out in dry solvents under argon. Solvents for these reactions were dried according to standard procedures.[1a] ^{1}H and ^{13}C NMR spectra were recorded on a Bruker Advance Spectrometer 200 instrument (200 MHz for ^{1}H and 50 MHz for ^{13}C). Chemical shifts (δ, ppm) were determined using solvent residual signals as the reference.[1b] TLC was conducted on "Silica gel 60 F_{254}" plates (MERCK). Flash chromatography [1a] was carried out on "Geduran Silica gel 60 (40-63 μm)" (MERCK). Regular mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative modes. High resolution mass spectra (HRMS) were recorded on Hybride LTQ Orbitrap (Thermoscientific) apparatus.

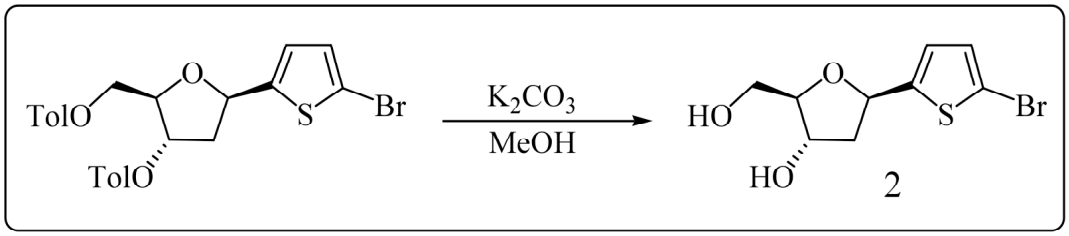
2.2.2. Synthesis of 3HC nucleoside phosphoramidite.

The strategy of the 3HC bearing fluorescent nucleoside synthesis was adapted from our recent publication.[2a] The starting 1β-(5-Bromothiophen-2-yl)-1,2-dideoxy-3,5-di-*O*-toluoyl-D-ribofuranose was synthesized as described by Barta and co-workers.[2b] Cbz-protected 2-Bromo-3-HC (**5**) was synthesized as described previously.[2c]



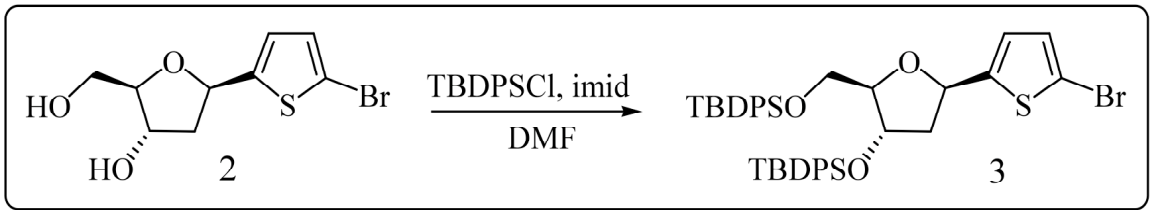
Synthesis of phosphoramidite **7**: a) TBDPSCl, Imidazole, DMF, RT, 97%; b) BuLi, THF, - 78 °C, 30 min; c) Bu_3SnCl, THF, - 78° C to 0 °C, 1h; d) $Pd(PPh_3)_4$, CuI, toluene, reflux, 2 h, 50%; e) i Pyridine-HF, THF, RT, 72-96 h, 78%; ii DMTrCl, Py, NEt_3, 24 h, 65%; iii $iPr_2NP(Cl)OEtCN$, iPr_2NEt, CH_2Cl_2, RT, 2h, 78%. TBDPS= *t*-Butyldiphenylsilyl, Cbz= Benzylcarbonate, DMTr= 4,4'-Dimethoxytrityl.

1-β-(5-Bromo-thiophen-2-yl)-deoxyribose (**2**)



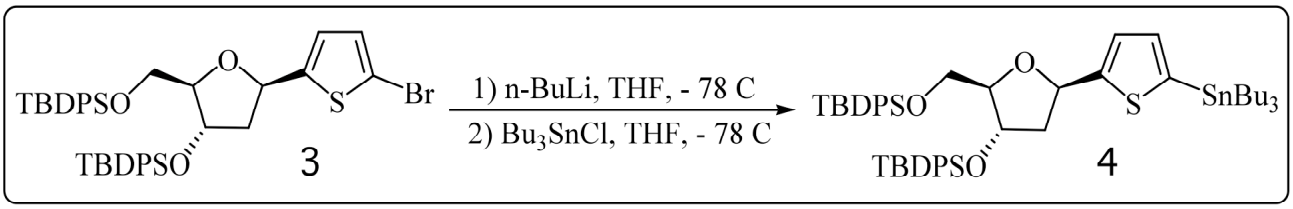
To a solution of Tol-protected C-glycoside [2b] (2.45 g, 4.76 mmol) in MeOH (28 mL), potassium carbonate (1.3 g, 9.42 mmol) was added at once. The mixture was stirred for 1 hour at room temperature (TLC, DCM / MeOH = 9/1). Then, the solid was removed by filtration and washed with DCM – MeOH mixture (1/1, v/v). The filtrate was concentrated under vacuum. The residue was purified by silica gel flash chromatography eluted with a gradient from DCM to DCM / MeOH (95/5). The product **2** was obtained as a pale-yellow solid (1.3 g, 98%). R_f = 0.43 (DCM / MeOH = 9 / 1). ^{1}H-NMR (DMSO, 200 MHz, δ): 7.05 (d, *J* = 3.7 Hz, 1H, H-4), 6.89 (dd, *J* = 3.8, 0.6 Hz, 1H, H-3), 5.19 (dd, *J* = 9.9, 5.5 Hz, 1H, H-1'), 5.11 (d, *J* = 3.8 Hz, 1H, HO-C-3'), 4.76 (t, *J* = 5.5 Hz, 1H, HO-C-5'), 4.30 – 4.07 (m, 1H, H-3'), 3.74 (ddd, *J* = 6.1, 5.0, 2.0 Hz, 1H, H-4'), 3.50 – 3.24 (m, 2H, H-5'), 2.11 (ddd, *J* = 12.7, 5.6, 1.7 Hz, 1H, H-2'a), 1.87 (ddd, *J* = 12.7, 10.0, 5.4 Hz, 1H, H-2'b). ^{13}C-NMR (CD_3OD, 50 MHz, δ): 148.94, 130.62, 125.98, 112.28, 89.20, 77.49, 74.22, 63.96, 44.74. MS (ESI, MeOH, neg) m/z: 277.0 + 278.8 (M – H).

(2*R*, 3*S*, 5*R*) 5-(5-Bromo-thiophen-2-yl)-3-(*tert*-butyl-diphenyl-silanyloxy)-2-(*tert*-butyl-diphenyl-silanyloxymethyl)-tetrahydrofuran **3**.



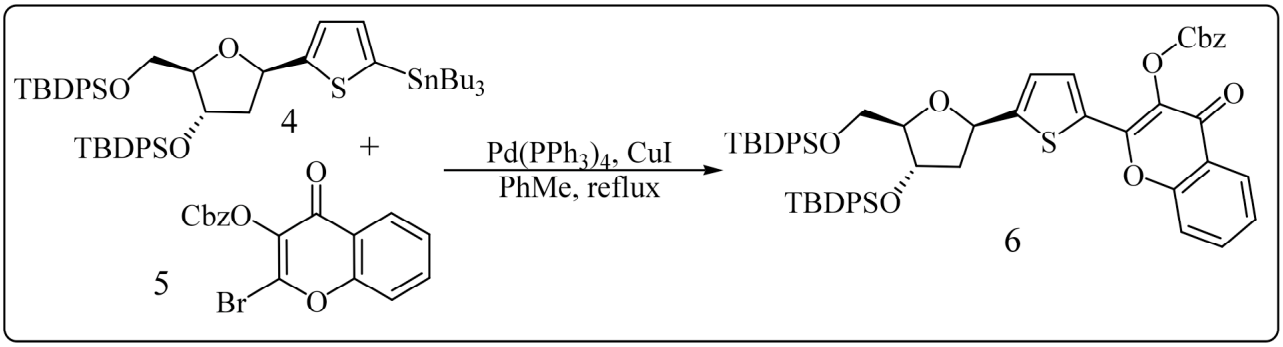
To a solution of the diol **2** (1.28 g, 4.59 mmol) in dry DMF (15 mL), TBDPSCl (2.98 mL, 11.46 mmol, 2.5 equiv.) and imidazole (1.25 g, 18.36 mmol, 4 eqiv) were added sequentially. The reaction mixture was stirred at ambient temperature overnight, quenched with brine (10 mL) and extracted tree times with DCM (3 × 50 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated under vacuum. The crude product was purified by silica gel flash chromatography with a gradient of CH to CH / EA (95/5). The product **3** was obtained as a viscous colorless oil (3.45 g, 97%). R_f = 0.85 (CH/EA = 9 / 1). ^{1}H-NMR ($CDCl_3$, 200 MHz, δ): 7.75 – 7.27 (m, 20H), 6.86 (d, *J* = 3.7 Hz, 1H), 6.72 (d, *J* = 3.7 Hz, 1H), 5.38 (dd, *J* = 10.7, 4.8 Hz, 1H), 4.55 (d, *J* = 4.8 Hz, 1H), 4.08 (m, 1H), 3.47 (dd, *J* = 11.0, 4.0 Hz, 1H), 3.32 (dd, *J* = 11.0, 3.8 Hz, 1H), 2.17 (dd, *J* = 12.7, 4.7 Hz, 1H), 1.88 (ddd, *J* = 12.7, 10.8, 5.1 Hz, 1H), 1.09 (s, 9H), 0.93 (s, 9H). ^{13}C-NMR ($CDCl_3$, 50 MHz, δ): 147.24, 135.87, 135.76, 133.87, 133.72, 133.28, 133.25, 129.98, 129.77, 129.74, 129.29, 127.91, 127.78, 127.74, 124.69, 111.55, 88.54, 75.58, 64.36, 44.69, 27.11, 26.92, 19.26. MS (MeOH) *m/z*: 777.3 + 779.3 $(M + H)^+$.

(2*R*, 3*S*, 5*R*) 3-(*tert*-Butyl-diphenyl-silanyloxy)-2-(*tert*-butyl-diphenyl-silanyloxymethyl)-5-(5-tributylstannylthiophen-2-yl)-tetrahydrofuran (**4**).



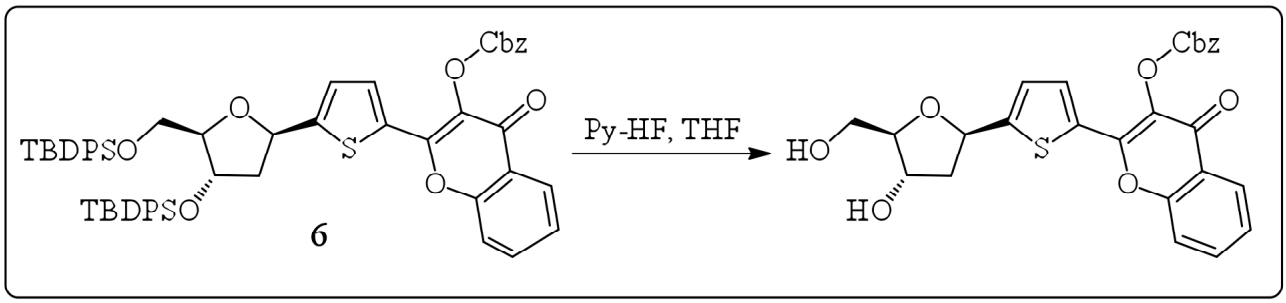
To a cooled solution (–78°C) of compound **3** (2.44 g, 3.23 mmol) in 20 mL of dry THF under argon, a BuLi solution in hexane (1.6 M, 2.34 mL, 1.15 equiv) was added dropwise. After 30 min stirring at –78°C, Bu_3SnCl (1.1 mL, 4.06 mmol, 1.25 equiv) was added dropwise. The cooling bath was then removed and the reaction mixture was slowly warmed up to room temperature. The solvent was evaporated under reduced pressure. The crude product **4** obtained was engaged into the next step without further purification.

(2*R*, 3*S*, 5*R*) Benzyl 2-{5-[4-(*tert*-butyl-diphenyl-silanyloxy)-5-(*tert*-butyl-diphenyl-silanyloxymethyl)-tetrahydrofuran-2-yl]-thiophen-2-yl}-4-oxo-4*H*-chromen-3-yl carbonate **6**



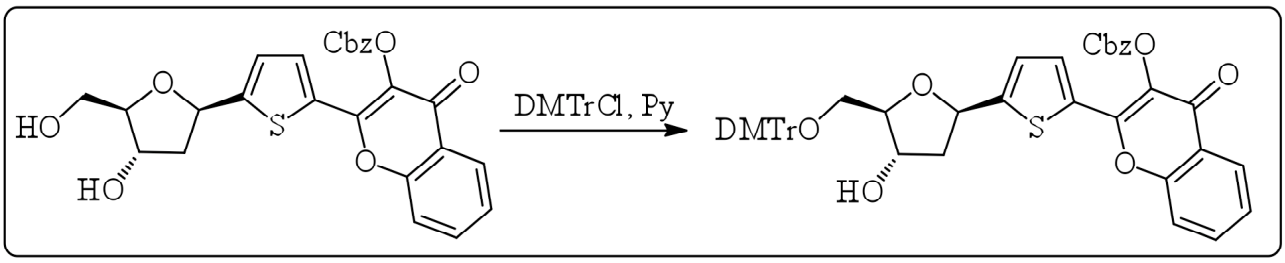
The crude product **4** and Cbz-protected 2-bromo-3-HC **5** [2c] (1.3 g, 3.55 mmol, 1.1. equiv) were dissolved in dry toluene (22 mL). The solution was degassed with argon before $Pd(PPh_3)_4$ (186 mg, 5% mol) and CuI (60 mg, 10% mol) were added at once. The reaction mixture was refluxed under argon for 2 hours. After completion of the reaction, as evidenced by TLC (CH/EA = 8/2), the mixture was cooled down to room temperature. The black solution was filtered through celite and concentrated under reduced pressure. The residue was subjected to flash chromatography and eluted with a mixture of CH/EA (9/1) to give the product **6** as a dark oil (1.56 g, 50% over two steps). (*The compound synthesized and purified by this way often retains tin-containing byproducts. Nevertheless, the tin byproducts do not interfere in the reaction of the next step, and are more conveniently removed at this step. Analytical data only for the main reaction product are given*). R_f = 0.43 (CH/EA = 8/2). ^{1}H-NMR ($CDCl_3$, 200 MHz, δ): 8.28 – 8.20 (m, 1H), 7.75 – 7.27 (m, 29 H), 7.04 (d, *J* = 3.9 Hz, 1H), 5.51 (dd, *J* = 10.8, 4.8 Hz, 1H), 5.31 (s, 2H), 4.59 (bd, *J* = 4.6 Hz, 1H), 4.17 – 4.09 (m, 1H), 3.53 (dd, *J* = 11.1, 3.7 Hz, 1H), 3.32 (dd, *J* = 11.1, 3.6 Hz, 1H), 2.35 – 2.20 (m, 1H), 2.03 – 1.86 (m, 1H), 1.12 (s, 9H), 0.93 (s, 9H). ^{13}C-NMR ($CDCl_3$, 50 MHz, δ): 171.55, 155.16, 153.07, 152.18, 151.83, 135.85, 135.72, 134.78, 134.00, 133.74, 133.62, 133.20, 133.15, 131.70, 131.18, 130.03, 129.90, 129.75, 128.81, 128.74, 128.42, 127.93, 127.80, 127.77, 126.13, 125.22, 123.75, 117.93, 88.85, 75.72, 71.17, 64.28, 43.57, 27.10, 27.02, 26.94, 26.88, 19.24. HRMS: *m/z* [M+H] calc for $C_{58}H_{59}O_8SSi_2$: 971.3464; found: 971.3461.

(2*R*, 3*S*, 5*R*) Benzyl 2-[5-(4-hydroxy-5-hydroxymethyl-tetrahydro-furan-2-yl)-thiophen-2-yl]-4-oxo-4*H*-chromen-3-yl carbonate.



Compound **6** (1.55 g) was dissolved in THF (15 mL) and the solution was poured in a plastic tube. Py-HF complex (1 mL) was added at once. The reaction mixture was kept at room temperature in the well capped plastic tube during 4 – 5 days. (TLC, DCM / MeOH = 9/1). After completion of the reaction (TLC), the mixture was taken in a plastic syringe and added dropwise to a suspension of silica gel (8 g, 100-200 µm particle size) in dry THF (25 mL) cooled to 0 ºC. The solid matter was isolated by filtration and washed with THF and DCM. The filtrate was concentrated under vacuum. The crude product was purified by flash chromatography and eluted using a gradient of DCM to DCM/MeOH (9/1) to give the diol as a pale yellow solid (m = 600 mg, 76%). R_f = 0.62 (DCM / MeOH = 9 / 1). ^{1}H-NMR (DMSO-d6, 200 MHz, δ): 8.08 (dd, *J* = 8.0, 1.1 Hz, 1H), 7.95 – 7.74 (m, 3H), 7.60 – 7.36 (m, 6H), 7.28 (d, *J* = 3.9 Hz, 1H, H-thiophene), 5.45 – 5.28 (m, 3H, H-1' + -O-$\underline{CH_2}$-Ph), 5.23 (d, *J* = 3.8 Hz, 1H, C-4' –O$\underline{H}$), 4.87 (t, *J* = 5.5 Hz, 1H, C-5' – O$\underline{H}$), 4.37 – 4.18 (m, 1H, H-3'), 3.93 – 3.78 (m, 1H, H-4'), 3.61 – 3.23 (m, 2H, H-5'), 2.23 (dd, *J* = 11.5, 5.6 Hz, 1H, H-2'), 1.92 (ddd, *J* = 12.7, 10.2, 5.3 Hz, 1H, H-2'). ^{13}C-NMR (DMSO-d6, 50 MHz, δ): 170.47, 154.80, 154.60, 151.52, 151.16, 134.82, 131.53, 130.47, 128.66, 128.61, 128.18, 128.00, 125.79, 125.38, 125.01, 122.77, 118.43, 88.18, 75.26, 72.40, 70.72, 62.39, 43.52. HRMS: *m/z* [M+H] calc for $C_{26}H_{23}O_8S$: 495.1108; found: 495.1107.

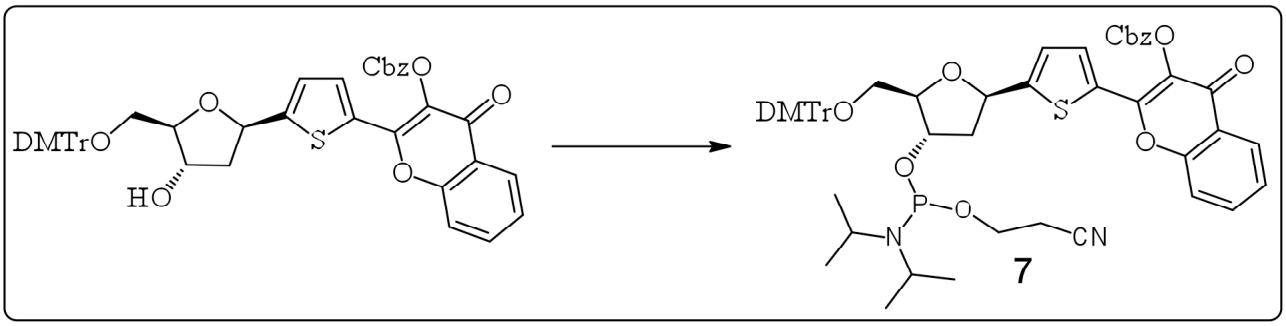
Benzyl 2-(5-{5-[bis-(4-methoxy-phenyl)-phenyl-methoxymethyl]-4-hydroxy-tetrahydrofuran-2-yl}-thiophen-2-yl)-4-oxo-4H-chromen-3-yl carbonate.



The diol (290 mg, 0.59 mmol) was azeotropically dried twice with dry pyridine (2 x1 mL) under reduced pressure. After that, it was dissolved in dry Py (5 mL) and DMTrCl (210 mg, 1.05 equiv) was added in one portion. The solution was stirred for 24 hours at ambient temperature. After completion of the reaction (TLC, DCM / EA, 9/1), TEA (0.12 mL, 1.5 equiv), DCM (30 mL) and brine (5 mL) were added sequentially. The organic layer was separated and the water phase was extracted twice with DCM (2 × 25 mL). The combined organic layers were dried over Na_2SO_4 and evaporated under reduced pressure to dryness. The residue was purified by FC eluting with a mixture of DCM / EA (99/1 + 1% TEA) to afford the trityl ether as a yellow solid (305 mg, 65%). R_f = 0.42 (DCM / EA = 9 / 1). ^{1}H-NMR ($CDCl_3$, 200 MHz, δ): 8.24 (dd, *J* = 8.0, 1.5 Hz, 1H), 7.77 (d, *J* = 4.0 Hz, 1H), 7.74 – 7.60 (m, 2H), 7.57 – 7.17 (m,

15H), 7.07 (d, *J* = 4.0 Hz, 1H), 6.83 (d, *J* = 8.9 Hz, 4H), 5.44 (dd, *J* = 9.7, 5.9 Hz, 1H, H-1'), 5.12 (s, 2H, O–$\underline{CH_2}$–Ph), 4.52 – 4.40 (m, 1H, H-3'), 4.14 – 4.01 (m, 1H, H-4'), 3.75 (s, 6H, 2 × –OCH_3), 3.35 – 3.22 (m, 2H, H-5'), 2.42 – 2.10 (m, 2H, H-2'). ^{13}C-NMR ($CDCl_3$, 50 MHz, δ): 171.55, 158.47, 155.08, 153.46, 152.09, 151.81, 144.95, 136.09, 136.06, 134.69, 134.05, 132.17, 132.11, 131.97, 131.51, 131.17, 130.13, 129.67, 128.74, 128.61, 128.50, 128.25, 128.21, 127.89, 126.97, 126.80, 125.93, 125.28, 125.05, 123.57, 117.97, 113.20, 86.95, 86.19, 76.19, 74.07, 70.93, 64.35, 55.21, 44.54. HRMS: *m/z* [M+H] calc for $C_{47}H_{41}O_{10}S$: 797.2415; found: 797.2416.

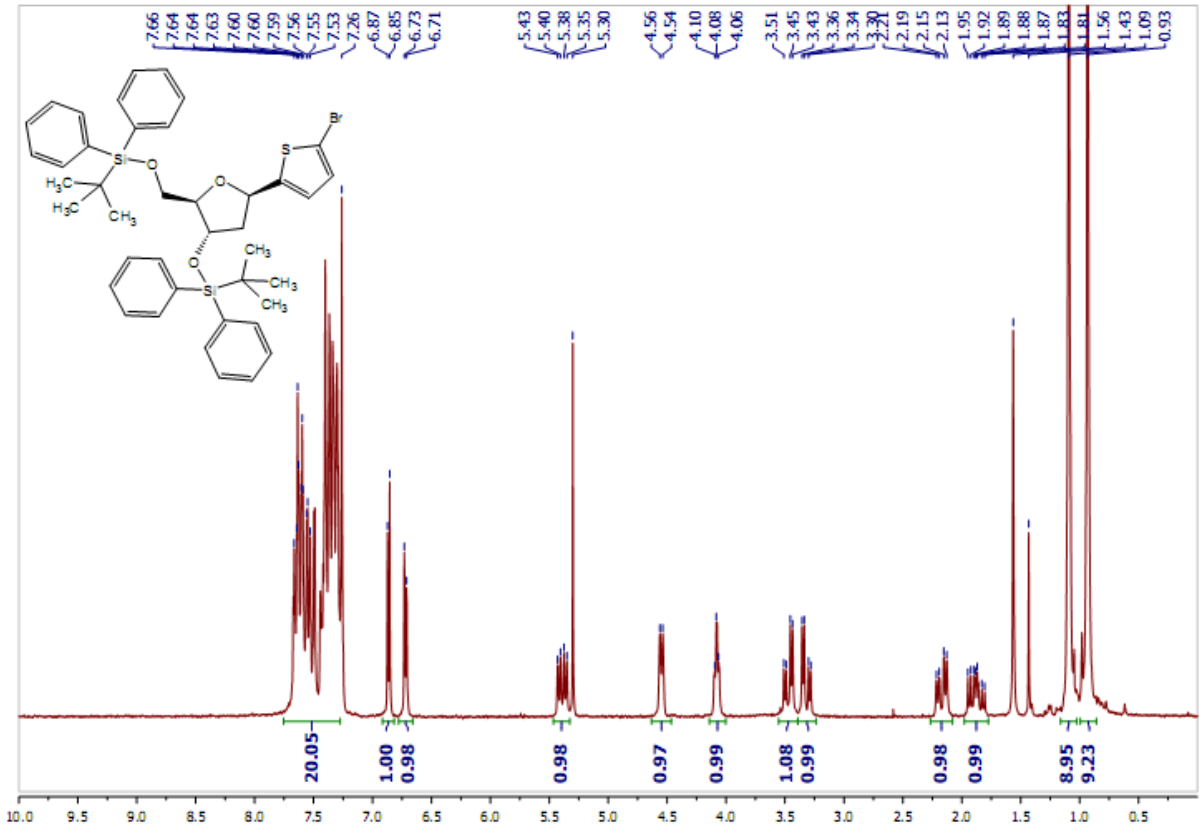
Benzyl 2-(5-{5-[bis-(4-methoxyphenyl)-phenyl-methoxymethyl]-4-[(2-cyano-ethoxy)-diisopropylamino-phosphanyloxy]-tetrahydrofuran-2-yl}-thiophen-2-yl)-4-oxo-4H-chromen-3-yl carbonate **7**.

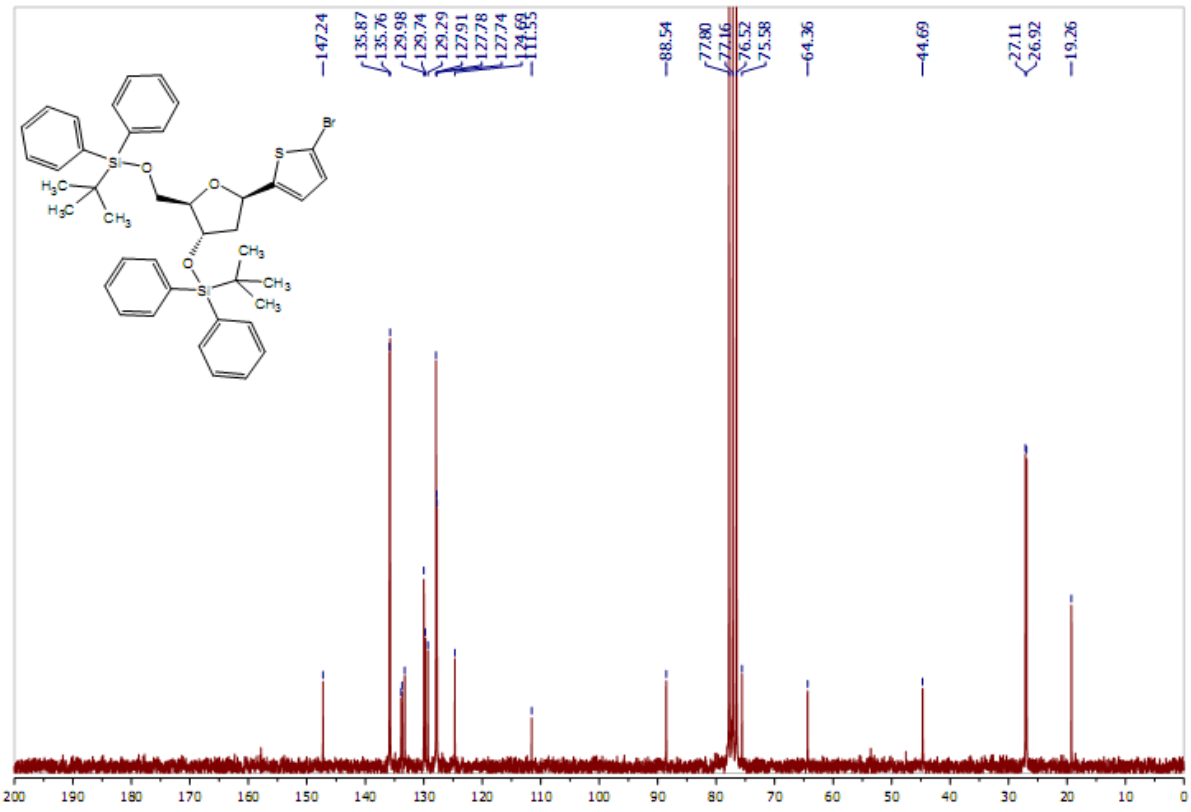


The trityl ether (300 mg, 0.38 mmol) was dissolved in dry DCM (3 mL). 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.1 mL, 1.2 equiv.) and DIEA (0.13 mL, 2 equiv) were added at once. The reaction mixture was stirred at room temperature for 2 hours (TLC, CH / EA = 6 / 4). The reaction mixture was subjected directly to silica gel FC eluting with a mixture of CH/EA (7/3 + 1% TEA) to give **7** as a yellow solid (290 mg, 77%, a mixture of two P-diastereomers). R_f = 0.43, 0.52 (CH/EA = 6/4). ^{1}H-NMR ($CDCl_3$, 200 MHz, δ): 8.24 (dd, *J* = 8.0, 1.5 Hz, 1H), 7.79 (d, *J* = 4.0 Hz, 1H), 7.75 – 7.61 (m, 1H), 7.55 – 7.16 (m, 16H), 7.11 (d, *J* = 4.0 Hz, 1H), 6.82 (d + d, *J* = 9.0 Hz, 4H), 5.43 (dd, *J* = 9.8, 4.6 Hz, 1H, H-1'), 5.07 + 5.06 (s + s, 2H, O–$\underline{CH_2}$–Ph), 4.67 – 4.46 (m, 1H), 4.34 – 4.16 (m, 1H), 3.94 – 3.47 (m, 10H), 3.46 – 3.12 (m, 2H, H-5'), 2.62 + 2.44 (t + t, *J* = 6.5 Hz, 2H, –$\underline{CH_2}$–CN), 2.54 – 2.11 (m, 2H, H-2') 1.28 – 0.88 (m, 12H). ^{31}P-NMR ($CDCl_3$, 81 MHz, δ): 148.09, 147.98. HRMS: *m/z* [M+H] calc for $C_{56}H_{58}N_2O_{11}PS$: 997.3499; found: 997.3515.

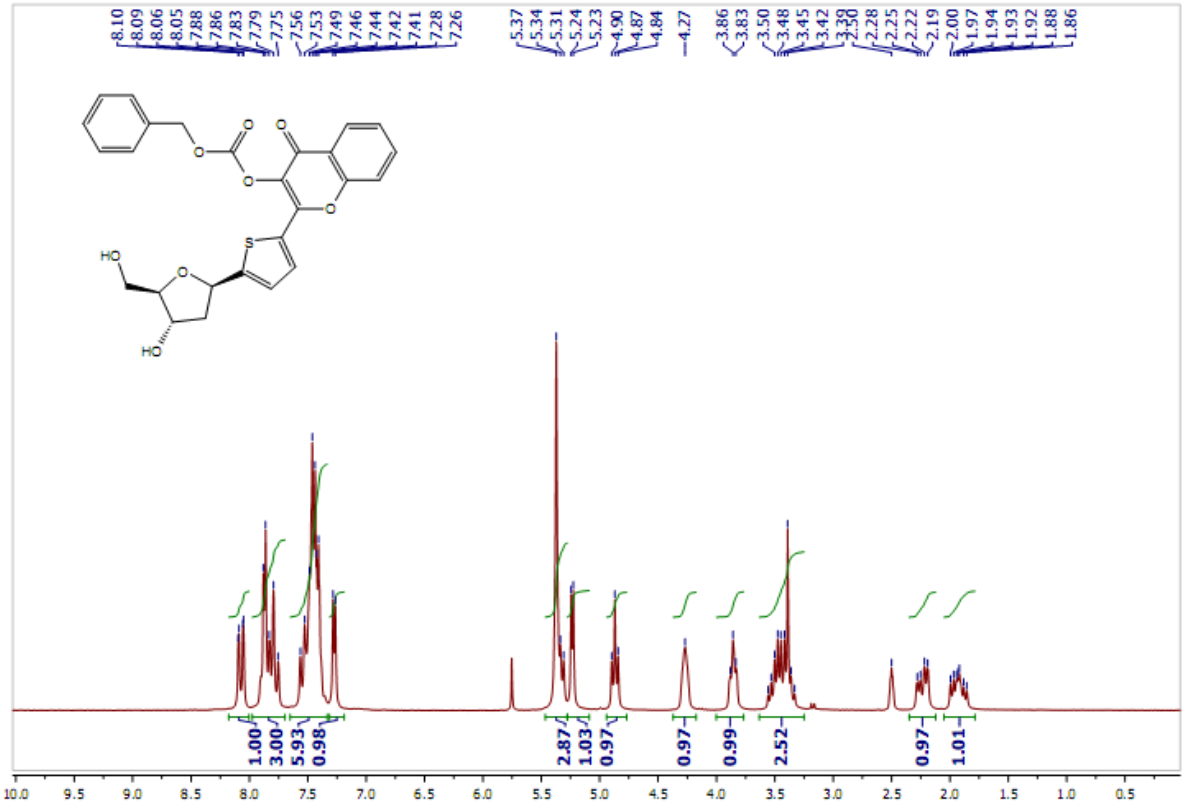
2.2.3. ^{1}H, ^{13}C and ^{31}P NMR figures.

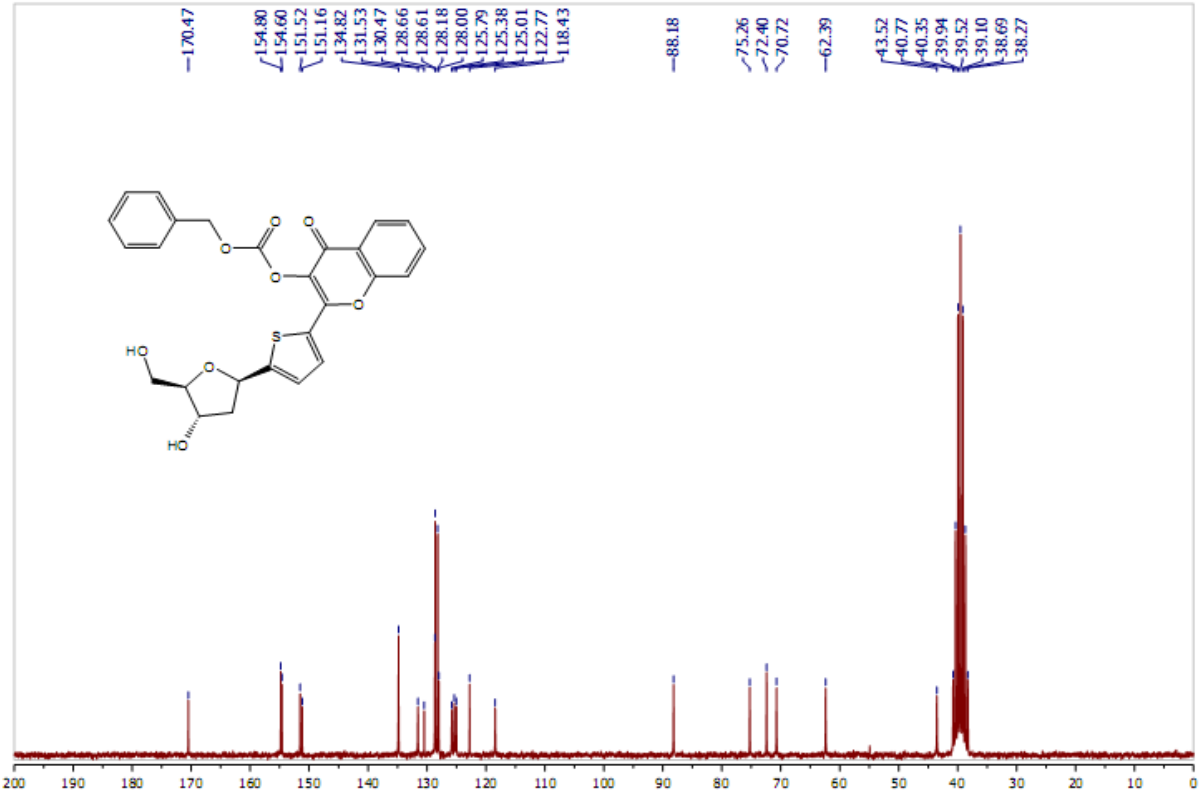
^{1}H and ^{13}C NMR spectra of the intermediate **3**.



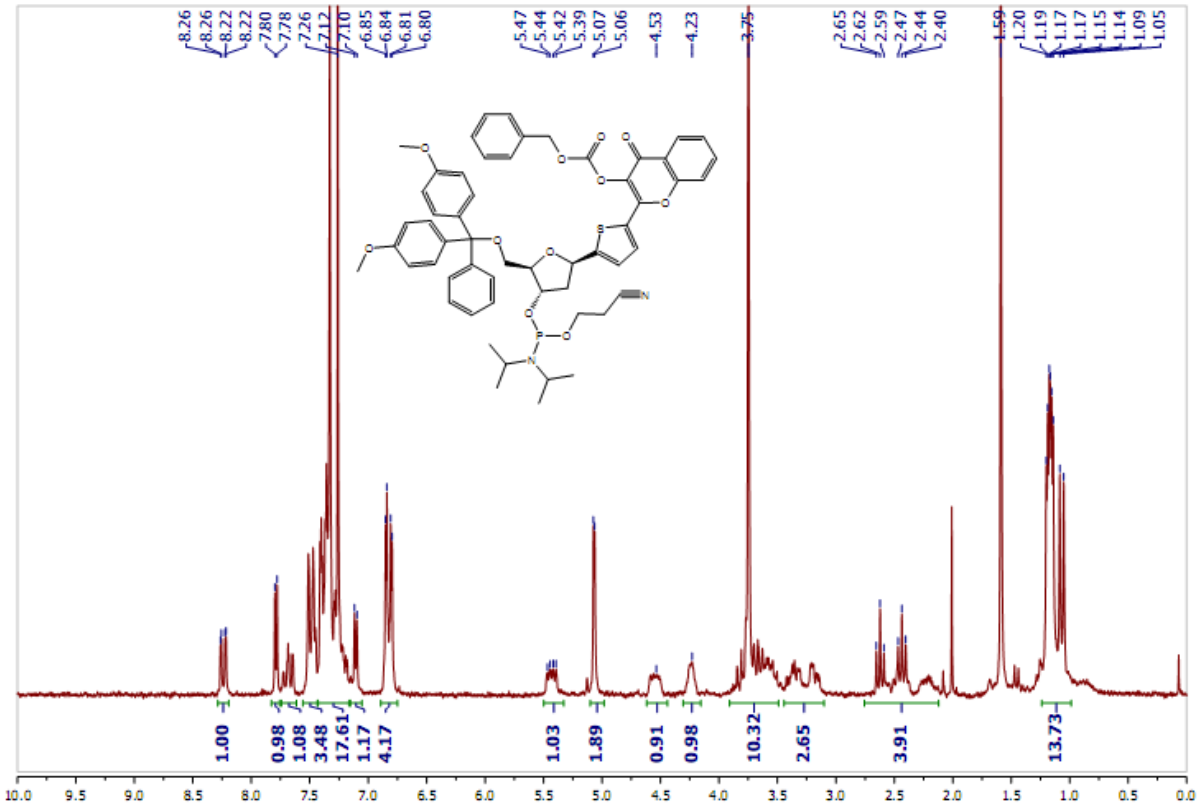


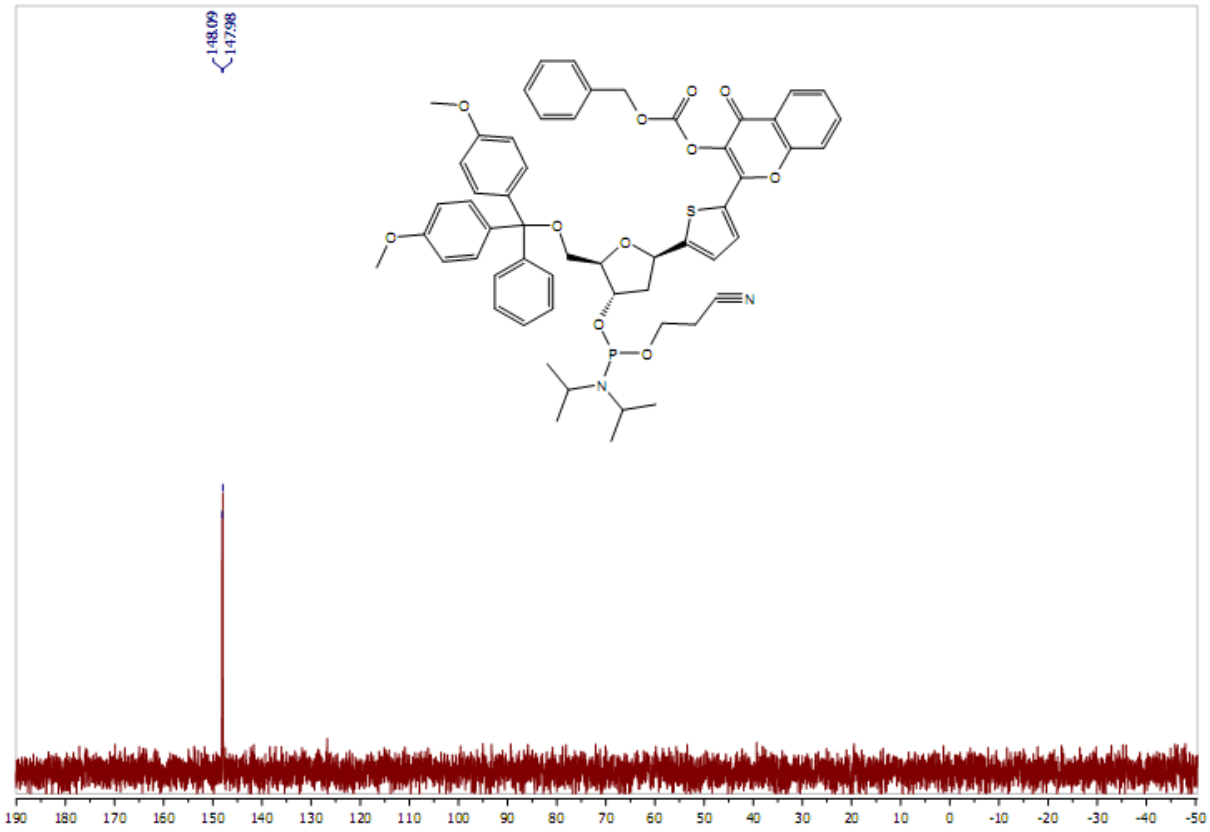
^{1}H and ^{13}C NMR spectra of the diol.





^{1}H and ^{31}P NMR spectra of the phosphoramidite **7**.





2.2.4. Oligodeoxyribonucleotides synthesis.

ODN synthesis was performed on an Expedite 8900 DNA synthesizer with Multiple oligo synthesis system (both from Applied Biosystem) using the "trityl on" mode and mild phosphoramidite chemistry on a 0.2 μmol scale. Reagents and solvents, as well as dT, Ac-dC, Pac-dA, and dmf-dG phosphoramidites were purchased from Link Technologies. The standard DNA assembly protocol "DMTr-on" was used except for the following modifications: 5-Ethylthio-1*H*-tetrazole (ETT) was used as activating agent; Pac-anhydride was used for capping; a longer coupling time (600 s) was applied to the 3HC phosphoramidite. Non labeled ODNs were purchased from IBA GmbH.

Table S1. Sequences of ODNs used in this work.

ODN	Notation[a]
d(CGT-TTT-**X(3HC)X**-TTT-TGC)	X = A (**AMA**), T (**TMT**), C(**CMC**), G(**GMG**)
5'-d(**3HC-CG**T-TTT-TAT-TTT-TGC)	**M-TAT**
d(CGT-TTT-**XAX**-TTT-TGC)	X = A (**AAA**), T (**TAT**), C(**CAC**), G(**GAG**)
d(GCA-AAA-**TXT**-AAA-ACG)	X = T (***TTT***), A (***TAT***), G (***TGT***), C (***TCT***), Ab (***TAbT***)
d(GCA-AAA-**AXA**-AAA-ACG)	X = T (***ATA***), A (***AAA***), G (***AGA***), C (***ACA***), Ab (***AAbA***)
d(GCA-AAA-**GXG**-AAA-ACG)	X = T (***GTG***), A (***GAG***), G (***GGG***), C (***GCG***), Ab (***GAbG***).
d(GCA-AAA-**CXC**-AAA-ACG)	X = T (***CTC***), A (***CAC***), G (***CGC***), C (***CCC***), Ab (***CAbC***)

a) 3HC **1** = **M**, Abasic site= ***Ab***.

2.2.5. ODN purification.

The ODNs were cleaved from the solid support and deprotected with concentrated aqueous ammonia at room temperature for 24 h. The DMT-derivatives of the ODNs were lyophilized to dryness and analyzed and purified by RP HPLC (HPLC apparatus: Waters™ 600 Controller with Waters™ 996 Photodiode Array Detector. Columns: analytical, Jupiter 5u C18 300A column 250*3 mm; semi preparative, Jupiter 5u C18 300A column 250*10 mm). The following gradient system was used: 100% A –(30 min)→ 60% A / 40% B –(5 min)→ 100% B –(5 min)→ 100%A with A = Buffer pH 7.0 (1.9 L of milliQ water, 160 mL acetonitrile, 28 mL triethylamine, 12 mL of acetic acid) and B = acetonitrile. After HPLC purification, the ODNs were detritylated in 50% acetic acid in H_2O for 1 hour at 22º C. The mixture was extracted 3 times with Et_2O (3 × 5 mL). The final labeled ODNs were obtained after further HPLC purification and lyophilisation.

Table S2. Retention times (RT) of the DMT-on and DMT-off ODNs

Sequence	RT (min)	
	"DMT on" ODNs	"DMT off" ODNs
TMT : CGT TTT T(3HC)T TTT TGC	21.0	17.5
AMA: CGT TTT A(3HC)A TTT TGC	20.5	18.0
CMC: CGT TTT C(3HC) TTT TGC	21.7	12.2
GMG: CGT TTT G(3HC)G TTT TGC	22.0	13.9
MCG: 5'-(3HC)-CGT TTT TA TT TGC-3'	29.5	17.6

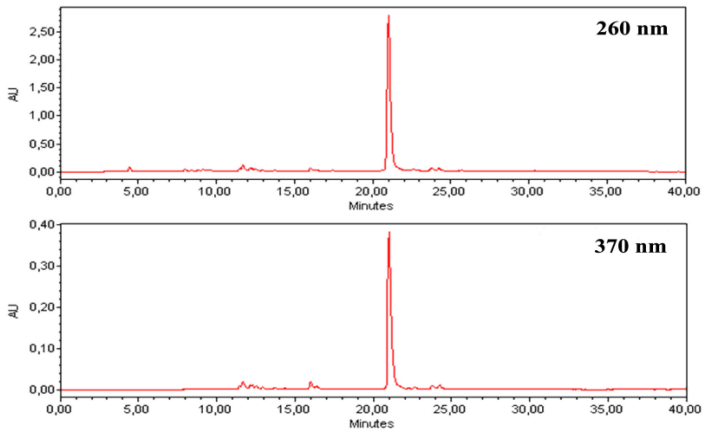


Figure S1. An example of HPLC profile of the crude DMT-on ODN: DMTr-**TMT**: DMTr-CGT-TTT-T(3HC)T-TTT-TGC. Absorption was monitored at 260 and 370 nm (absorptions of DNA bases and of the thiophenyl-3-hydroxychromone, respectively).

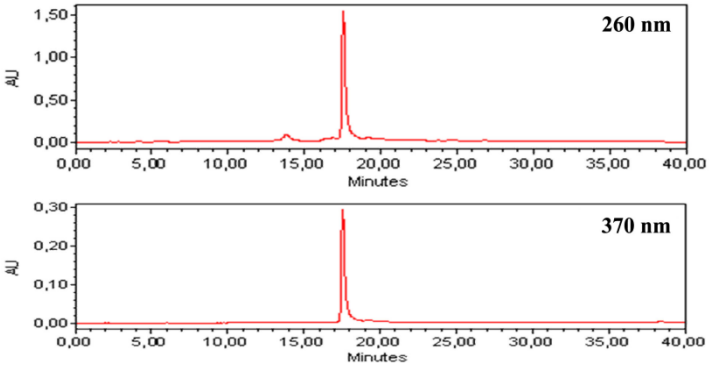


Figure S2. The HPLC profile of the oligonucleotide after cleavage of DMTr group. **TMT**: CGT-TTT-T(3HC)T-TTT-TGC Absorption was monitored at 260 and 370 nm (absorptions of DNA bases and of the thiophenyl-3-hydroxychromone, respectively).

2.2.6. MALDI – TOF spectra of ODNs.

A PerSeptive Voyager DE-STR MALDI-TOF mass spectrometer (PerSeptive Biosystems, Framingham, MA, USA), equipped with a 337-nm pulsed nitrogen laser (20 Hz) and an Acqiris® 2 GHz digitizer board, was used for all experiments. Mass spectra were obtained in the linear negative ion mode with the following settings: accelerating voltage 20kV, grid voltage 91 % of accelerating voltage, extraction delay time of 300 ns. The laser intensity was set just above the ion generation threshold to obtain peaks with the highest possible signal-to-noise (S/N) ratio without significant peak broadening. All data were processed using the Data Explorer software package (Applied Biosystems). 3-Hydroxypicolinic acid (3-HPA, used as the matrix for MALDI-TOF experiments, was of the highest grade available and used without further purification) was purchased from Sigma Aldrich Co.

Table S3. MS results of final ODNs

Sequence.	Formula of the compound	Calculated mass [M]	Experimental mass
TMT : CGT TTT T(3HC)T TTT TGC	$C_{156}H_{194}N_{36}O_{100}P_{14}S$	4639	4640 (pos)
AMA: CGT TTT A(3HC)A TTT TGC	$C_{156}H_{192}N_{42}O_{96}P_{14}S$	4657	4657 (pos)
CMC: CGT TTT C(3HC) TTT TGC	$C_{154}H_{192}N_{38}O_{98}P_{14}S$	4609	4609 (neg)
GMG: CGT TTT G(3HC)G TTT TGC	$C_{156}H_{192}N_{42}O_{98}P_{14}S$	4689	4690 (neg)
MCG: 5'-3HC CGT TTT TA TT TGC	$C_{166}H_{206}N_{41}O_{105}P_{15}S$	4952	4951 (neg)

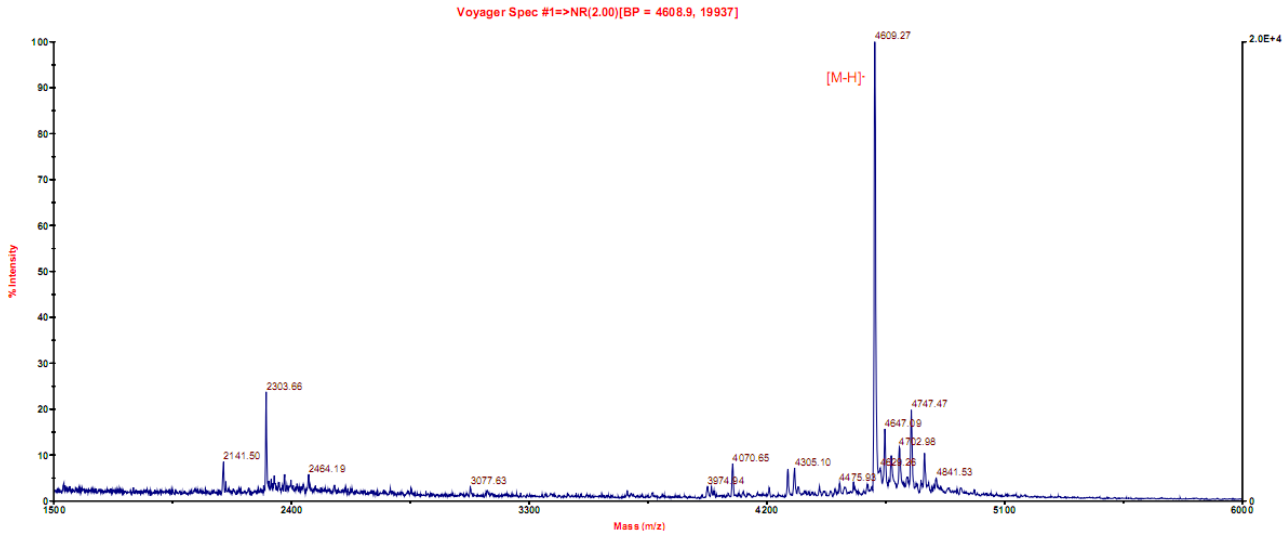


Figure S3. An example of a MALDI-TOF spectrum of the ODN **CMC**

2.2. Physical characterization

2.2.1. Preparation of the double stranded DNA samples.

Complementary DNA strands as well as reference natural strands were purchased from IBA GmbH. The absorption coefficient of ODN at 260 nm was used to calculate the concentration of the ODN solution. For unlabeled ODNs, the absorption coefficients provided by the supplier were used. The absorption coefficient of 3HC (ε_{260}= 10000 $M^{-1}cm^{-1}$) and the ODN calculator [3] were used to determine the absorption coefficients of labeled ODNs. The double stranded ODNs were prepared by mixing stock solutions of both ODN strands (32 µM of each strand) with 50 µL of buffer (10 mM cacodylate, 150 mM NaCl, pH 7), heating for 5 min at 80 ºC and cooling slowly down to room temperature. Then, the resulting solutions were diluted with buffer to a final volume of 800 µL to get the 2µM concentration needed for further experiments.

2.2.2. Thermal denaturation measurements.

Melting curves were recorded by following the temperature-dependence of the absorbance changes of the sample (2 µM concentration of each strand). Absorption spectra were recorded in a Peltier thermostated cell holder on a Cary 4 spectrophotometer (Varian). Wavelength for detection was 260 nm. The pathlength of cell was 1 cm. The temperature range for denaturation measurement was 15 – 80° C. Speed of heating was 0.5 °C/min. 10mM cacodylate buffer, 150mM NaCl, pH = 7.0 was used. The melting curves were converted into a plot of α versus temperature, where α represents the fraction of single-strands in the duplex state. The melting temperatures were extracted from these curves after differentiation as described elsewhere.[4]

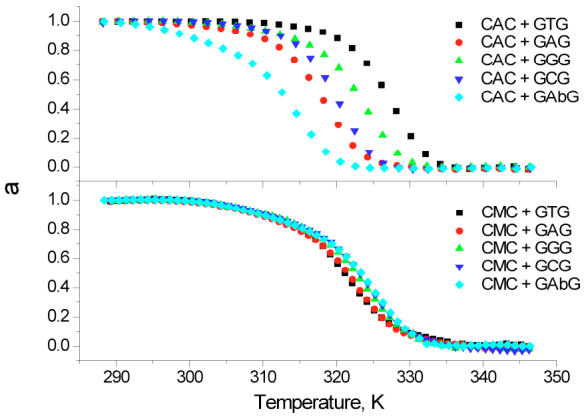


Figure S4. Examples of thermal denaturation curves expressed in fraction α of single strand versus temperature for control (top) and labeled (bottom).

2.2.3. Circular dichroism spectra.

Circular dichroism spectra were recorded with 2 µM solution of canonical dsDNA (control) and labeled dsDNA (3HC (**1**) opposite T or Ab) in buffer pH 7 (10mM cacodylate buffer, 150mM NaCl) at 25°C on a Jasco J-810 spectropolarimeter. Two maxima were observed in CD spectra: one negative one at 249 nm, the other positive at 282 nm.

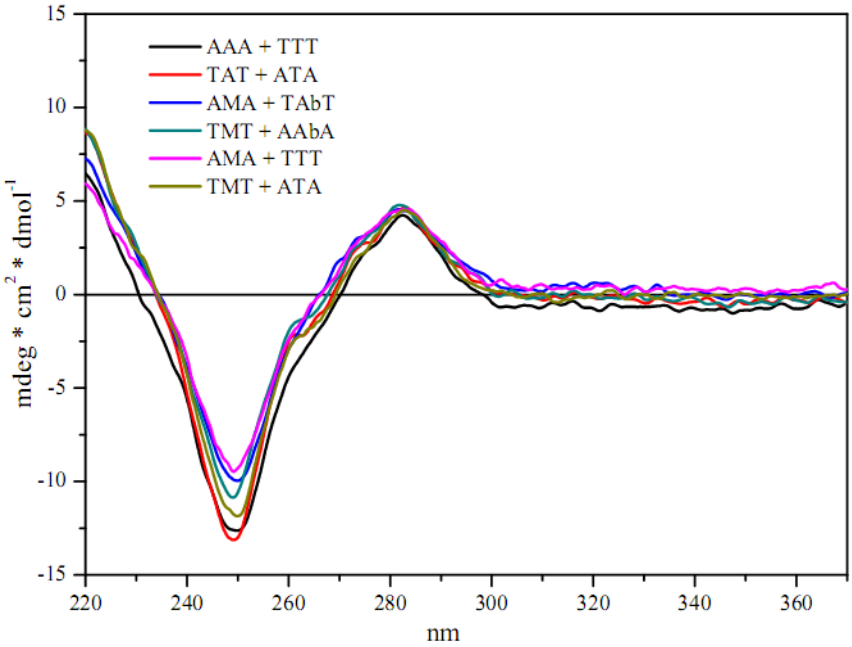


Figure S5. Representative CD spectra obtained for non-labeled (control) and labeled ODNs in buffer pH 7.

2.3. Spectroscopic studies.

Absorption and fluorescence spectra were recorded on a Cary 4 spectrophotometer (Varian) and FluoroMax 3.0 spectrofluorometer (Jobin Yvon, Horiba), respectively. Fluorescence emission spectra were systematically recorded at 20°C using an excitation wavelength of 360 nm. All the spectra were corrected from the fluorescence of the corresponding blank solution. Fluorescence quantum yields were determined by taking quinine sulfate in 0.5 M sulfuric acid (quantum yield, QY = 0.577) as a reference.

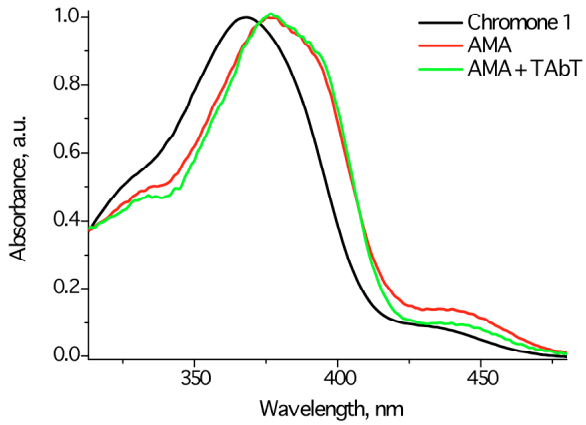


Figure S6. Representative UV spectra of the free chromone (**1**) in buffer pH 6.5 (10 mM phosphate) and of single and double strand forms of labeled ODNs in buffer pH 7 (10 mM cacodylate buffer, 150 mM NaCl).

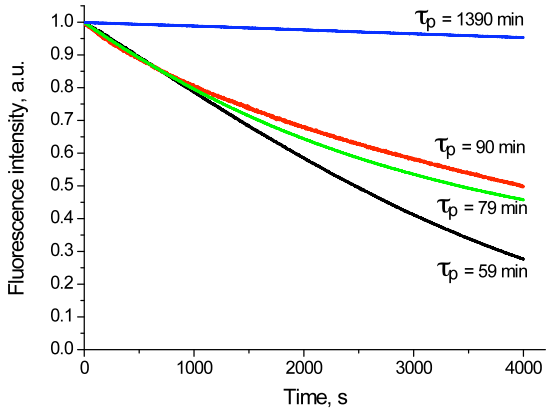


Figure S7. Photodegradation kinetics. Photodegradation of 2 µM solution of each sample was performed in a quartz micro-cuvette (50 µL volume) at 360 nm with the xenon lamp of the spectrofluorometer (slits were open to 8 nm) in buffer pH 7 (10 mM cacodylate, 150 mM NaCl) or EtOH. Black curve: Prodan in EtOH; red: nucleoside **1** in buffer; green: **TMT** in buffer; blue: **TMT+ATA** in buffer. During the time of illumination (4000 seconds), the fluorescence at the maximum (490 nm for Prodan and 540 nm for 3HC derivatives) was recorded as a function of time. τp is the time constant of photo-degradation (min) obtained by fitting the curves to a mono-exponential decay function.

For investigation of ODN interactions with viral peptides, NC and its mutant without zinc fingers (NC-SSHS) were used. In this mutant all cysteines were substituted with serines. NC and NC-SSHS were synthesized using solid phase peptide synthesis as described elsewhere.[5]

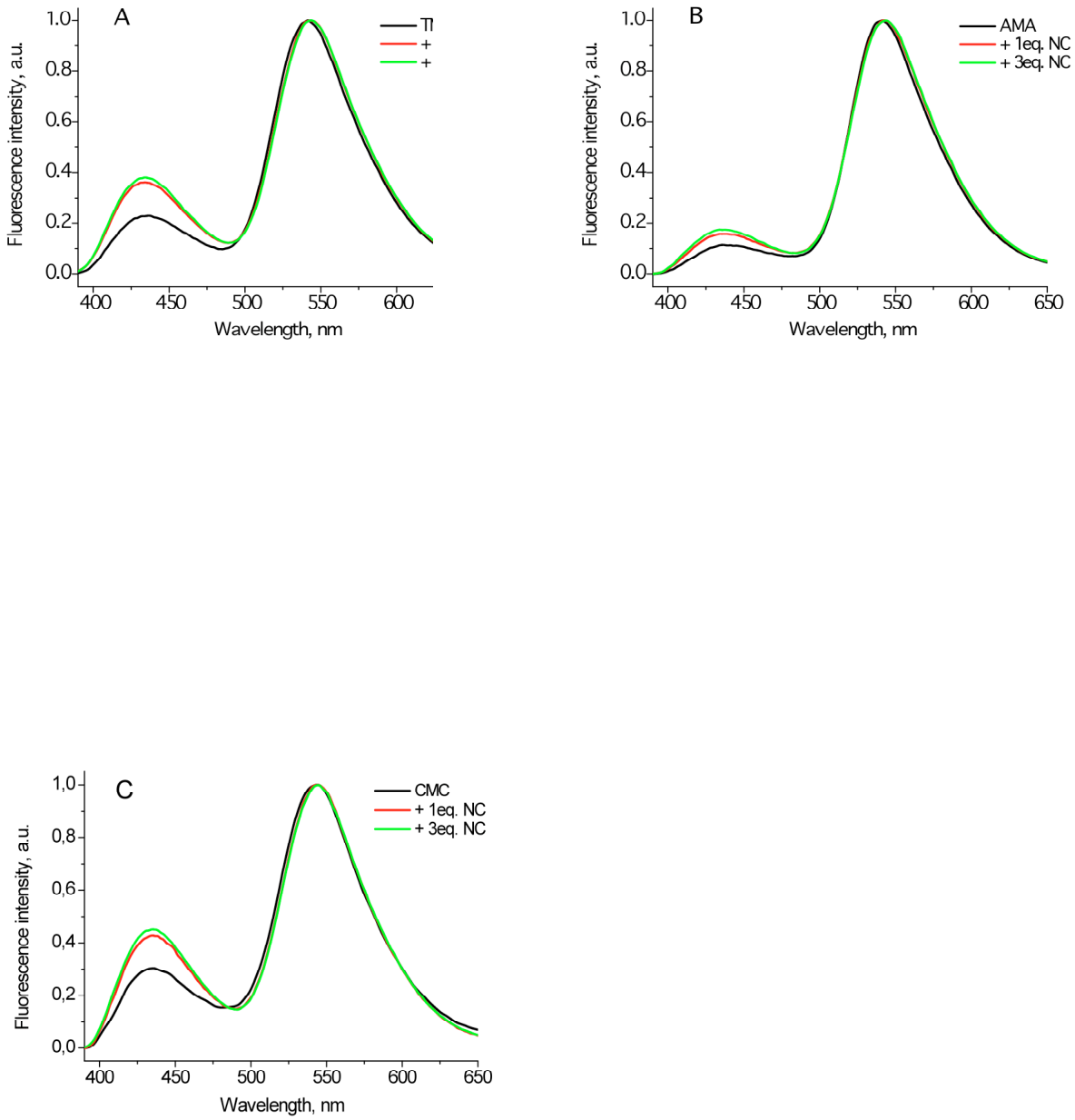


Figure S8. Fluorescence spectra of **TMT, AMA** and **CMC** ODNs in the absence (black) and presence of the 1eq. (red) and 3eq. (green) of NC peptide. The spectra were normalized at the T* band. The experiments were performed in 10 mM cacodylate buffer, 150 mM NaCl, pH 7.

3.0. References.

(1) a) Armarego, W. L. F.; Perrin, D. D. *Purification of Laboratory Chemicals*; 4th ed.; Butterworth Heinemann, 1996; b) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512−7515.

(2) a) Spadafora, M.; Postupalenko, V. Y.; Shvadchak, V. V.; Klymchenko, A. S.; Mely, Y.; Burger, A.; Benhida, R. *Tetrahedron* **2009**, *65*, 7809−7816; b) Barta, J.; Pohl, R.; Klepetarova, B.; Ernsting, N. P.; Hocek, M. *J. Org. Chem.* **2008**, *73*, 3798−3806; c) Dziuba, D.; Benhida, R.; Burger, A. *Synthesis* **2011**, 2159−2164.

(3) http://biophysics.idtdna.com/

(4) Breslauer, K. J. *Methods Enzymol.* **1995**, *259*, 221−242.

(5) V. V. Shvadchak, A. S. Klymchenko, H. de Rocquigny, Y. Mély, *Nucleic Acids Res.* **2009**, *37*, e25.

2.2. Investigation of peptide – membrane interactions by a 3-hydroxyflavone label

2.2.1. 3-Hydroxyflavone label for monitoring peptide – membrane interactions

The second part of my work consisted in developing an approach for monitoring membrane binding and insertion of peptides using a fluorescent environment-sensitive label of the 3-hydroxyflavone (3HF) family (**Publication 6**). For this purpose, we selected the 4'-(dimethylamino)-3-hydroxyflavone fluorophore which due to its excited state intramolecular proton transfer (ESIPT) and charge transfer presents a strong sensitivity of its dual emission to solvent polarity and hydration (Chou, Martinez et al. 1993; Klymchenko and Demchenko 2003). This fluorophore has already been used for characterizing the hydration and polarity of model lipid membranes (Klymchenko, Duportail et al. 2004; Klymchenko, Mely et al. 2004; Demchenko, Mely et al. 2009) and cell plasma membranes (Shynkar, Klymchenko et al. 2005; Shynkar, Klymchenko et al. 2007).

In this work, a functionalized fluorescent label (MFL) based on 4'-(dimethylamino)-3-hydroxyflavone (Figure 2.15) was synthesized in three steps starting from 5-N-acetylamino-2-hydroxyacetophenone. At the final step, the 6-amino-4'-(dimethylamino)-3-hydroxyflavone was reacted with succinic anhydride affording the MFL label. In aprotic environment, the probe showed a dual emission, with a short- and a long-wavelength band that could be unambiguously assigned to the normal (N*) and ESIPT tautomer (T*) forms (Figure 2.15). In protic environment, MFL exhibited a single red-shifted band which corresponds to the emission of the H-bonded N* form (H-N*). In this form, a strong H-bond between the 4-carbonyl group of the flavone and the protic solvent is thought to inhibit the ESIPT reaction.

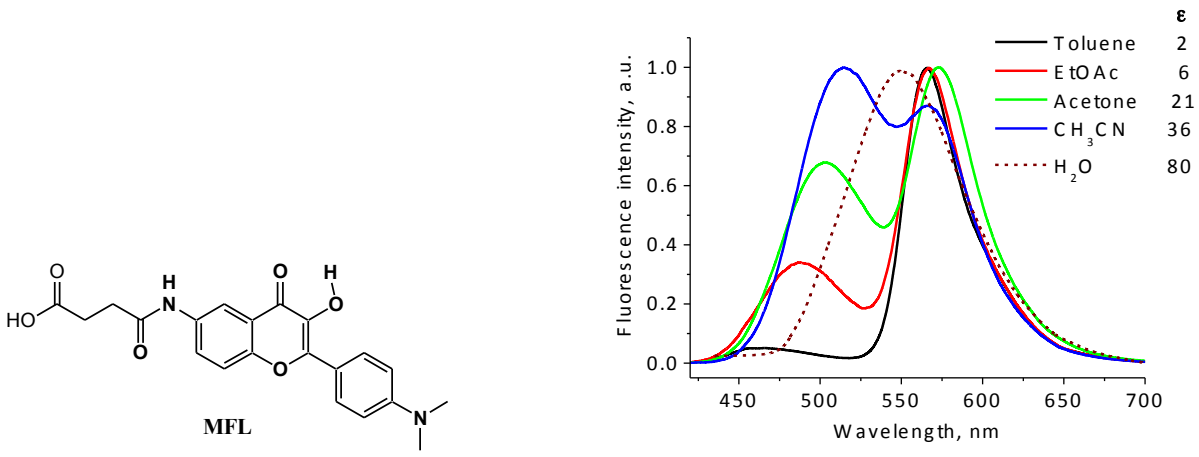


Figure 2.15. Structure and spectra of the MFL label in different solvents.

For testing the ability of MFL to sense peptide-membrane interactions, three model peptides labeled by MFL at their N-terminus were synthesized using solid-phase peptide synthesis: melittin, magainin 2 and poly-L-lysine. These peptides present different numbers of basic and hydrophobic amino acids, and bind differently to lipid membranes. In buffer, all three labeled peptides showed a

single red-shifted emission band similar to that of the free label, though the fluorescence quantum yield is significantly higher with the peptides. We speculate that the proximal peptide backbone could partially screen the probe from water and, thus reduce its quenching by water.

Upon addition of neutral vesicles, the fluorescence of the labeled melittin and magainin 2 was changed dramatically while the fluorescence of poly-L-lysine remained the same as in buffer, suggesting that only the two first peptides can interact with neutral lipid bilayers. For both melittin and magainin 2, a large increase in the fluorescence quantum yield accompanied by the appearance of the two-band emission characteristic for hydrophobic environment was observed. Since the fluorescence of the free peptide differs significantly from that of the membrane-bound peptide, the binding stoichiometry was determined by a simple titration (addition of the labeled peptide to vesicles or vice-versa). We observed that the fluorescence intensity reaches saturation at a ratio of 1 peptide per 100 lipids. Using this lipid-peptide binding stoichiometry, the titration curve allowed us to evaluate the binding constant of melittin to the DOPC membrane ($K = 4 \pm 2 \times 10^7\ M^{-1}$). As a next step, we determined the appropriate lipid composition required for efficient binding of poly-L-lysine to lipid membranes. For this purpose, the labeled poly-L-lysine was studied in the presence of vesicles containing mixtures of neutral lipid with different molar fractions of negatively charged lipids. It was found that the critical concentration of negatively charged lipids required for binding of poly-L-lysine is 20% which was used in the further experiments.

The high sensitivity of MFL to the polarity and hydration of the environment was then used for evaluating the insertion depth of the peptide N-terminus to which the probe is covalently bound. From comparison of the dual emission of the various peptides bound to lipid vesicles, it can be observed that the relative intensity of the short-wavelength band is much larger for poly-L-lysine than for the two other peptides (Figure 2.16), indicating a significantly larger polarity and/or hydration for poly-L-lysine. To further quantify the hydration and polarity of the label environment in these peptides bound to DOPC/DOPS vesicles, the two-band emission spectrum was deconvolved into three bands, with the N* and T* bands corresponding to the non-hydrated form of the dye and the H-N* band corresponding to the hydrated form (Figure 2.16). The H-N* form is usually localized at the hydrated membrane interface and its emission band is strongly red-shifted with respect to the N* band. From the deconvolved spectra, the "polarity" parameter was evaluated from the I_{N*}/I_{T*} ratio, showing that the polarity parameter was much larger for poly-L-lysine than for the two other peptides. Next, we used the previously established linear dependence of the I_{N*}/I_{T*} ratio of 3HF dye in aprotic solvents versus the dielectric constant function (Klymchenko and Demchenko 2003) to estimate the environment polarity of the non-hydrated form of the label for the three labeled peptides bound to the membranes. We found that for melittin and magainin 2, the observed I_{N*}/I_{T*} ratios correspond to a dielectric constant ε~10-11, while for poly-L-lysine – to $\varepsilon \sim 39$, which correspond to depths of ~10 and ~ 20 Å from the bilayer center, respectively (from the dielectric model of lipid bilayer suggested by Griffith et al (Griffith, Dehlinger et al. 1974)).

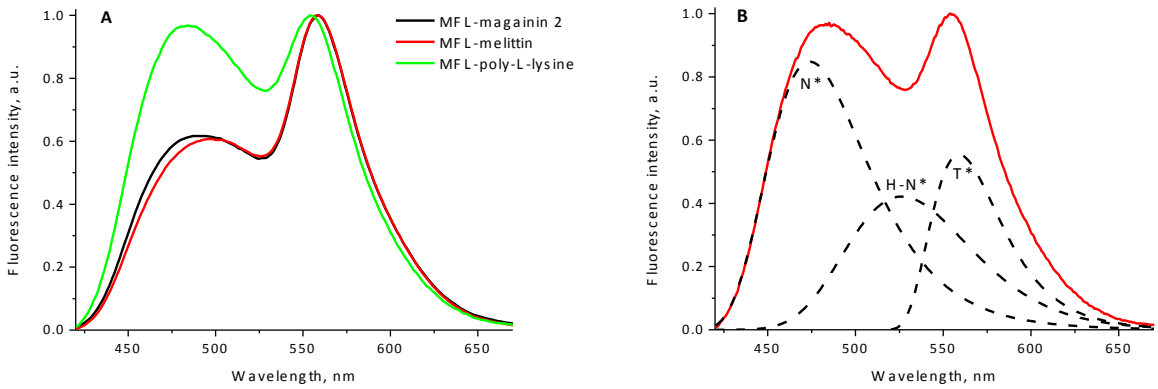


Figure 2.16. (A) Normalized fluorescence spectra of labeled peptides bound to DOPC:DOPS vesicles. (B) Emission spectrum of poly-L-lysine (solid red line) and its deconvolution into N*, H-N*, and T* spectral components.

To confirm the localization of the labeled peptides bound to the bilayer, the parallax quenching method using spin-labeled lipids was performed. Three quenchers of deep (5.85 Å from the bilayer center), middle (12.15 Å) and shallow location (19.5 Å) were used. Parallax experiments showed that for melittin and magainin 2, the deep quencher was the most efficient, while for poly-L-lysine the quenching was the most efficient with the shallow quencher. The quenching of the hydrated and non-hydrated forms of the label was analyzed independently after deconvolution of the corresponding emission spectra. The non-hydrated form of the label for melittin and magainin 2 bound to DOPC/DOPS vesicles was found to be located at 8.0-8.1 Å from the bilayer center, in the region of the fatty acid chains (Figure 2.17), in line with our estimations based on the relatively low dielectric constant of its environment. The location of the hydrated form of the label for both proteins was 16.4 Å, indicating a shallow position next to the head groups (Figure 2.17). This large difference in the locations of the two forms of the label could be explained by the relative freedom of the label, due to the presence of the flexible linker (Figure 2.17). Another possible explanation is that melittin and magainin 2 present both tilted and surface location in the lipid bilayer, thus producing the observed heterogeneity in depths. Remarkably, for poly-L-lysine, both hydrated and non-hydrated forms present a shallow location close to the head groups, around 16.5 Å from the bilayer center (Figure 2.17), in line with the high dielectric constant of its environment.

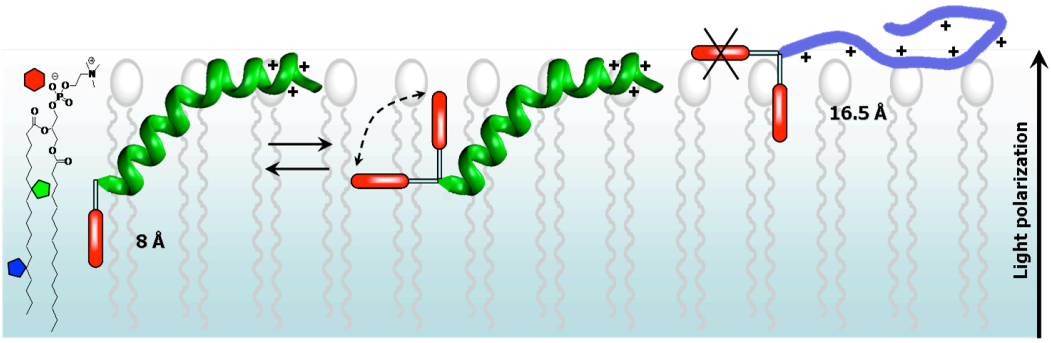


Figure 2.17. Localization of the N-terminus of melittin (green) and poly-L-lysine (blue) labeled by MFL probe and bound to DOPC:DOPS (80:20) membrane. The positions of the quenching residues of the deep (blue), middle (green) and shallow (red) quenchers are pointed by the colored polygons. In the case of melittin, two locations of the label are shown: deep vertical (non-hydrated) and shallow with a distribution of orientations (hydrated).

Moreover, the orientation of the label in the lipid bilayer was evaluated. Labeled melittin and poly-L-lysine were added to giant vesicles (GUVs) composed of DOPC/DOPS and studied by two-photon microscopy using the polarized laser excitation. We recorded the images in the blue and red regions, which were then used to calculate the ratio images. With the labeled melittin, the poles and equator appeared in different colors corresponding to lower and higher I_{blue}/I_{red} ratios, respectively suggesting two populations for the label in the bilayer. One population is associated with a low I_{blue}/I_{red} ratio and orients vertically in the bilayer, while the second population is associated with a high I_{blue}/I_{red} ratio and does not exhibit any preferential orientation. This conclusion is in line with the parallax data, which also suggest two populations for the melittin label: a deep non-hydrated population and a shallow hydrated population. In contrast, a vertical orientation of the fluorophore and no heterogeneity in the probe orientation were observed for the labeled poly-L-lysine.

Thus, this label constitutes an interesting new tool for monitoring membrane binding and insertion of peptides.

Publication 6

Monitoring membrane binding and insertion of peptides by two-color fluorescent label

Biochimica et Biophysica Acta 1808 (2011) 424–432

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Monitoring membrane binding and insertion of peptides by two-color fluorescent label

V.Y. Postupalenko [a], V.V. Shvadchak [a], G. Duportail [a], V.G. Pivovarenko [b], A.S. Klymchenko [a], Y. Mély [a,*]

[a] Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de Pharmacie, 74, Route du Rhin, 67401 ILLKIRCH Cedex, France
[b] Department of Chemistry, Kyiv National Taras Shevchenko University, 01033 Kyiv, Ukraine

ARTICLE INFO

Article history:
Received 22 July 2010
Received in revised form 16 September 2010
Accepted 20 September 2010
Available online 13 October 2010

Keywords:
Fluorescence
3-hydroxyflavone
Lipid membrane
Melittin
Peptide–membrane interaction

ABSTRACT

Herein, we developed an approach for monitoring membrane binding and insertion of peptides using a fluorescent environment-sensitive label of the 3-hydroxyflavone family. For this purpose, we labeled the N-terminus of three synthetic peptides, melittin, magainin 2 and poly-L-lysine capable to interact with lipid membranes. Binding of these peptides to lipid vesicles induced a strong fluorescence increase, which enabled to quantify the peptide–membrane interaction. Moreover, the dual emission of the label in these peptides correlated well with the depth of its insertion measured by the parallax quenching method. Thus, in melittin and magainin 2, which show deep insertion of their N-terminus, the label presented a dual emission corresponding to a low polar environment, while the environment of the poly-L-lysine N-terminus was rather polar, consistent with its location close to the bilayer surface. Using spectral deconvolution to distinguish the non-hydrated label species from the hydrated ones and two photon fluorescence microscopy to determine the probe orientation in giant vesicles, we found that the non-hydrated species were vertically oriented in the bilayer and constituted the best indicators for evaluating the depth of the peptide N-terminus in membranes. Thus, this label constitutes an interesting new tool for monitoring membrane binding and insertion of peptides.

1. Introduction

Interactions of membrane proteins, peptide toxins and delivery vectors with lipid bilayers play a fundamental role in their activity. However, the precise insertion of the proteins in the bilayer, which is a key element for characterizing the mechanism of these proteins, is difficult to determine experimentally. Fluorescence and notably, fluorescence resonance energy transfer (FRET) is one of the most well-established technique to monitor protein–membrane interactions, through the proximity between fluorescently labeled lipids and peptide. However, the FRET method can hardly be used for precisely localizing a protein/peptide in the membrane, since the FRET signal depends not only on the precise positioning of the FRET partners, but also on their distribution in the lipid bilayer [1]. To localize peptides in membranes, the parallax fluorescence quenching is commonly used. With this method, the Trp residues of a peptide/protein in the bilayer are localized using spin-labeled [2,3] or brominated lipids [4] bearing a quencher at a precise depth. This approach though being universal and rather precise requires a large quantity of quencher (ca 20 mol%) and is limited mainly to model lipid bilayers. Another method that becomes more and more attractive for studying protein insertion is based on environment-sensitive (polarity-sensitive) fluorescent labels. These labels being covalently attached to a peptide chain can change their emission color as a function of the insertion depth (Fig. 1), since the environment polarity varies steeply from the top of the bilayer (dielectric constant $\varepsilon = 80$) to the middle of the bilayer ($\varepsilon \sim 2$). Trp residues, being environment-sensitive fluorophore, can also provide some information about the protein insertion [5,6], but this approach is limited to proteins containing a single Trp residue. Moreover, the fluorescence characteristics and environment-sensitivity of Trp residues show some limitations. Acrylodan and Badan, two reactive derivatives of the common environment-sensitive dye Prodan [7], are among the most popular labels of proteins and peptides used for studying the interactions with lipid membranes. For instance, Acrylodan was used to map the insertion of equinatoxin II into membranes. When labeled mutants of this protein were added to lipid vesicles, only some of them exhibited blue-shifts of the Acrylodan fluorescence, due to insertion of the corresponding peptide labeling site into the hydrophobic membrane environment. These data allowed identification of two protein domains embedded within the lipid membrane [8]. In another study, Acrylodan was used to monitor the insertion of the N-terminal domain of Annexin 2 into lipid bilayers. This insertion was evidenced by blue shifts in its emission and its fluorescence quenching by doxyl-labeled phospholipids [9]. Finally, the Badan label was successfully used to monitor the insertion and orientation of model transmembrane proteins in lipid

* Corresponding author. Tel.: +33 368 854263; fax: +33 368 854313.
E-mail address: yves.mely@unistra.fr (Y. Mély).

0005-2736/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamem.2010.09.013

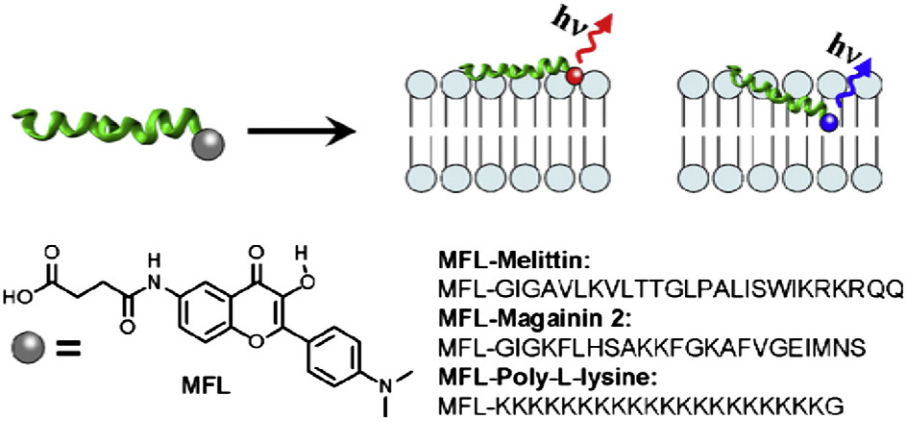


Fig. 1. Principle of the monitoring of the interaction of MFL-labeled peptides with the lipid membranes. The peptide (in green) is labeled with the fluorescent MFL label. The interaction of the peptide with lipid membranes changes the fluorescence intensity and color of the label. The extent of these changes depends on the peptide insertion. The chemical structure of the MFL label and its conjugates with peptides are also shown.

membranes, providing complementary information with respect to NMR data and molecular-dynamics simulations [10].

In our studies, we selected a 3-hydroxyflavone fluorophore, 4′-(dimethylamino)-3-hydroxyflavone, which due to excited–state intramolecular proton transfer (ESIPT) and charge transfer presents a strong sensitivity of its dual emission to solvent polarity and hydration [11,12]. This fluorophore has already been used for characterizing the hydration and polarity of model lipid membranes [13–15] and cell plasma membranes [16,17]. In this respect, it was tempting to apply this fluorophore for monitoring interactions of peptides with lipid membranes. As model membrane binding peptides, we selected melittin, magainin 2 and poly-L-lysine. The first two are membranolytic toxins, which bind neutral and charged membranes and insert deeply into the bilayer, producing pores [18–20]. For both melittin and magainin 2, it was reported that their C-terminus is located at the membrane surface, while their N-terminus is inserted rather deeply into the bilayer [21–26]. In contrast, the non viral gene delivery vector, poly-L-lysine [27] binds mainly negatively charged membranes and localizes at the surface next to polar head groups [28]. In the present work, a functionalized fluorescent label, MFL, based on 4′-(dimethylamino)-3-hydraxy-flavone was synthesized and attached to the N-terminus of melittin, magainin 2 and poly-L-lysine (Fig. 1). Studies of the interaction of the labeled peptides with lipid vesicles revealed that peptide–membrane interactions result in a dramatic increase in the fluorescence intensity of the label. All peptides bound to lipid membranes showed a dual emission, which correlated with the depth of the label insertion, as evidenced from parallax quenching measurements. We confirmed that while the N-terminus of melittin and magainin 2 present a rather deep insertion in the bilayer, the poly-L-lysine N-terminus is localized at the interface. Thus, in this work we present a new methodology for monitoring insertion of peptides into lipid bilayers.

2. Materials and methods

All chemicals and solvents for synthesis and spectroscopic measurements were from Sigma-Aldrich. Dioleoylphosphatidylcholine (DOPC), dioleoylphosphatidylserine (DOPS) were from Sigma-Aldrich. 1,2-dipalmitoyl-*sn*-glycero-3-phospho(tempo)choline (TempoPC), 1-palmitoyl-2-stearoyl-(5- and 12-doxyl)- *sn*-glycero-3-phosphocholine (5- and 12-SLPC) were from Avanti Polar Lipids (Alabaster, AL, USA). The concentration of phospholipid stock solutions in chloroform was determined by dry weight. The nitroxide content of nitroxide-labeled lipids was calculated using the electron spin resonance integrated spectra of the corresponding diluted stock solutions in chloroform by comparing with a tempocholine reference solution in the same solvent.

2.1. Synthesis of MFL label

5′-Acetamido-2′-hydroxyacetophenone (15.5 mmol, 3 g) and 4′-dimethylaminobenzaldehyde (17.1 mmol, 2.55 g) were dissolved in 20 mL of dry DMF. Finally, sodium methoxide (92.6 mmol, 5 g) was added in two portions. The mixture was stirred for 24 h at 60 °C. Then, the reaction mixture was diluted with 100 mL of ethanol followed by addition of 10 mol excess of hydrogen peroxide (16 mL) and 15 mol excess of sodium methoxide (14 g). The reaction mixture was refluxed for 10 min. After cooling, the mixture was poured into water and neutralized with conc. HCl to pH = 6–7. The formed precipitate was filtered off, washed with water, and dried under vacuum. The yellow solid was purified by crystallization from methanol/ethanol mixture to give 6-acetamido-4′-(dimethylamino)-3-hydroxyflavone (2.8 g, 54 %) as yellow crystals. ^{1}H NMR (400 MHz, DMSO-d6) δ 10.22 (s, 1 H), 9.12 (s, 1 H), 8.39 (d, 1 H, J = 2.51 Hz), 8.11 (d, 2 H, J = 9.03 Hz), 7.87 (dd, 1 H, J_1 = 9.03 Hz, J_2 = 2.51 Hz), 7,68 (d, 1 H, J = 9.03 Hz), 6.85 (d, 2 H, J = 9.03 Hz), 3.02 (s, 6 H) , 2.10 (s, 3 H); *m/z* ($M+H^+$) calculated for $C_{19}H_{19}N_2O_4$: 339.1; found: 339.0.

6-acetamido-4′-(dimethylamino)-3-hydroxyflavone (3 mmol, 1 g) was refluxed in 10% HCl (30 mmol, 11 mL) for 7 h. Then, the reaction mixture was concentrated in vacuo and dissolved in 15 mL dry THF. Then, 0.3 g (3 mmol) of succinic anhydride and 2.6 mL of diisopropylethylamine (15 mmol) were added to this solution. The reaction mixture was stirred overnight at 60 °C. Then the solvent was evaporated and the residue was treated with water. The precipitate was filtered off and washed with water. The product was recrystallized from propanol-2 to give 1 g (yield 84%) of the final acid. ^{1}H NMR (300 MHz, DMSO-d6) δ 12.5–12.0 (1 H, br s), 10.24 (s, 1 H), 9.12 (s, 1 H), 8.40 (d, 1 H, J = 2.64 Hz), 8.09 (d, 2 H, J = 9.04 Hz), 7.84 (dd, 1 H, J_1 = 9.04 Hz, J_2 = 2.64 Hz), 7,67 (d, 1 H, J = 9.04 Hz), 6.84 (d, 2 H, J = 9.04 Hz), 3.01 (s, 6 H) , 2.65–2.50 (m, 4 H); *m/z* ($M+H^+$) calculated for $C_{21}H_{21}N_2O_6$: 397.1; found: 397.1.

2.2. Peptide synthesis

Peptides were synthesized by solid phase peptide synthesis on a 433A synthesizer (ABI, Foster City, CA). The synthesis was performed at a 0.1-mmol scale using standard side-chain protected fluorenylmethoxycarbonyl

(Fmoc)-amino acids and HBTU/HOBt coupling protocol. LL-HMP resin (ABI, 0.44 mmol/g reactive group concentrations) was used as solid support. At the end of the synthesis, peptidylresin was isolated and washed twice by NMP.

Two equivalents (0.15 mmol) of the label (MFL) were dissolved in 1 mL of NMP mixed with two eq. of HBTU/HOBt coupling solution (in DMF) and added to Fmoc-deprotected peptidylresin (0.075 mmol) swelled in 1 mL of NMP. After a few minutes of shaking, five eq. of DIEA solution was added. Then, the reaction mixture was stirred overnight at 40 °C. Resin was filtrated and washed by NMP, methanol and DCM.

Cleavage and deprotection of the peptidylresin were performed for 2 h using a 10 mL trifluoroacetic acid (TFA) solution containing water (5%, v/v), TIS (iPr_3SiH, 2.5%, v/v), phenol (1%, w/v), thioanisole (5%, v/v) and ethanedithiol (2.5%, v/v). The solution was concentrated in vacuo and the peptide was precipitated by using ice-cold diethyl ether and then pelleted by centrifugation. The pellet was washed with diethyl ether and dried. The peptides were solubilized with aqueous TFA (0.05 %, v/v). HPLC purification was carried out on a C8 column (uptisphere 300A, 5 μm; 250X10, Interchim, France) in water/acetonitrile mixture containing 0.05% TFA with linear gradients depending on the peptide (typically 10 to 60% of acetonitrile for 30 min) and monitored at 210 nm (detection of all peptides including non-labeled) and 370 nm (detection of labeled peptides only). Molecular masses obtained by ion spray mass spectrometry were: 2844, 3225 and 3017 for MFL-magainin 2, MFL-melittin and MFL-poly-L-lysine, respectively, in agreement with the expected theoretical masses. Prior to use, peptides were dissolved in distilled water, aliquoted and stored at −20 °C. Concentrations of the labeled peptides were determined from the label absorbance at 400 nm using $\varepsilon = 33000\ M^{-1}\ cm^{-1}$.

2.3. Sample preparation

Large unilamellar vesicles (LUVs) were obtained by the classical extrusion method [29] or by ethanol dilution [30]. In the first method, a suspension of multilamellar vesicles was extruded by using a Lipex Biomembranes extruder (Vancouver, Canada). The size of the filters was first 0.2 μm (7 passages) and thereafter 0.1 μm (10 passages). This protocol leads to monodisperse LUVs with a mean diameter of 0.11 μm as measured with a Malvern Zetamaster 300 (Malvern, UK).

Giant unilamellar vesicles (GUVs) were generated by electroformation in a home-built liquid cell (University of Odense, Denmark), using previously described procedures [31–33]. 1 mM solution of lipids in chloroform was deposited on the platinum wires of the chamber, and the solvent was evaporated under vacuum for 30 min. The chamber was filled with a 300 mM sucrose solution, and a 2-V, 10-Hz alternating electric current was applied to this capacitor-like configuration for ca. 1.5 h. Then, a 50 μL aliquot of the obtained stock solution of GUVs in sucrose (cooled down to room temperature) was added to 200 μL of 300 mM glucose solution to give the final suspension of GUVs used in microscopy experiments. The staining of GUVs was performed by addition of an aliquot of the peptide solution to obtain a 0.05 μM final concentration.

All measurements were done in 20 mM phosphate buffer containing 150 mM NaCl (pH = 7.4) at 20 °C. To incorporate peptides into the membranes, an aliquot of a peptide stock solution in distilled water was added to a suspension of lipid vesicles. Then the measurements were performed after 5 min of incubation.

2.4. Parallax Quenching Method

The fluorescence intensity of labelled vesicles, either DOPC or DOPC with 15% nitroxide lipids, was measured in a 1-cm semi-micro quartz cuvette. Using the corrected F/F_0 values, the distance of the fluorophores from the center of the bilayer was calculated using the parallax equation originally proposed by London and collaborators [34–36]:

$$Z_{cf} = L_{cl} + \left[-\ln(F_1 / F_2) / \pi C - L_{21}^2\right] / 2L_{21} \quad (1)$$

where Z_{cf} is the distance of the fluorophore from the center of the bilayer; F_1 and F_2 are the fluorescence intensities in the presence of the shallow quencher (quencher 1) or the deeper quencher (quencher 2), respectively; L_{cl} is the distance of the shallow quencher from the center of the bilayer, L_{21} is the distance between the shallow and deep quenchers, and C the concentration of quencher in molecules/Å^2 (equals the mole fraction of nitroxide-labeled phospholipid divided by area per phospholipid; presently $C = 0.15/70$ Å^2 [37]). The quenching by the two most efficient quenchers (TempoPC/5-SLPC or 5-SLPC/12-SLPC) is used to calculate Z_{cf} [34–36]. The values used for the distances of the nitroxide group from the bilayer center were 5.85 Å for 12-SLPC, 12.15 Å for 5-SLPC, and 19.5 Å for TempoPC [34–36].

For quenching experiments by the parallax method, lipid vesicles were prepared according to Kachel et al. [30] with small modifications. DOPC (85%) and nitroxide-labeled PCs (15%) were mixed in chloroform in order to obtain a final concentration of 400 μM. The mixtures were dried under N_2 and kept under vacuum for 30 min, and then resuspended in ethanol (120 μL) by continuous rotation using a rotary evaporator for 30 min. Finally, 6 mL of buffer (20 mM phosphate, 150 mM NaCl, pH=7.4) was added and vortexed briefly. The sizes of vesicles were determined by light scattering using a N4SD Coultronics Nanosizer. The vesicles obtained by the ethanol dilution method were of similar size as those obtained by extrusion (0.12 μm). Concentrations of the labeled peptides and lipids were 1 μM (in case of poly-L-lysine −0.05 μM) and 400 μM, respectively.

2.5. Deconvolution of spectra

Deconvolution of the fluorescence spectra of the MFL-labeled peptides into three bands, corresponding to the normal (N*), H-bonded normal (H-N*) and tautomer (T*) forms, was performed using the Siano software kindly provided by Dr. A.O. Doroshenko (Kharkov, Ukraine), as previously described [14,15]. The program is based on an iterative nonlinear least-squares method, where the individual emission bands were approximated by a log-normal function accounting for several parameters: maximal amplitude, I_{max}, spectral maximum position, ν_{max}, and position of half-maximum amplitudes, ν_1 and ν_2, for the blue and red parts of the band, respectively. These parameters determine the shape parameters of the log-normal function, namely the full width at the half-maximum, $FWHM = \nu_1 - \nu_2$, and the band asymmetry, $P = (\nu_1 - \nu_{max})/(\nu_{max} - \nu_2)$. For the iteration process, the FWHM of the two short-wavelength bands (N* and H–N*) were fixed at 3000 cm^{-1}. For the H-N* band, the asymmetry and the band position were fixed at 0.9 and 19000 cm^{-1}, respectively. The other parameters, i.e. asymmetry of N* and T* bands, the band width of the T* band and the relative intensities of the bands, were allowed to vary in the iteration process. The resulting fluorescence intensities of the separated N*, H-N* and T* bands (I_{N^*}, $I_{H\text{-}N^*}$ and I_{T^*}) were used for calculation of the hydration parameter, which was expressed as the ratio of the peak emission intensity of the hydrated (H-N*) form to the summed intensities of the non-hydrated (N* and T*) forms. Taking into account that the FWHM for the T* band is ca 2-fold narrower than for the N* and H-N* bands, the hydration was estimated as $I_{H\text{-}N^*}/(I_{N^*} + 0.5 \times I_{T^*})$. The "polarity" parameter was expressed as the I_{N^*}/I_{T^*} ratio [14,15].

2.6. Instrumentation

Proton NMR spectra were recorded on a Bruker spectrometer and mass spectra on a LC/MSD SL Agilent Technologies mass spectrometer using the electrospray ionization (ESI) method. Absorption spectra

were measured on a Cary 4 spectrophotometer (Varian) and fluorescence spectra on a FluoroLog (Jobin Yvon, Horiba) spectrofluorometer. Fluorescence emission spectra were recorded at 400 nm excitation wavelength. All the spectra were corrected for Raman scattering and background fluorescence measured before addition of the labeled peptide. Fluorescence quantum yields were determined by taking 4′-(dialkylamino)-3-hydroxyflavone in ethanol (quantum yield, QY = 0.51) as a reference [12].

To calculate the vesicle concentration, the external radius of the vesicles (R) was considered to be 535 Å, as determined by DLS measurement. The thickness of the lipid bilayer (t) and the average lipid density (d) were assumed to be 40 Å and 70 Å^2/lipid, respectively [38]. The number of lipids per vesicle was thus determined as $n = 4\pi^*(R^2 + (R-t)^2)/d = 9.54\times10^4$ lipids /vesicle. The vesicles concentration can thus be obtained by using: $C\ (vesicles) = C\ (lipids)/n$.

To determine the affinity of MFL-melittin for neutral (DOPC) vesicles, a fixed amount of the peptide was titrated with lipids by monitoring the two-band fluorescence of MFL. The fluorescence intensity versus vesicle concentration was plotted and the affinity constant was determined from direct fitting of the curve by the following equation [39]:

$$I = I_0 - \frac{(I_0 - I_t)}{P_t} \times \frac{(1 + (P_t + nN_t)K_a) - \sqrt{(1 + (P_t + nN_t)K_a)^2 - 4P_t nN_t K_a^2}}{2K_a} \quad (2)$$

where I and I_t are the integrated intensity of the whole emission spectrum at a given and a saturating vesicles concentration, respectively, I_0 is the corresponding intensity in the absence of vesicles. N_t is the total vesicles concentration, P_t is the total concentration of peptide, K_a is the apparent affinity constant, and n is the number of peptides per vesicle.

Fluorescence microscopy experiments were performed by using a home-built two-photon laser scanning setup based on an Olympus IX70 inverted microscope with an Olympus 60× 1.2NA water immersion objective [40,41]. Two-photon excitation was provided by a titanium-sapphire laser (Tsunami, Spectra Physics), and photons were detected with Avalanche Photodiodes (APD SPCM-AQR-14-FC, Perkin-Elmer) connected to a counter/timer PCI board (PCI6602, National Instrument). Imaging was carried out using two fast galvo-mirrors in the descanned fluorescence collection mode. Typical acquisition time was 5 s with an excitation power around 2.5 mW (830 nm) at the sample. Images corresponding to the blue and red channels were recorded simultaneously using a dichroic mirror (Beamsplitter 585 DCXR) and two APDs. The images were processed with a home-made program under LabView that generates a ratiometric image by dividing the image of the blue channel by that of the red channel. For each pixel, a pseudocolor scale is used for coding the ratio, while the intensity is defined by the integrated intensity recorded for both channels at the corresponding pixel [33].

3. Results and discussion

3.1. Label design and characterization

The MFL label (Fig. 1) was designed on the basis of 4′-(dimethylamino)-3-hydroxyflavone, a dye displaying high sensitivity to polarity in low polar media [11,42]. It was synthesized in three steps starting from 5-N-acetylamino-2-hydroxyacetophenone. At the final step, the 6-amino-4′-(dimethylamino)-3-hydroxyflavone was reacted with succinic anhydride affording the MFL label. In organic solvents, the MFL label showed absorption and fluorescence properties similar to its parent analogue F. In aprotic media, the probe presented a dual emission, with a short- and a long-wavelength band that could be unambiguously assigned to the normal (N*) and ESIPT tautomer (T*) forms (Fig. 2). While increasing the solvent polarity, the N* band shifted to the red and its relative intensity, I_{N^*}/I_{T^*}, increased (Fig. 2, Table 1), in line with our previous studies [11]. In water and other protic media, MFL exhibited a single red-shifted band, corresponding to the emission of the H-bonded N* form (H-N*). In this form, a strong H-bond between the 4-carbonyl group of the flavone and the protic solvent is thought to inhibit the ESIPT reaction [43,44]. Importantly, the fluorescence quantum yield of the label is very low in neat water (Table 1), but not in other protic media like alcohols.

3.2. Labeled peptides and peptide–membrane interactions

Using solid-phase peptide synthesis, three peptides labeled by MFL at their N-terminus were obtained: melittin, magainin 2 and poly-L-lysine, which contain 26, 23 and 21 amino acids, respectively (Fig. 1). These peptides present different numbers of basic and hydrophobic amino acids, and bind differently to lipid membranes. In buffer, all three labeled peptides showed a single red-shifted emission band similar to that of the free label, though the fluorescence quantum yield is significantly higher with the peptides. We speculate that the proximal peptide backbone could partially screen the probe from water and, thus reduce its quenching by water.

Addition of DOPC vesicles changed dramatically the fluorescence of the labeled melittin and magainin 2 but not of poly-L-lysine, confirming that only the two first peptides can interact with neutral lipid bilayers. For both melittin and magainin 2, a large increase in the fluorescence quantum yield was accompanied with the appearance of the two-band emission characteristic for the probe in hydrophobic environment (Table 2, Fig. 3). Thus, the interaction of the two peptides with lipid membranes results in the transfer of the label from water into the hydrophobic membrane environment which affects dramatically its emission properties.

The binding stoichiometry is an important characteristic of peptide–lipid interactions. Since the fluorescence of the free peptide differs significantly from that of the membrane-bound peptide, the binding stoichiometry could be determined in a simple titration experiment. For these studies, we selected the interaction between melittin and DOPC vesicles. On addition of the labeled peptide to a suspension of DOPC vesicles, the fluorescence increased linearly with the peptide concentration reaching saturation at a ratio of 1 peptide per 100 lipids. Above this ratio, the intensity remained constant indicating a saturation of the membrane surface by the peptide (Fig. 4A). We performed also a reverse titration by adding increasing

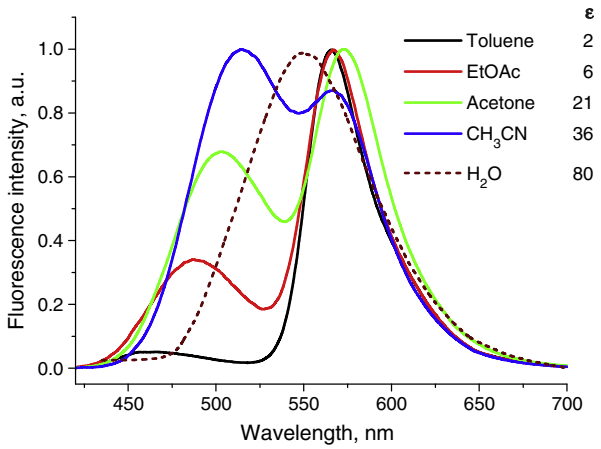


Fig. 2. Normalized fluorescence spectra of the MFL label in different solvents. Concentration of MFL was 1 μM. Excitation wavelength was 400 nm.

Table 1
Spectroscopic properties of the MFL label in solvents.

No	Solvent	ε	λ_{ABS}, nm	λ_{N^*}, nm	λ_{T^*}, nm	I_{N^*}/I_{T^*}	QY
1	Toluene	2.37	403	466	566	0.05	0.226
2	EtOAc	5.99	396	487	567	0.34	0.095
3	Acetone	20.5	398	503	573	0.68	0.101
4	CH_3CN	35.7	399	516[a]	575[a]	2.32[a]	0.154
5	EtOH	24.9	408	529	–	–	0.482
6	MeOH	32.6	406	533	–	–	0.262
7	Buffer	78.4	394	550	–	–	0.004

ε—dielectric constant, λ_{ABS}—position of absorption maxima, λ_{N^*} and λ_{T^*}—position of fluorescence maxima of N* and T* forms, respectively. I_{N^*}/I_{T^*}—ratio of the intensities of the two emission bands at their peak maxima (errors are ±2%). QY— fluorescence quantum yield (errors are ±5%), measured using FE in EtOH (QY = 51%) as a reference. 20 mM phosphate buffer, 150 mM NaCl, pH = 7.4 was used. Excitation wavelength was 400 nm.

[a] The values were evaluated from deconvolution of the emission spectra.

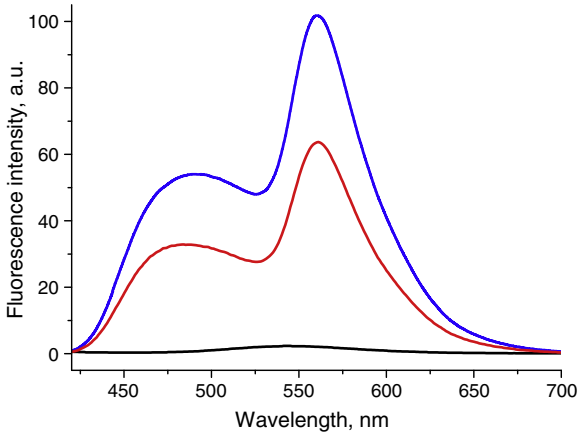


Fig. 3. Fluorescence changes on binding of MFL-labeled melittin and magainin 2 to DOPC vesicles. Emission spectra of both peptides in the absence (black) and in the presence of LUVs composed of neutral DOPC lipids (magainin 2—red, melittin—blue). Concentration of peptide and lipids was 0.3 and 100 μM, respectively. Buffer was 20 mM phosphate, 150 mM NaCl, pH = 7.4. Excitation wavelength was 400 nm.

lipid concentrations to a constant concentration of the labeled peptide (Fig. 4B). We observed that the fluorescence intensity reaches saturation at a ratio of 1 peptide per 100 lipids, confirming the lipid/peptide binding stoichiometry. Moreover, this titration curve allowed us to evaluate the binding constant $K = 4 \pm 2 \times 10^7\ M^{-1}$ of melittin to the DOPC membrane.

The ability of the label to monitor the binding of the peptide to the membrane was then used to determine the appropriate lipid composition required for efficient binding of poly-L-lysine to lipid membranes. For this purpose, we studied the fluorescence spectra of the labeled poly-L-lysine in the presence of vesicles containing mixtures of neutral lipid DOPC with different molar fraction of negatively charged lipid DOPS (Fig. 5). A strong (ca 20 times) increase in the fluorescence intensity of the labeled poly-L-lysine was recorded for vesicles, where the content of DOPS exceeded 10 mol% (Fig. 5), which thus represents the critical concentration of negatively charged lipids required for binding of poly-L-lysine. In the further experiments, we used 20 mol% of DOPS in DOPC vesicles, which provides efficient poly-L-lysine binding.

It was also important to monitor by dynamic light scattering the possible aggregation of vesicles in the presence of cationic peptides [45]. Neither melittin, nor magainin 2 affected the size of DOPC/DOPS (80/20 mol%) vesicles at a lipid/probe ratio 400:1, indicating that these peptides do not induce vesicle aggregation in these conditions. In contrast, an aggregation was observed with the highly charged poly-L-lysine (data not shown) at lipid to peptide ratios < 8000/1. Therefore, in further experiments, an 8000:1 ratio was used with poly-L-lysine.

3.3. Localizing peptide N-terminus in the membrane

Lipid membranes are complex media, where the polarity and hydration parameters exhibit a steep dependence on the depth in the bilayer [46]. Our label, being sensitive to both of these parameters, can therefore be used for evaluating the insertion depth of the peptide N-terminus to which the probe is covalently bound. Comparison of the dual emission of the peptides bound to lipid vesicles of the same composition (DOPC/DOPS) shows that the relative intensity of the short-wavelength band is much larger for poly-L-lysine than for the two other peptides (Fig. 6A), suggesting that the polarity and/or hydration of the peptide N-terminus is significantly larger for poly-L-lysine. To evaluate quantitatively the hydration and polarity of the label environment in these peptides, we used a methodology well established for membrane probes based on the same fluorophore [14,15]. In this methodology, the apparent two-band emission spectrum is deconvolved into three bands, with the N* and T* bands corresponding to the non-hydrated form of the dye and the H–N* band corresponding to the hydrated form. The H–N* form is usually localized at the hydrated membrane interface and its emission band is strongly red-shifted with respect to the N* band [14,15]. Using this approach, we deconvolved the emission spectra of the peptides bound to DOPC/DOPS vesicles into N*, T* and H–N* bands (Fig. 6B).

Table 2
Spectroscopic properties of MFL-labeled peptides.[a]

Peptide	Media	λ_{ABS}, nm	λ_{N^*}, nm	λ_{T^*}, nm	I_S/I_L	I_{N^*}/I_{T^*}	Hydration	QY
MFL–Magainin 2	Buffer	414	547	–	–	–	–	0.044
	DOPC	406	475	562	0.55	0.62	0.42	0.459
	DOPC/DOPS(8/2)	406	474	562	0.62	0.74	0.41	0.464
MFL–Melittin	Buffer	408	544	–	–	–	–	0.022
	DOPC	406	476	562	0.54	0.58	0.44	0.596
	DOPC/DOPS(8/2)	406	477	562	0.61	0.68	0.44	0.406
MFL–Poly-L-lysine	Buffer	410	553	–	–	–	–	0.032
	DOPC	414	556	–	–	–	–	0.048
	DOPC/DOPS(8/2)	~413	474	559	0.97	1.52	0.37	0.669

[a] λ_{ABS} is the position of absorption maxima; λ_{N^*} and λ_{T^*} are the maxima of N* and T* emission bands, respectively; I_S/I_L is the intensity ratio of the short- and long-wavelength bands measured at the peak maxima; I_{N^*}/I_{T^*} is the intensity ratio of the two emission bands measured at the peak maxima obtained from deconvolution of the corresponding spectra into three bands; all data, except quantum yields, refer to LUVs prepared by the ethanol dilution method; quantum yield (QY) measurements were done in LUVs prepared by the extrusion method. Excitation wavelength was 400 nm. Peptide and lipids concentrations were 0.3 and 100 μM, respectively. Buffer is 20 mM phosphate, 150 mM NaCl, pH = 7.4. Estimated errors: $\lambda_{ABS} \pm 2$ nm; λ_{N^*}, $\lambda_{T^*} \pm 1$ nm; $I_S/I_L \pm 2\%$; $I_{N^*}/I_{T^*} \pm 3\%$; Hydration ± 3%; QY ± 5%.

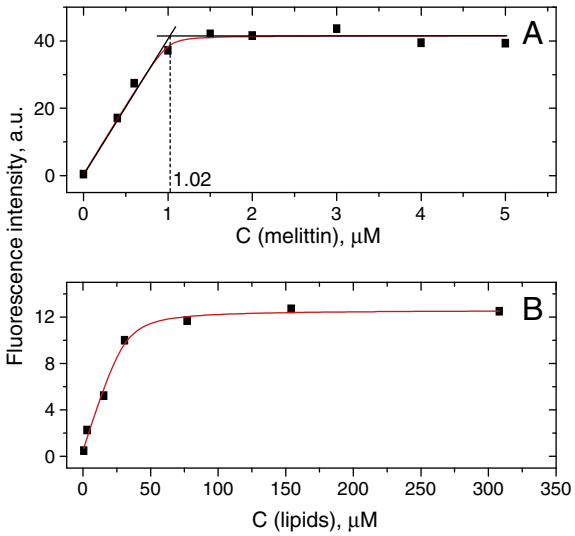


Fig. 4. Determination of the binding parameters of MFL-melittin to DOPC vesicles. Titration of 100 μM DOPC vesicles by MFL-melittin (A). Titration of 0.3 μM MFL-melittin by DOPC vesicles (B). The binding constant of MFL-melittin to the DOPC vesicles was determined by fitting (red curve) the data points in B to equation 1 in the materials and methods section.

From the deconvolved spectra, the hydration parameter can be evaluated from the ratio of the peak emission intensity of the hydrated (H–N*) form to the summed intensities of the non-hydrated (N* and T*) forms, while the "polarity" parameter can be evaluated from the I_{N^*}/I_{T^*} ratio. The obtained results show that the hydration parameter is nearly the same for all three peptides, while the polarity parameter is much larger for poly-L-lysine than for the two other peptides (Table 2). Previously, we have established for the same 3HF fluorophore that its I_{N^*}/I_{T^*} ratio in aprotic solvents varies linearly with the dielectric constant function, $f(\varepsilon)=(\varepsilon-1)/(2\varepsilon+1)$ [11] (Fig. 7). Here, we used this linear dependence to estimate the environment polarity of the label (non-hydrated form) in the membrane for the three labeled peptides. We found that for melittin and magainin 2, the observed I_{N^*}/I_{T^*} ratios correspond to a dielectric constant $\varepsilon \sim 10$–11, while for poly-L-lysine, the I_{N^*}/I_{T^*} ratio corresponds to an ε value close to 39 (Fig. 7). Using the dielectric model of lipid bilayer suggested by Griffith et al [46], these values correspond to depths of ~10 and ~20 Å from the bilayer center, i.e. to the regions of the fatty acid chains and phosphate residues, respectively.

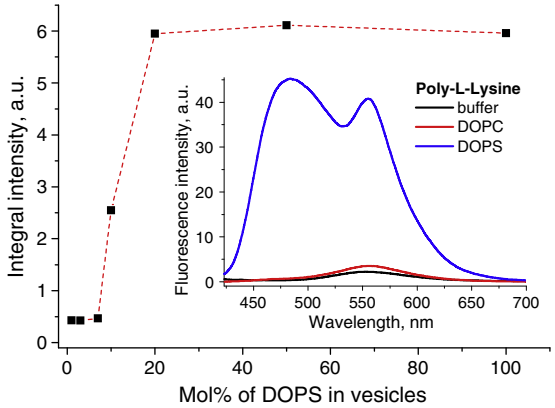


Fig. 5. Dependence of poly-L-lysine/LUV binding (integral fluorescence intensity of the label) on the content of negatively charged lipids (DOPS). The dashed red curve is used to connect the data points (squares). Fluorescence spectra of the peptide with and without lipid vesicles are shown in the inset. Concentration of peptide and lipids was 0.3 and 100 μM, respectively. Excitation wavelength was 400 nm.

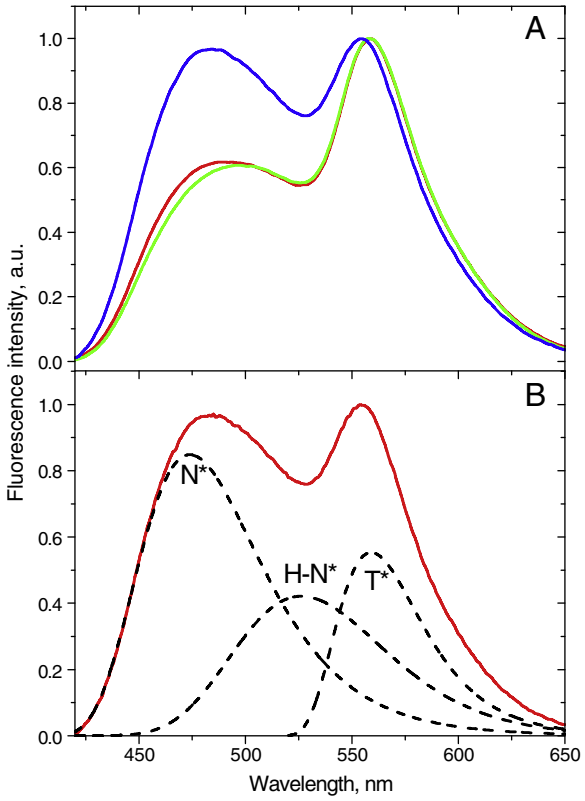


Fig. 6. Fluorescence spectra of MFL—melittin, MFL—magainin 2 and MFL—poly-L-lysine bound to DOPC/DOPS (80/20 mol%) vesicles. (A) Normalized fluorescence spectra of MFL—melittin (green), MFL—magainin 2 (red) and MFL—poly-L-lysine (blue) bound to DOPC/DOPS vesicles. (B) Emission spectrum of poly-L-lysine (solid red line) and its deconvolution into N*, H–N*, and T* spectral components. Experiments were performed in 20 mM phosphate buffer, 150 mM NaCl, pH = 7.4, using an excitation wavelength of 400 nm.

An independent measure of the localization of a fluorescent dye in a bilayer can be provided by the parallax quenching method using spin-labeled lipids [34,35]. Previously, using this method we were

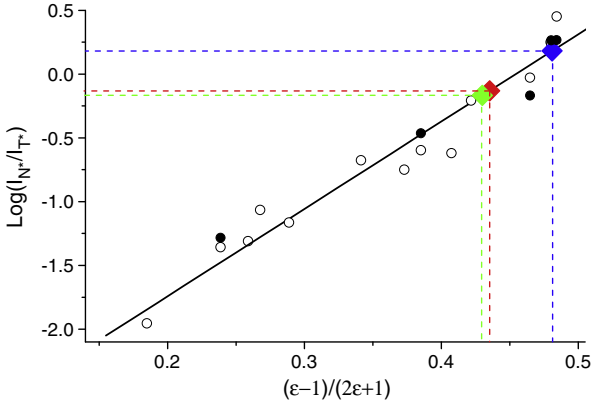


Fig. 7. Evaluation of the polarity of the probe environment for the three labeled peptides bound to DOPC/DOPS vesicles. The logarithm of I_{N^*}/I_{T^*} was plotted as a function of the polarity function, $f(\varepsilon)$, for FE (○) (data from ref. [11]) and MFL (●) in different solvents as well as for MFL-magainin 2 (red), MFL-melittin (green) and MFL-poly-L-lysine (blue) bound to DOPC/DOPS (80/20 mol%) vesicles. The data points were fitted (black line) with the following equation: $\log(I_{N^*}/I_{T^*}) = -3.111 + 6.849 f(\varepsilon)$, $r^2 = 0.976$, SD = 0.159.

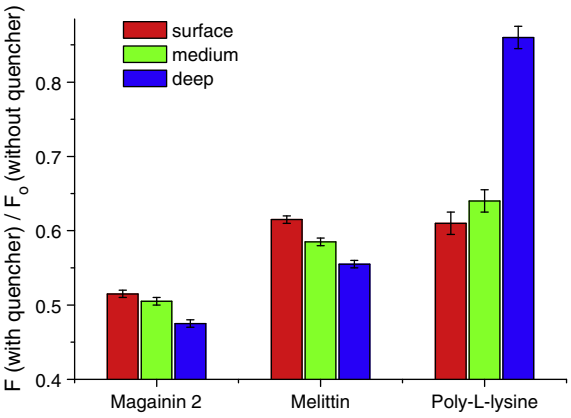


Fig. 8. Fluorescence quenching of MFL-magainin 2, MFL-melittin and MFL-poly-L-lysine bound to DOPC/DOPS (80/20 mol%) vesicles containing 15 mol % TempoPC (red), 5-SLPC (green), or 12-SLPC (blue), respectively. The fluorescence intensities of the labeled peptides bound to vesicles in the presence of quenchers were compared to the fluorescence intensities in their absence.

able to localize the fluorophore of membrane probes in lipid membranes [47]. Presently, we used this method to evaluate directly the depth of the labels in the bilayer for all three peptides. Three quenchers of deep (5.85 Å from the bilayer center), middle (12.15 Å) and shallow location (19.5 Å) were used. In a first step, we performed the parallax analysis on the basis of the integral intensity of the whole emission spectrum. For both melittin and magainin 2, the deep quencher was the most efficient, while the shallow quencher was the least efficient (Fig. 8). Using the parallax equation (see Materials and methods), we obtained average depths for melittin and magainin 2 of 8.3–8.4 Å from the bilayer center. In contrast, for poly-L-lysine the quenching was the most efficient with the shallow quencher and the least efficient for the deep quencher (Fig. 8), giving an estimated depth of 16.6 Å. Thus, the parallax analysis of the integral intensity of the whole emission spectra gave us a clear difference in the depths of the label for melittin and magainin 2 in comparison to poly-L-lysine. In a next step, the quenching of the hydrated and non-hydrated forms of the label was analyzed independently after deconvolution of the corresponding emission spectra. The non-hydrated form of the label for melittin and magainin 2 in DOPC/DOPS membrane was found to be located at 8.0–8.1 Å from the bilayer center (Table 3), in the region of the fatty acid chains (Fig. 9), in line with our estimations based on the relatively low dielectric constant of its environment (Fig. 7). The location of the hydrated form of the label for both proteins was 16.4 Å, indicating a shallow position next to the head groups (Fig. 9). Two reasons could explain this large difference in the locations of the two forms of the label. One is the relative freedom of the label, due to the presence of the flexible spacer, which could allow both deep and shallow localizations of the label (Fig. 9). The other reason could be that melittin and magainin 2 present both tilted and surface location in the lipid bilayer, thus producing the observed heterogeneity in depths. Remarkably, for poly-L-lysine, both hydrated and non-hydrated forms present a shallow location close to the head groups, around 16.5 Å from the bilayer center (Table 3, Fig. 9), in line with the high dielectric constant of its environment (Fig. 7). Noticeably, the fluorescence quantum yields were rather close for deeply inserted and shallow peptides, in line with the poor correlation of the fluorescence quantum yield of MFL (Table 1) and its analogues [11] with solvent polarity.

3.4. Label orientation

Orientation of the label is a parameter that could also help to understand the behavior of the protein and its label in the bilayer. A direct measurement of the fluorophore orientation in the bilayer of giant vesicles (GUVs) can be performed by fluorescence microscopy using a polarized laser excitation. Fluorophores with a transition dipole moment parallel to the polarization plane are excited preferentially. This method was recently used to determine the orientation of membrane probes based on Prodan [48] and 3HF [33]. In the present work, labeled melittin and poly-L-lysine were added to GUVs composed of DOPC/DOPS and studied by two-photon microscopy. We recorded the images in the blue and red regions, which were then used to calculate the ratio images.

In the case of poly-L-lysine, the intensity recorded for both blue and red channels was highest at the poles of the GUVs, while the lowest intensity was observed on their equator (Fig. 10A and B). This intensity distribution is typical for a fluorophore oriented vertically in the bilayer (orthogonal to the lipid surface, Fig. 9 [33,48]). Moreover, in the ratio image, the color of the poles and equator were almost identical (Fig. 10C), so that the dual emission of the label in this case is independent of the direction of the light polarization. For the labeled melittin, it was found that for the blue channel, the intensity was the same at the poles and equator (Fig. 10D), while for the red channel GUVs are brighter at the poles than at the equator (Fig. 10E). In the ratiometric images, the poles and equator appeared in different colors corresponding to lower and higher I_{blue}/I_{red} ratios, respectively (Fig. 10F). The observed effect of polarization on the emission color suggests two populations for the label in the bilayer. One population is associated with a low I_{blue}/I_{red} ratio (observed mainly in the red channel, Fig. 10E) and orients vertically in the bilayer, while the second population is associated with a high I_{blue}/I_{red} ratio (observed mainly in the blue channel, Fig. 10D) and does not exhibit any preferential orientation. This conclusion corroborates the parallax data, which also suggest two populations for the melittin label: a deep non-hydrated population and a shallow hydrated population. Therefore, we could conclude that the deeply located label population is oriented vertically in the bilayer, while the shallow population presents no preferential orientation. A similar behavior was previously observed for the parent fluorophore F in lipid membranes and to some extent for membrane probes based on the same fluorophore [14,15,33]. In contrast, poly-L-lysine showed no heterogeneity in the probe orientation. Since the N-terminus of this peptide is at the

Table 3
Parallax quenching data of non–hydrated and hydrated forms of MFL-labeled magainin 2, melittin and poly-L-lysine in DOPC/DOPS (80/20 mol%) vesicles.[a]

Peptide	F_{TC}/F_0	F_5/F_0	F_{12}/F_0	Z_{cf}, Å	F_{TC}/F_0	F_5/F_0	F_{12}/F_0	Z_{cf}, Å	$<Z_{cf}>$, Å
Forms	Non-hydrated form				Hydrated form				Both forms
Magainin 2	0.53	0.52	0.48	8.0	0.43	0.45	0.44	16.4	8.3
Melittin	0.67	0.62	0.57	8.1	0.49	0.52	0.52	16.4	8.4
Poly-L-lysine	0.63	0.67	0.89	16.5	0.54	0.57	0.85	16.3	16.6

[a] F_{TC}/F_0, F_5/F_0, and F_{12}/F_0 are the ratios of the fluorescence intensities of the labeled peptides bound to DOPC/DOPS (80/20 mol%) vesicles containing 15 mol % TempoPC, 5-SLPC, or 12-SLPC, respectively, to the corresponding fluorescence intensities in the absence of the nitroxide-labeled lipids. The intensities were obtained by deconvolution of the fluorescence quenching data. The obtained integral intensities of the H–N* form and the sum of the integral intensities of N* and T* forms were used to calculate the quenching for hydrated and non-hydrated forms, respectively. Z_{cf} is the distance between the middle of the bilayer and the chromophore center calculated from the parallax equation. $<Z_{cf}>$ is the average distance calculated from the integral intensity of the label without deconvolution. The estimated error for Z_{cf} is ± 1 Å.

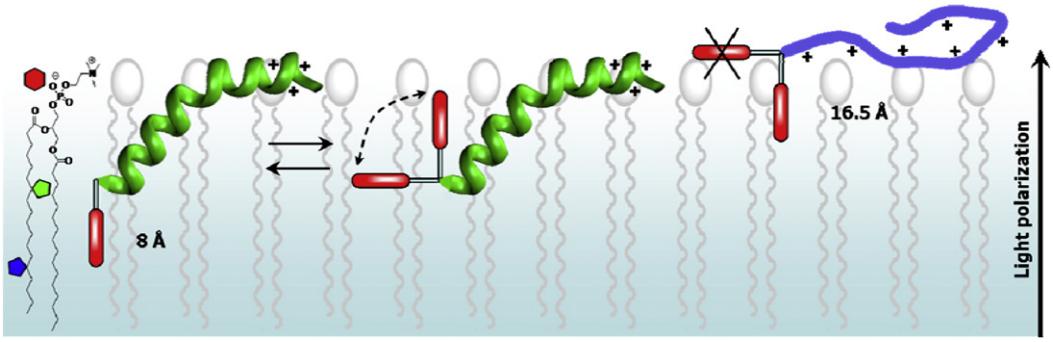


Fig. 9. Localization of the N-terminus of melittin (green) and poly-L-lysine (blue) labeled by MFL probe and bound to DOPC/DOPS (80/20 mol%) membrane. The positions of the quenching residues of the deep (blue), middle (green) and shallow (red) quenchers are pointed by the colored polygons. In the case of melittin, two locations of the label are shown: deep vertical (non-hydrated) and shallow with a distribution of orientations (hydrated).

bilayer surface [45], the relatively hydrophobic label likely orients vertically towards the hydrophobic interior of the bilayer (Fig. 9).

The present data illustrate the use of the present label for monitoring membrane binding and insertion of peptides. With this label, the binding of the label peptides to membranes can be readily detected by the strong increase in their fluorescence intensity and appearance of dual emission. Moreover, this label can provide a rapid evaluation of the approximate location of the peptide N-terminus by simply recording the ratio of its two emission bands. Importantly, the N-terminus localization of both deeply inserted (melittin and magainin 2) and surface located (poly-L-lysine) peptides, estimated by our approach, is in agreement with the literature data [21–26,28,45], suggesting that the MFL label did not change significantly the peptide insertion. However, the precision of this location is limited by the flexibility of the linker group, which allows delocalization of the probe, with different depths and orientations. This problem of label delocalization is partly avoided by distinguishing non-hydrated and hydrated species. Indeed, we observed that the non-hydrated species of the dye presents mainly a vertical orientation, and its depth correlates well with the protein insertion. In contrast, hydrated species are located at the surface, independently of the peptide insertion, allowing multiple orientations of the probe (Fig. 9). Thus, it is important to remove the hydrated species from the analysis, which can successfully be realized by spectral deconvolution. As a consequence, the N^*/T^* ratio, describing the polarity of non-hydrated species appears as a good indicator of protein insertion. In the future, further improvement in the location of the probe will be achieved by attaching it rigidly to the peptide backbone, as it has already been done with fluorine- and spin-labels for NMR [49,50] and EPR [51] spectroscopy, respectively. Moreover, the precision of the depth estimation based on the measured dielectric constant will also benefit from a better definition of the exact profile of the dielectric constant across the lipid membrane, which is still a matter of discussion [46,52,53].

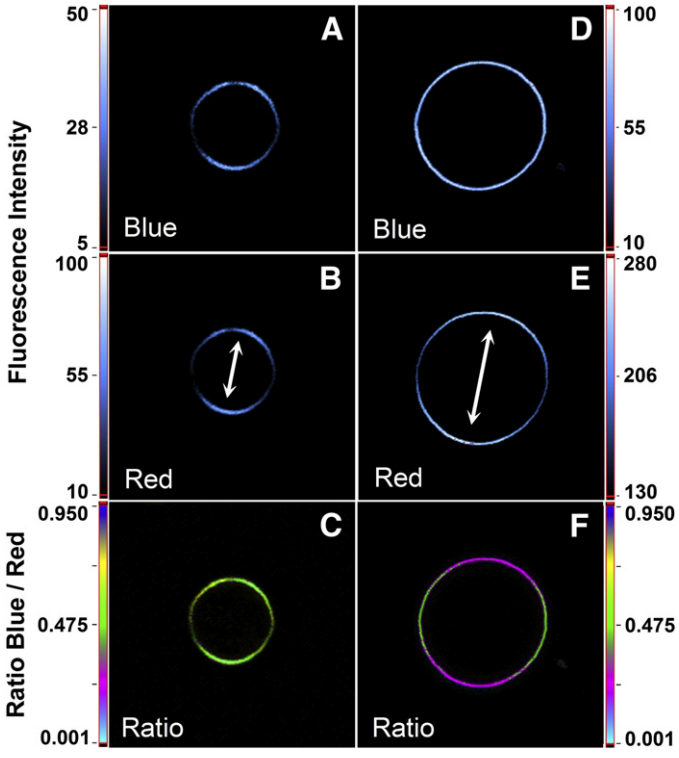


Fig. 10. Fluorescence microscopy imaging of MFL-labeled poly-L-lysine (A, B, C) and melittin (D, E, F) bound to DOPC/DOPS (80/20 mol%) GUVs. Intensity images at the blue (<585 nm, A and D) and red (>585 nm, B and E) channels. In the ratiometric images (C and F), the color of each pixel represents the value of the intensity ratio I_{blue}/I_{red}, while the pixel intensity corresponds to the total intensity at both channels. Two-photon excitation wavelength was at 830 nm. Arrows indicate the orientation of the light polarization. Sizes of the images were 70 μm × 70 μm. Concentration of poly-L-lysine and melittin was 0.05 and 0.1 μM, respectively.

4. Conclusion

The dual-fluorescence label MFL being attached at the N-terminus of peptides allows monitoring interactions of peptides with lipid bilayers, through a nearly two orders of magnitude increase in its emission intensity and appearance of two bands in its fluorescence spectrum. The high environmental sensitivity of the label allows separating the spectrum into individual components and determining the amounts of the non-hydrated and hydrated populations of the label. Then, the intensity ratio of the two emission bands of the non-hydrated species enables evaluation of the dielectric constant of the label environment and thus, estimation of its position in the lipid bilayer. We found that the N-terminus of melittin and magainin 2 is immersed into the bilayer, while that of poly-L-lysine is localized at the membrane surface. These conclusions are fully in line with parallax quenching measurements. Fluorescence microscopy measurements in giant vesicles further revealed differences in the orientation of the label bound to these two peptides. The proposed methodology for monitoring peptide–membrane interactions can in principle be extended to other peptides and environment-sensitive labels.

Acknowledgments

This work and VYP were supported by ANR (ANR-07-BLAN-0287, FLU-AA-NT), CNRS, Université de Strasbourg, and the ARCUS program between France, Ukraine and Russia. VGP was supported by Université de Strasbourg for short stays in YM's laboratory. We thank H. de Rocquigny for his help in peptide synthesis.

References

[1] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, 3rd ed.Springer-Verlag, New York, 2006.
[2] J. Ren, S. Lew, J. Wang, E. London, Control of the transmembrane orientation and interhelical interactions within membranes by hydrophobic helix length, Biochemistry 38 (1999) 5905–5912.
[3] J. Ren, S. Lew, Z. Wang, E. London, Transmembrane orientation of hydrophobic α-helices is regulated both by the relationship of helix length to bilayer thickness and by the cholesterol concentration, Biochemistry 36 (1997) 10213–10220.
[4] H. Zhao, P.K.J. Kinnunen, Binding of the antimicrobial peptide temporin L to liposomes assessed by Trp fluorescence, J. Biol. Chem. 277 (2002) 25170–25177.
[5] K. Matsuzaki, O. Murase, H. Tokuda, S. Funakoshi, N. Fujii, K. Miyajima, Orientational and aggregational states of magainin 2 in phospholipid bilayers, Biochemistry 33 (1994) 3342–3349.
[6] A.S. Ladokhin, S. Jayasinghe, S.H. White, How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? Anal. Biochem. 285 (2000) 235–245.
[7] G. Weber, F.J. Farris, Synthesis and spectral properties of a hydrophobic fluorescent probe: 6-propionyl-2-(dimethylamino)naphthalene, Biochemistry 18 (1979) 3075–3078.
[8] G. Anderluh, A. Barlic, Z. Podlesek, P. Macek, J. Pungercar, F. Gubensek, M.L. Zecchini, M.D. Serra, G. Menestrina, Cysteine-scanning mutagenesis of an eukaryotic pore-forming toxin from sea anemone. Topology in lipid membranes, Eur. J. Biochem. 263 (1999) 128–136.
[9] M. Zibouche, M. Vincent, F. Illien, J. Gallay, J. Ayala-Sanmartin, The N-terminal domain of annexin 2 serves as a secondary binding site during membrane bridging, J. Biol. Chem. 283 (2008) 22121–22127.
[10] A. Holt, R.B.M. Koehorst, T. Rutters-Meijneke, M.H. Gelb, D.T.S. Rijkers, M.A. Hemminga, J.A. Killian, Tilt and rotation angles of a transmembrane model peptide as studied by fluorescence spectroscopy, Biophys. J. 97 (2009) 2258–2266.
[11] A.S. Klymchenko, A.P. Demchenko, Multiparametric probing of intermolecular interactions with fluorescent dye exhibiting excited state intramolecular proton transfer, Phys. Chem. Chem. Phys. 5 (2003) 461–468.
[12] P.-T. Chou, M.L. Martinez, H. Clements, Reversal of excitation behavior of proton-transfer vs. charge-transfer by dielectric perturbation of electronic manifolds, J. Phys. Chem. 97 (1993) 2618–2622.
[13] A.P. Demchenko, Y. Mely, G. Duportail, A.S. Klymchenko, Monitoring biophysical properties of lipid membranes by environment-sensitive fluorescent probes, Biophys. J. 96 (2009) 3461–3470.
[14] A.S. Klymchenko, Y. Mely, A.P. Demchenko, G. Duportail, Simultaneous probing of hydration and polarity of lipid bilayers with 3-hydroxyflavone fluorescent dyes, Biochim. Biophys. Acta 1665 (2004) 6–19.
[15] A.S. Klymchenko, G. Duportail, A.P. Demchenko, Y. Mely, Bimodal distribution and fluorescence response of environment-sensitive probes in lipid bilayers, Biophys. J. 86 (2004) 2929–2941.
[16] V.V. Shynkar, A.S. Klymchenko, C. Kunzelmann, G. Duportail, C.D. Muller, A.P. Demchenko, J.M. Freyssinet, Y. Mely, Fluorescent biomembrane probe for ratiometric detection of apoptosis, J. Am. Chem. Soc. 129 (2007) 2187–2193.
[17] V.V. Shynkar, A.S. Klymchenko, G. Duportail, A.P. Demchenko, Y. Mely, Two-color fluorescent probes for imaging the dipole potential of cell plasma membranes, Biochim. Biophys. Acta 1712 (2005) 128–136.
[18] C.E. Dempsey, The actions of melittin on membranes, Biochim. Biophys. Acta Rev. Biomembr. 1031 (1990) 143–161.
[19] K. Matsuzaki, S. Yoneyama, K. Miyajima, Pore formation and translocation of melittin, Biophys. J. 73 (1997) 831–838.
[20] S.J. Ludtke, K. He, W.T. Heller, T.A. Harroun, L. Yang, H.W. Huang, Membrane pores induced by magainin, Biochemistry 35 (1996) 13723–13728.
[21] S. Frey, L.K. Tamm, Orientation of melittin in phospholipid bilayers. A polarized attenuated total reflection infrared study, Biophys. J. 60 (1991) 922–930.
[22] D. Sengupta, L. Meinhold, D. Langosch, G.M. Ullmann, J.C. Smith, Understanding the energetics of helical peptide orientation in membranes, Proteins 58 (2005) 913–922.
[23] S. Toraya, K. Nishimura, A. Naito, Dynamic structure of vesicle-bound melittin in a variety of lipid chain lengths by solid-state NMR, Biophys. J. 87 (2004) 3323–3335.
[24] H. Raghuraman, A. Chattopadhyay, Orientation and dynamics of melittin in membranes of varying composition utilizing NBD fluorescence, Biophys. J. 92 (2007) 1271–1283.
[25] B. Bechinger, Structure and functions of channel-forming peptides: Magainins, cecropins, melittin and alamethicin, J. Membr. Biol. 156 (1997) 197–211.
[26] H. Raghuraman, A. Chattopadhyay, Melittin: a membrane-active peptide with diverse functions, Biosci. Rep. 27 (2007) 189–223.
[27] W. Zauner, M. Ogris, E. Wagner, Polylysine-based transfection systems utilizing receptor-mediated delivery, Adv. Drug Deliv. Rev. 30 (1998) 97–113.
[28] W. Hartmann, H.J. Galla, Binding of polylysine to charged bilayer membranes. Molecular organization of a lipid peptide complex, Biochim. Biophys. Acta 509 (1978) 474–490.
[29] M.J. Hope, M.B. Bally, G. Webb, P.R. Cullis, Production of large unilamellar vesicles by a rapid extrusion procedure. Characterization of size distribution, trapped volume and ability to maintain a membrane potential, Biochim. Biophys. Acta Biomembr. 812 (1985) 55–65.
[30] K. Kachel, E. Asuncion-Punzalan, E. London, The location of fluorescence probes with charged groups in model membranes, Biochim. Biophys. Acta 1374 (1998) 63–76.
[31] M.I. Angelova, D.S. Dimitrov, Liposome electroformation, Faraday Discuss. Chem. Soc. 81 (1986) 303–311.
[32] M. Fidorra, L. Duelund, C. Leidy, A.C. Simonsen, L.A. Bagatolli, Absence of fluid-ordered/fluid-disordered phase coexistence in ceramide/POPC mixtures containing cholesterol, Biophys. J. 90 (2006) 4437–4451.
[33] A.S. Klymchenko, S. Oncul, P. Didier, E. Schaub, L. Bagatolli, G. Duportail, Y. Mely, Visualization of lipid domains in giant unilamellar vesicles using an environment-sensitive membrane probe based on 3-hydroxyflavone, Biochim. Biophys. Acta 1788 (2009) 495–499.
[34] F.S. Abrams, E. London, Extension of the parallax analysis of membrane penetration depth to the polar region of model membranes: use of fluorescence quenching by a spin-label attached to the phospholipid polar headgroup, Biochemistry 32 (1993) 10826–10831.
[35] A. Chattopadhyay, E. London, Parallax method for direct measurement of membrane penetration depth utilizing fluorescence quenching by spin-labeled phospholipids, Biochemistry 26 (1987) 39–45.
[36] R.D. Kaiser, E. London, Location of diphenylhexatriene (DPH) and its derivatives within membranes: comparison of different fluorescence quenching analyses of membrane depth, Biochemistry 37 (1998) 8180–8190.
[37] D.P. Tieleman, S.J. Marrink, H.J.C. Berendsen, A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems, Biochim. Biophys. Acta Rev. Biomembr. 1331 (1997) 235–270.
[38] B.A. Lewis, D.M. Engelman, Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles, J. Mol. Biol. 166 (1983) 211–217.
[39] H. Beltz, E. Piemont, E. Schaub, D. Ficheux, B. Roques, J.L. Darlix, Y. Mely, 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 (2004) 711–723.
[40] J. Azoulay, J.P. Clamme, J.L. Darlix, B.P. Roques, Y. Mely, Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations, J. Mol. Biol. 326 (2003) 691–700.
[41] J.P. Clamme, J. Azoulay, Y. Mely, Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy, Biophys. J. 84 (2003) 1960–1968.
[42] A.S. Klymchenko, V.G. Pivovarenko, T. Ozturk, A.P. Demchenko, Modulation of the solvent-dependent dual emission in 3-hydroxychromones by substituents, New J. Chem. 27 (2003) 1336–1343.
[43] V.V. Shynkar, A.S. Klymchenko, E. Piemont, A.P. Demchenko, Y. Mely, Dynamics of intermolecular hydrogen bonds in the excited states of 4′-dialkylamino-3-hydroxyflavones. On the pathway to an ideal fluorescent hydrogen bonding sensor, J. Phys. Chem. A 108 (2004) 8151–8159.
[44] A.S. Klymchenko, V.G. Pivovarenko, A.P. Demchenko, Elimination of the hydrogen bonding effect on the solvatochromism of 3-hydroxyflavones, J. Phys. Chem. A 107 (2003) 4211–4216.
[45] D. Volodkin, V. Ball, P. Schaaf, J.C. Voegel, H. Mohwald, Complexation of phosphocholine liposomes with polylysine. Stabilization by surface coverage versus aggregation, Biochim. Biophys. Acta Biomembr. 1768 (2007) 280–290.
[46] O.H. Griffith, P.J. Dehlinger, S.P. Van, Shape of the hydrophobic barrier of phospholipid bilayers (evidence for water penetration in biological membranes), J. Membr. Biol. 15 (1974) 159–192.
[47] A.S. Klymchenko, G. Duportail, T. Ozturk, V.G. Pivovarenko, Y. Mely, A.P. Demchenko, Novel two-band ratiometric fluorescence probes with different location and orientation in phospholipid membranes, Chem. Biol. 9 (2002) 1199–1208.
[48] L.A. Bagatolli, To see or not to see: lateral organization of biological membranes and fluorescence microscopy, Biochim. Biophys. Acta 1758 (2006) 1541–1556.
[49] S. Afonin, R.W. Glaser, M. Berditchevskaia, P. Wadhwani, K.H. Guhrs, U. Mollmann, A. Perner, A.S. Ulrich, 4-Fluorophenylglycine as a label for 19F NMR structure analysis of membrane-associated peptides, ChemBioChem 4 (2003) 1151–1163.
[50] P.K. Mikhailiuk, S. Afonin, A.N. Chernega, E.B. Rusanov, M.O. Platonov, G.G. Dubinina, M. Berditsch, A.S. Ulrich, I.V. Komarov, Conformationally rigid trifluoromethyl-substituted alfa-amino acid designed for peptide structure analysis by solid-state 19F NMR spectroscopy, Angew. Chem. Int. Ed. 45 (2006) 5659–5661.
[51] G.E. Fanucci, D.S. Cafiso, Recent advances and applications of site-directed spin labeling, Curr. Opin. Struct. Biol. 16 (2006) 644–653.
[52] H. Nymeyer, H.X. Zhou, A method to determine dielectric constants in nonhomogeneous systems: application to biological membranes, Biophys. J. 94 (2008) 1185–1193.
[53] I. Konopasek, P. Kvasnicka, E. Amler, A. Kotyk, G. Curatola, The transmembrane gradient of the dielectric constant influences the DPH lifetime distribution, FEBS Lett. 374 (1995) 338–340.

2.2.2. Investigation of the interaction of nucleocapsid protein with lipid membranes

One of the important steps in the process of retroviral infection is transport of the preintegration complex (PIC) into the nucleus of nondividing cells. However, a controversy still exists on the mechanism of nuclear entry (Yamashita and Emerman 2005). Nevertheless, several viral proteins are suspected to intervene in this process such as the matrix protein, integrase, Vpr protein and the central flap region (Zennou, Petit et al. 2000; Arhel, Munier et al. 2006). Mutations of these proteins affect the infectivity of HIV-1 but did not inhibit it. One hypothesis is that these proteins have complementary roles for the nuclear import of the viral genome. Another possibility is that there may be other additional mediators of HIV-1 nuclear import not yet identified.

The nucleocapsid protein (NC) is required at many steps of viral life cycle: reverse transcription, integration and assembly due to its nucleic acid chaperone properties (Thomas and Gorelick 2008). In addition, NC is also thought to bind to several viral and host proteins. Lipid membranes may be an additional binding target of NC, even though this has never been clearly demonstrated. NC could play a role in the process of the entry of PIC through the nuclear membrane. Its role at this step is totally unknown. The presence of NC in the PIC is expected as a consequence of its presence in multiple copies in the reverse-transcription complex (RTC). However, the presence of NC in the PIC was not clearly demonstrated. The failure to detect NC could be either due to its weak binding to the viral DNA, which would be responsible for the loss of NC during the process of isolation of the PIC or the amount of NC bound in the PIC may be below the limit of detection by Western blotting (Thomas and Gorelick 2008). However, experiments have shown nuclear localization of NC after infection (Gallay, Swingler et al. 1995; Zhang and Crumpacker 2002) and its role in the integration step (Carteau, Gorelick et al. 1999). Thus, a thorough characterization of the interaction of NC with membranes would provide new clues for understanding the possible role of NC in the entry of the PIC in the nucleus.

Labeled peptides

In this work, to study the possible interaction of NC with membranes, NC(1-55) peptide was synthesized by solid-phase peptide synthesis and N-terminally labeled with the 3-hydroxyflavone dye, MFL (Figure 2.18). As it was described before, the MFL label is highly sensitive to polarity in low polar media and represents a good tool for studying peptide-membrane interactions (Postupalenko, Shvadchak et al. 2011). The chaperone properties of the labeled NC were tested using previously described protocols (Bernacchi, Stoylov et al. 2002; Beltz, Azoulay et al. 2003; Godet, De Rocquigny et al. 2006) and were found to be close to those of the native one indicating that the MFL label marginally perturbs the activity of the peptide. Labeled NC in the buffer showed a single red-shifted emission band similar to that of the free MFL label. Upon incorporation of MFL label into NC peptide, the fluorescence quantum yield was significantly increased (Table 2.1), probably, due to the close proximity of peptide backbone which could partially screen the probe from water and, thus reduce its quenching by water.

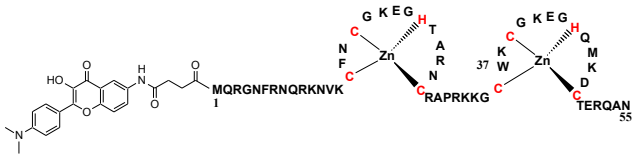


Figure 2.18. Structure of NC peptide N-terminally labeled with the MFL probe.

Binding of NC to membrane

For studying NC-membrane interactions, as a model membrane system we first used large unilamellar vesicles (LUVs), which are thermodynamically stable and uniform in size. Addition of negatively charged LUVs (composed of DOPS or DOPG lipids) to the labeled NC increased its fluorescence quantum yield by a factor of eight and induced the appearance of two separated N* and T* bands indicating an interaction of the peptide with the negatively charged LUVs (Figure 2.19). The observed effects clearly showed that the dye was transferred to an aprotic environment. In contrast to the negatively charged LUVs, the neutral LUVs (composed of DOPC) did not significantly affect the fluorescence properties of MFL-NC (Figure 2.19, Table 4), suggesting an absence of interaction in this case In a next step, we used a mixture of lipids (DOPC:DOPE:DOPS:PI:PIP_2 in a ratio of 38:30:5:24:3 (mole % of lipids)) which was shown to mimic nuclear membrane (Byrne, Gamier-Lhomme et al. 2007). Addition of labeled NC to this mixture (nuclear membrane like composition – NMLC) led to an effect similar to that observed with negatively charged vesicles, with an increase of the fluorescent quantum yield and the appearance of a two-band emission (Figure 2.19). Based on these data, it appears that NC binds efficiently to negatively charged lipids. This binding mainly depends on the charge of the lipids, as could be seen from the similar responses of the labeled peptide upon binding to DOPS, DOPG or NMLC. Nevertheless, careful comparison of the spectra indicate that the NMLC mixture induced a lower N*/T* ratio and a blue-shifted N* band that indicate a significantly more apolar environment of the MFL label in this case.

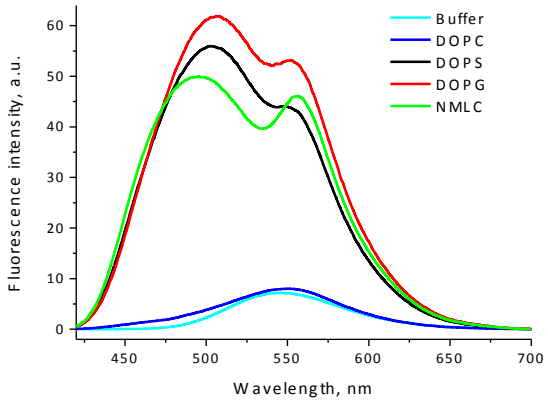


Figure 2.19. Fluorescence of MFL-NC in the absence (cyan) and in the presence of LUVs composed of either neutral DOPC (blue), negatively charged DOPS (black), DOPG (red) or NMLC (green) lipids. Experiments were performed in 20 mM phosphate buffer, pH 7.4 containing 150 mM

of NaCl to reduce the impact of electrostatic interactions. Concentration of peptide and lipids was 0.25 and 100 μM, respectively. Excitation wavelength was 400 nm.

Table 4. Spectroscopic properties of labeled NC bound to LUVs of various compositions

Peptide	Media	λ_{Abs}, nm	λ_{N*}, nm	λ_{T*}, nm	I_{N*}/I_{T*}	QY
MFL-NC	Buffer	419	545	-	-	0.07
	DOPC	420	547	-	-	0.08
	DOPS	410	504	552	1.27	0.56
	DOPG	408	506	553	1.16	0.62
	NMLC	410	495	556	1.08	0.50

λ_{Abs}, λ_{N*} and λ_{T*} are the maxima of absorption, N* and T* emission bands, respectively. I_{N*}/I_{T*} is the intensity ratio of the two emission bands measured at the peak maxima; QY is the fluorescence quantum yield (estimated error of QY is ± 10%). Excitation wavelength was 400 nm. Concentration of peptide and lipids was 0.25 and 100 μM, respectively. Measurements were done in 20 mM phosphate buffer containing 150 mM NaCl (pH 7.4).

Since cell membranes consist of various lipids with various percentages of neutral and negatively charged lipids, we performed experiments to find the appropriate lipid composition which provides efficient binding of NC. For this, we prepared vesicles containing mixtures of neutral lipid DOPC with increasing molar fractions of negatively charged DOPS or DOPG lipids. A strong increase in the fluorescence intensity of the MFL-NC was observed for vesicles, where the content of DOPS or DOPG exceeded 30-40 mol% (Figure 2.20), which thus represents the critical concentration of negatively charged lipids required for binding of NC.

In the further experiments, we used 50-60 mol% of DOPS or DOPG in DOPC vesicles as a model membrane system, which provides efficient NC binding. This percentage is close to that in nuclear membranes (Byrne, Gamier-Lhomme et al. 2007) as well as in the inner leaflet of the plasma membrane (Op den Kamp 1979; Zachowski 1993) which consists of around 40 mol% of negatively charged lipids.

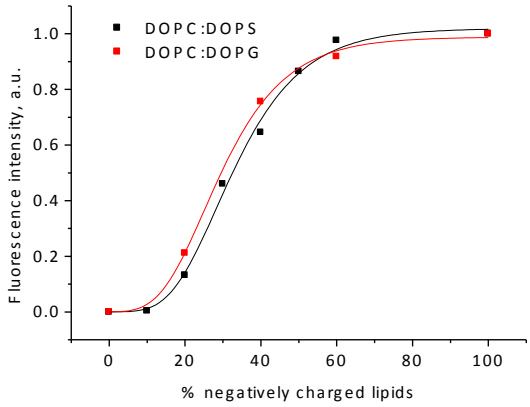


Figure 2.20. Dependence of MFL-NC/LUVs binding (integral fluorescence intensity of the label) on the percentage of negatively charged lipids (DOPS or DOPG). The black and red curves

were used to connect the data points (squares). Concentration of peptide and DOPS, DOPG lipids was 0.3 and 200 μM, 100 μM, respectively. Excitation wavelength was 400 nm.

As a next step, to further determine the appropriate working conditions to study NC-membrane interaction we performed dynamic light scattering experiments to monitor the possible aggregation of vesicles by NC. Several types of vesicles were used for this experiment: pure negatively charged one (DOPG), neutral (DOPC) and their mixtures (DOPG:DOPC (60:40) and DOPG:DOPC (40:60)). In case of DOPG vesicles, NC started to aggregate LUVs at lipid to peptide ratio < 300:1 (Figure 2.21). Therefore, in further experiments, a ratio >500:1 was used to work in non-aggregating conditions. A similar ratio of lipids to NC was needed to aggregate the DOPG:DOPC (60:40) lipid vesicles while in the case of DOPG:DOPC (40:60) vesicles, an aggregation occurred at higher NC concentration suggesting weaker binding of peptide (Figure 2.21). The size of neutral vesicles was not affected by the peptide, confirming the lack of binding. At high peptide concentrations, the cationic NC molecules likely neutralize the negative charges of the vesicles, leading to vesicle aggregation. Moreover, we observed that with increasing content of DOPC in the vesicles, higher concentration of NC was needed to cause an aggregation of LUVs due to the less efficient binding. By using these curves, the binding constant of NC to vesicles could be estimated. For example, we found that $K_{bind} > 10^7\ M^{-1}$ for DOPG vesicles.

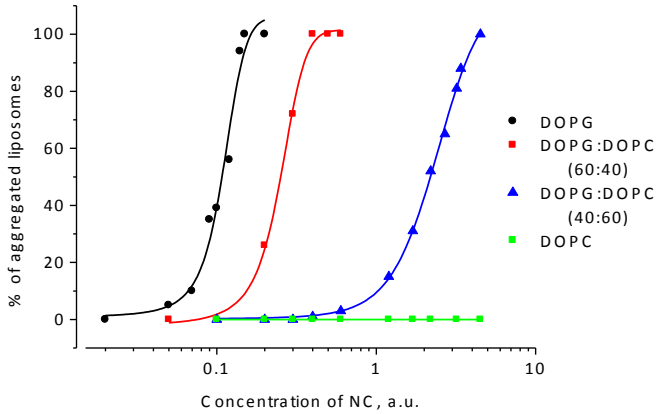


Figure 2.21. Dependence of the percentage of aggregated vesicles on the concentration of NC in various lipid compositions: DOPG (black), DOPG:DOPC (60:40) (red), DOPG:DOPC (40:60) (blue), DOPC (green). The curves were used to connect the data points. The concentration of lipids was 10 μM in all cases.

Determinants of binding to membrane

To identify the determinants of the NC-membrane interaction, three NC mutants were selected. Mutations were generated in one or both of Zn^{2+} fingers. The NC_{H23C} mutant has the first CCHC zinc finger mutated to a CCCC motif, while the NC_{SSHS} mutant has its both zinc fingers mutated to SSHS motifs, preventing zinc binding (Gorelick, Gagliardi et al. 1999; Guo, Wu et al. 2000). Previous studies have shown that these mutations greatly reduce the viral activity of NC, causing a defect in reverse transcription and a lack of integration (Thomas, Gagliardi et al. 2006). As a third mutant, we choose NC without its basic N-terminal domain (NC(11-55)). These

mutations should allow us to determine the contribution of the zinc fingers and the N-terminal basic domain in the interaction of NC with the bilayers and evaluate the importance of the folding of the fingers in the process. The structures of these peptides are presented on Figure 2.22.

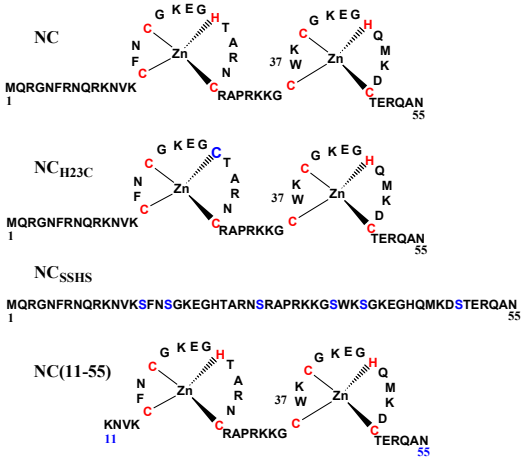


Figure 2.22. Structures of NC peptide and its mutants: NC_{H23C}, NC_{SSHS}, NC(11-55) .

These three NC mutants were synthesized by solid-phase peptide synthesis and N-terminally labeled with the MFL label. All mutated peptides showed a single red-shifted band in buffer similar to that of the native peptide. The fluorescence quantum yields of mutants were of the same range as for the native NC except for the NC_{SSHS} peptide, which showed a somewhat lower value (4%) (Table 5). This could be due to the higher exposure of the label to water in this non-structured peptide as compared to the other more structured peptides. Addition of DOPS or DOPG LUVs changed dramatically the fluorescence of all labeled mutants, indicating that all of these peptides are able to interact with LUVs. For all peptides, a strong increase in the fluorescence quantum yield was accompanied with the appearance of the two separated N* and T* bands. Similar effects were observed upon addition of the NMLC LUVs to these peptides (Table 5). NC_{H23C} and NC_{SSHS} mutants showed quantum yields comparable to that of the as the native NC when bound to LUVs, while in the case of NC(11-55) peptide, values are lower, probably due to its lower affinity to LUVs, as a consequence of its reduced number of basic residues (Table 5).

In contrast to the quantum yields, the positions of the emission bands and the values of the N*/T* ratio were indistinguishable for NC(11-55) and NC, indicating that the two peptides exhibit similar binding modes, with similar location of the MFL probe in the membrane. Significantly lower values of the N*/T* ratio were observed with the NC_{H23C} and NC_{SSHS} mutants bound to LUVs, indicating that the MFL probe in these peptides experienced a less polar environment than that of the native NC, when the peptides are bound to LUVs.

Table 5. Spectroscopic properties of labeled NC peptide and its mutants: NC_{H23C}, NC_{SSHS}, NC(11-55)

Peptide	Media	λ_{Abs}, nm	λ_{N*}, nm	λ_{T*}, nm	I_{N*}/I_{T*}	QY
MFL-NC	Buffer	419	545	-	-	0.07
	DOPS	410	504	552	1.27	0.56
	DOPG	408	506	553	1.16	0.62
	NMLC	410	495	556	1.08	0.50
$MFL\text{-}NC_{H23C}$	Buffer	421	549	-	-	0.09
	DOPS	410	508	552	1.07	0.55
	DOPG	409	510	555	0.98	0.58
	NMLC	409	498	557	0.83	0.50
$MFL\text{-}NC_{SSHS}$	Buffer	421	551	-	-	0.04
	DOPS	408	506	552	1.14	0.54
	DOPG	410	505	555	0.98	0.60
	NMLC	408	495	557	0.82	0.54
MFL-NC(11-55)	Buffer	420	545	-	-	0.09
	DOPS	410	505	552	1.24	0.43
	DOPG	408	505	553	1.17	0.49
	NMLC	410	496	556	1.04	0.43

$\lambda_{Abs,}$ λ_{N*} and λ_{T*} are the maxima of absorption, N* and T* emission bands, respectively. I_{N*}/I_{T*} is the intensity ratio of the two emission bands measured at the peak maxima; QY is the fluorescence quantum yield. Estimated errors: $\lambda_{Abs} \pm 2$ nm; λ_{N*}, $\lambda_{T*} \pm 1$ nm; $I_{N*}/I_{T*} \pm 3\%$; QY ± 10%. Excitation wavelength was 400 nm. Concentration of peptides and lipids were 0.25 and 100 μM, respectively. Measurements were done in 20 mM phosphate buffer containing 150 mM NaCl (pH 7.4).

As a next step, we compared the binding of the four labeled peptides to LUVs containing different molar fractions of negatively charged lipids. To this end, we performed experiments with vesicles consisting of mixtures of DOPC and DOPS lipids in different ratios as it was described above for the native NC. Interestingly, we found that NC_{H23C} and NC_{SSHS} mutants show the same dependence on the ratio of negative to neutral lipids than the native NC (Figure 2.23). Optimal binding for all these peptides was reached with LUVs containing at least 60% of DOPS. In contrast, NC(11-55) needs higher percentages of DOPS lipids to bind efficiently to the LUVs (Figure 2.23).

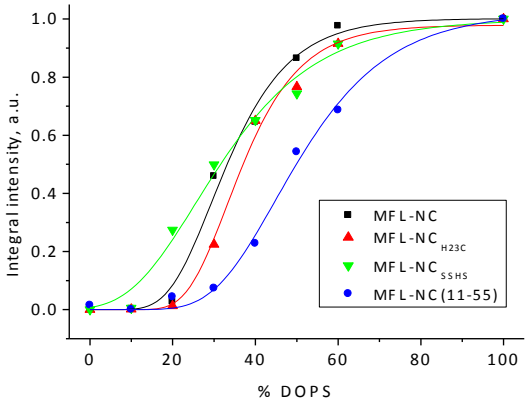


Figure 2.23. Dependence of binding of MFL-NC, MFL-NC_{H23C}, MFL-NC_{SSHS} and MFL-NC(11-55) peptides to LUVs on the percentage of negatively charged DOPS lipids. The curves were used to connect the data points (squares). Concentration of peptides and LUVs was 0.25 and 200 μM, respectively. Excitation wavelength was 400 nm.

Further, dynamic light scattering experiments were performed with NC derivatives to compare their aggregation ability. For this purpose, DOPS:DOPC (60:40) vesicles were used. From figure 2.24, it appears that mutants NC_{H23C} and NC_{SSHS} cause aggregation of vesicles in the same concentration range than native NC. In contrast, six times higher concentration of NC(11-55) was needed to get the same aggregation level (Figure 2.24), confirming that this mutant shows reduced membrane binding ability .

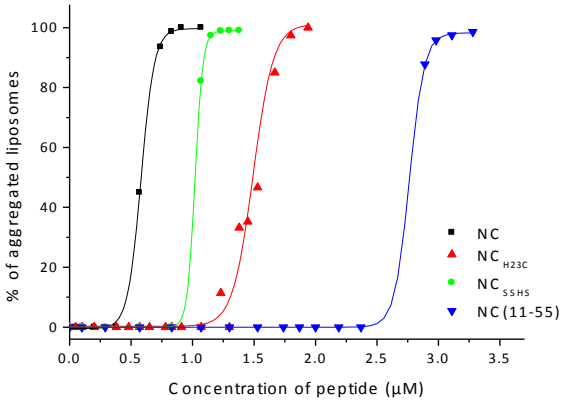


Figure 2.24. Dependence of the percentage of aggregated vesicles on the concentration of NC derivatives in DOPS:DOPC (60:40). The curves were used to connect the data points. The concentration of lipids was 50 μM in all cases.

Taken together, our data with the mutant peptides show that the interaction of NC with negatively charged membranes is mainly driven by electrostatic forces. Nevertheless, the folded fingers seem to provide a more polar location for the MFL probe, when the labeled peptide is bound to the membrane.

Localization of peptide N-terminus in the membrane

In order to further investigate the localization of NC bound to the membrane, we compared its fluorescence with the above mentioned model peptides (Postupalenko, Shvadchak et al. 2011). The fluorescence response of NC is close to that of poly-L-lysine, indicating that NC is located on the water-membrane interface. To further prove this assumption, we performed a parallax quenching experiment using spin-labeled lipids. As it was shown above (Postupalenko, Shvadchak et al. 2011), this method allows determination of the depth of the label in the bilayer. We performed this experiment with the NC peptide bound to DOPS vesicles and found that the N-terminus is located 17 Å from the center of the bilayer (phospho-group region). Afterwards, the depth of the label coupled to the NC peptide and its three mutants were studied in DOPC:DOPS (40:60) vesicles. For all peptides, the shallow quencher was the most efficient (Table 6). As a consequence, by using the parallax equation we obtained average depths of 16.2-17.7 Å from the center of bilayer. These depths pointed on a shallow location close to the head groups for NC, NC_{H23C}, NC_{SSHS} and NC(11-55) peptides.

Table 6. Parallax quenching data of MFL-labeled NC, NC_{H23C}, NC_{SSHS} and NC(11-55) peptides in DOPC:DOPS (40:60 mol%) vesicles

Peptide	F_{TC}/F_0	F_5/F_0	F_{12}/F_0	Z_{cf}, Å
MFL-NC	0.83	0.86	0.86	**16.2**
MFL-NC_{H23C}	0.77	0.83	0.84	**16.6**
MFL-NC_{SSHS}	0.78	0.94	0.95	**17.7**
MFL-NC(11-55)	0.80	0.84	0.85	**16.3**

F_{TC}/F_0, F_5/F_0, and F_{12}/F_0 are the ratios of the fluorescence intensities of the labeled peptides bound to DOPC:DOPS (40:60 mol%) vesicles containing 15 mol % TempoPC, 5-SLPC, or 12-SLPC, respectively, to the corresponding fluorescence intensities in the absence of the nitroxide-labeled lipids. The integral intensities of emission spectra were used to calculate the quenching. Z_{cf} is the distance between the middle of the bilayer and the chromophore center calculated from the parallax equation. $<Z_{cf}>$ is the average distance calculated from the integral intensity of the label without deconvolution. The estimated error for Z_{cf} is ±1 Å.

Destabilization of the lipid bilayer by NC peptide (AFM studies)

Atomic force microscopy (AFM) was used to further characterize the interaction of NC with lipids. As model membranes, we used supported lipid bilayers (SLBs) prepared by fusion of lipid vesicles on mica. Experiments were performed by using three lipid compositions: DOPC, DOPS and DOPC:DOPS (1:1) mixture. In order to characterize the behavior of supported lipid bilayers after the addition of NC, the protein concentration was increased over time, from the nM to the μM range. First, pure DOPC bilayer was formed on mica. Figure 2.25A shows a typical topography of a complete DOPC bilayer with very smooth and uniform surface. Increasing concentrations of NC were added until 1μM without notable change in the DOPC bilayer which remained smooth and uniform (Figure 2.25B), confirming that NCp7 does not interact with DOPC membranes.

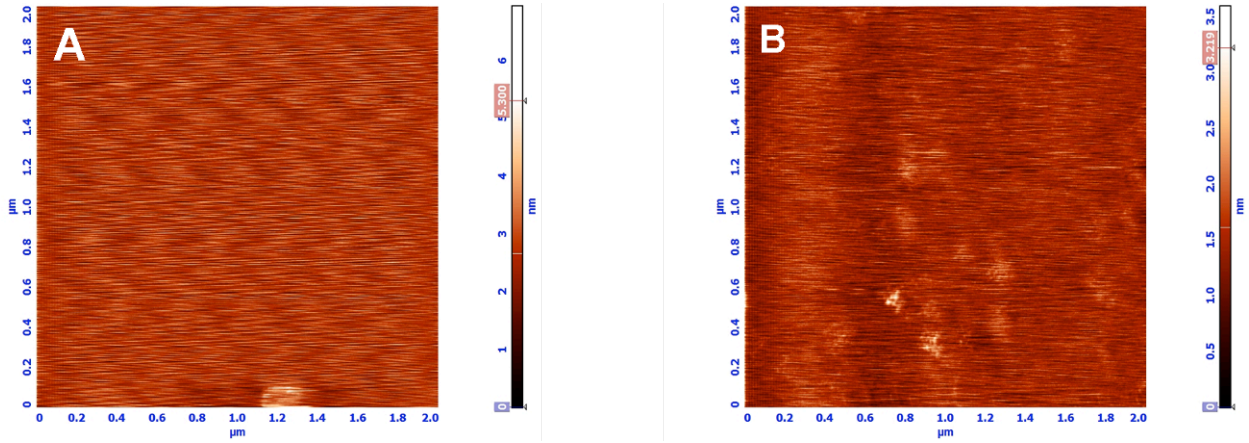


Figure 2.25. AFM images of a DOPC bilayer recorded prior (A) and after addition of 1µM NC (B).

Subsequently, we tested DOPS bilayers. Figure 2.26A shows an AFM image of the smooth and continuous DOPS bilayer on the mica surface. After addition of low concentrations of NC (0.01µM), particles (~2-4 nm height) could be observed on the surface but without affecting the stability of bilayer (Figure 2.26B). In the presence of 0.1µM NC, a destabilization of the DOPS bilayer was observed (Figure 2.26C), since holes appear progressively in the bilayer. Depth of the holes was about 4 nm indicating that the bilayer was removed from its support in these places. Moreover, large particles appeared on the surface. These particles are thought to be lipoprotein complexes formed of NC molecules complexed to lipids. Addition of 1µM NC further increased the bilayer fluidity. At this NC concentration, holes increased in size and the surface of the bilayer appeared severely destroyed. Moreover, the lipoprotein complexes increased in size and in number on the bilayer surface (Figure 2.26D).

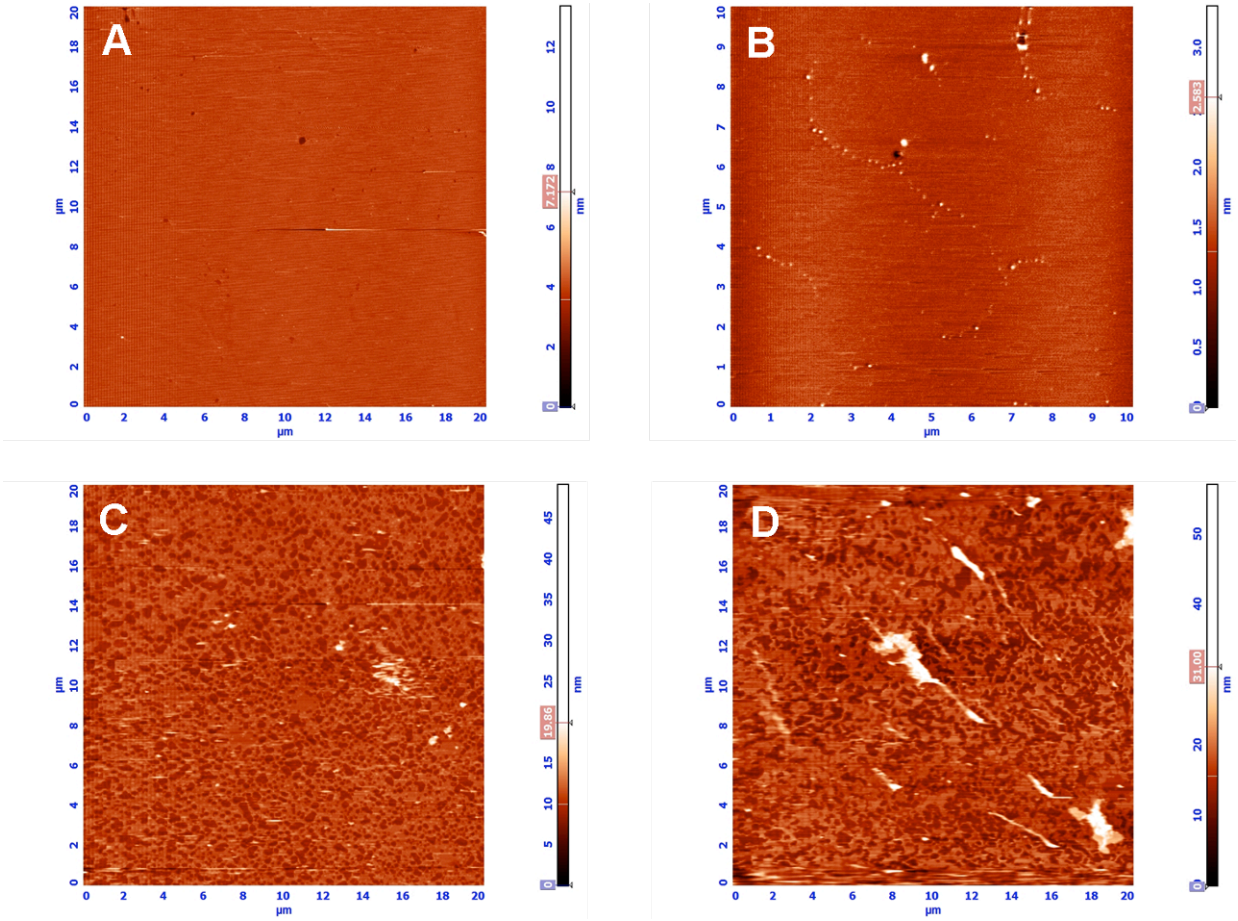


Figure 2.26. AFM images of DOPS supported lipid bilayer alone (20mM Hepes, 150mM NaCl, pH 7.4) (A) and in presence of different concentrations of NC: 0.01µM (B), 0.1µM (C), 1µM (D).

Afterwards, a 1:1 mixture of DOPS and DOPC lipids was used in AFM studies. As for pure DOPS bilayers, we observed a destabilization of the bilayers with the appearance of small defects in the DOPS:DOPC bilayer structure when 0.1 µM NC was added. An increase of NC concentration till 1µM led to the destruction of a large part of the bilayer with the appearance of many holes and large particles (data not shown). It therefore appears that at high concentrations, NC strongly destabilizes and even solubilizes the bilayer in some places. Thus, NC leads to the destabilization of DOPS bilayer as well as DOPC:DOPS, suggesting that the major determinant of lipid binding and destabilization is the negative charge of the DOPS lipids.

Interestingly, for both pure DOPS and 1:1 mixed DOPS:DOPC lipid bilayer, large scale topographic images showed clearly differences between areas where several scans were processed and the neighboring areas where only a single scan was performed (Figure 2.27, black dashed box). These differences of morphology could be explained by the interaction between the AFM tip and the lipid bilayers, leading to a spreading of the lipids on the surface of the mica. This artifact was not observed in the absence of NC. One possible interpretation is that at high concentrations, NC changes the mechanical properties of the lipids, so that they become sensitive to the interaction with the tip.

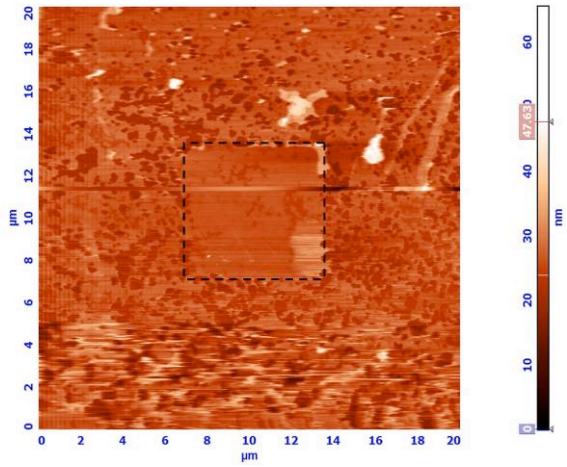


Figure 2.27. AFM image of a DOPS bilayer after addition of 1µM NC. The area inside the black dashed box corresponds to a repetitive scan.

As a next step, the three NC mutants were tested on supported lipid bilayers of DOPC:DOPS (1:1) composition to compare their ability of binding and membrane disruption. NC_{H23C} and NC_{SSHS} mutants showed similar destabilisation of the supported lipid bilayers as the native NC. Defects of surface started to appear in the same concentration range (0.1µM). In contrast, NC(11-55) mutant did not change the surface of supported lipid bilayers even for high concentration of peptide (3µM). These data are in agreement with DLS investigations presented above, suggesting that aggregation of NC with lipids strongly relies on the positive charges of the NC peptide.

Destabilization of the lipid bilayer by NC peptide (two-photon microscopy studies)

In order to confirm the ability of NC to permeate and disrupt lipid membranes, we performed experiments using two-photon fluorescence microscopy on giant unilamellar vesicles (GUVs). To visualize GUVs, we labeled them with F2N12SM probe based on 3-methoxychromone which is a fluorescent environment-sensitive membrane probe recently developed in the laboratory. For this experiment, we prepared GUVs composed of DOPS:DOPC (60:40) lipids and labeled them with 0.4 µM F2N12SM probe (Figure 2.28A).

To observe the effects of NC on the labeled GUVs, NC labeled with tetramethylrhodamine (Rh-NC) was added at increasing concentrations to the labeled GUVs. This double labeling strategy allowed distinguishing GUVs with and without NC attached. Binding of NC to the membrane was observed after incubation of GUVs with 0.4 µM Rh-NC. However, no detectable alteration of the GUVs occurred with this NC concentration. When a higher concentration of peptide (1 µM) was added, an aggregation of GUVs induced by NC was observed. More dramatic changes in the morphology of DOPS:DOPC GUVs occurred as a result of addition of 2 µM of Rh-NC (Figure 2.28B, C). At this concentration, spherical structures likely corresponding to proteolipidic particles were observed at the surface of the GUVs. After a few minutes, complete rupture of the membrane occurred (data not shown). Thus, NC binds to the membrane surface, which leads to GUV aggregation and membrane disruption.

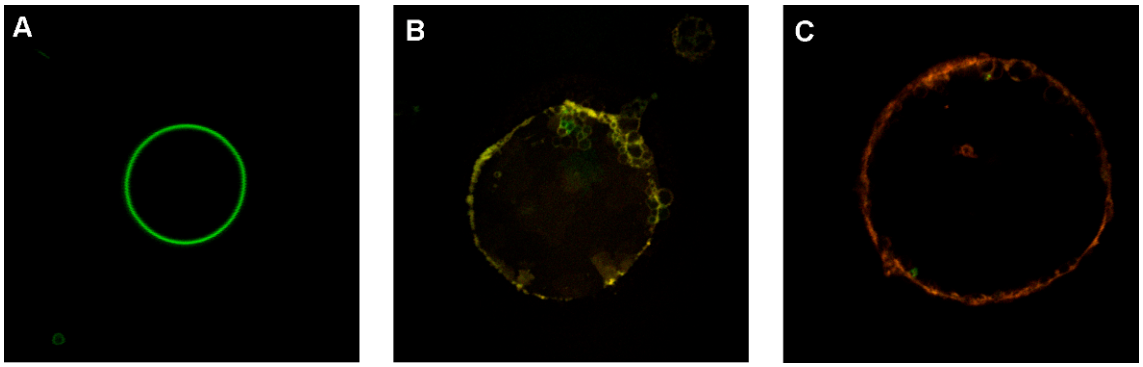


Figure 2.28. Two-photon fluorescence microscopy imaging of DOPS:DOPC (6:4) GUVs labeled with 0.4 μM F2N12SM probe in the absence (A) and in the presence of Rh-NC (2μM) (B, C). Two-photon excitation wavelength was at 830 nm. Sizes of the images were 70 μm x 70 μm.

Interaction of NC-DNA complex with membrane

Since NC is likely bound to the viral DNA in the PIC, it is important also to investigate the interaction of NC-DNA complexes with the membrane. As a model DNA sequence, we used the dTAR sequence (55 bases) corresponding to the equivalent DNA transactivation sequence of HIV-1, located at the end of the non-coding viral RNA.

First, we investigated the ability of the complex to interact with membrane by fluorescence spectroscopy. We used the NC peptide labeled with the MFL label. As it was described before, the labeled NC in the buffer showed a single red-shifted emission band. Upon addition of increasing amounts of dTAR, the maximum of the emission band was gradually shifted (Figure 2.29) from 546 nm to 532 nm, a value close to that of the free label in alcohols (λ_{em} = 533 nm of the label in MeOH). This blue shift indicates a decrease in the polarity of the label environment due to the close proximity to DNA bases, suggesting a screening of the label from water in the NC-dTAR complex. The progressive blue shift observed with increasing dTAR concentrations may have two origins. The first one is that at low dTAR concentrations, there is still a mixture of bound and free NC species. The presence of free NC shifts the spectrum to the red. The second reason is that cTAR contains about 8 nonequivalent binding sites for NC (Shvadchak, Klymchenko et al. 2009; Bazzi, Zargarian et al. 2011). At high concentrations of dTAR (ratio of TAR to NC = 1:1 or 2:1), the NC will only bind to the highest affinity sites of cTAR. In this high affinity site, the label seems to be well screened from the solvent, as shown by the most blue-shifted spectr at these concentrations. In contrast, at lower dTAR concentrations, NC is in excess over dTAR and likely occupies also lower affinity sites. NC likely binds to these sites mainly by electrostatic interactions, which does not enable the probe to intercalate with the bases, leaving it in a solvent exposed environment, as suggested by the rather red-shifted spectra in these conditions. In further experiments, we used a ratio of three NC per one dTAR, since at this ratio NC appears fully bound to DNA, as suggested by the blue-shifted spectrum at this ratio (Beltz, Clauss et al. 2005).

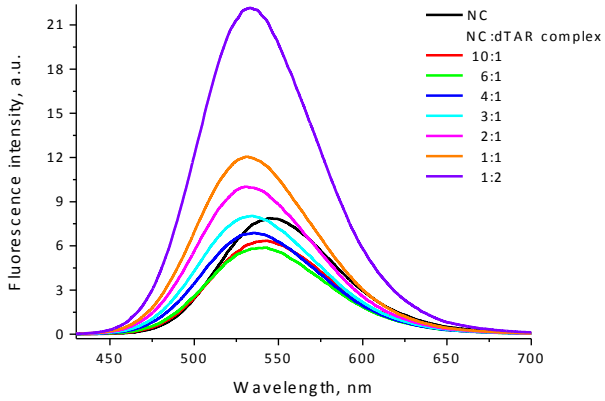


Figure 2.29. Fluorescence spectra of the MFL-NC alone and added to dTAR at different molar ratios. Experiments were performed in 20 mM phosphate buffer pH 7.4, 150 mM of NaCl. Concentration of peptide was 0.3 μM. Excitation wavelength was 400 nm.

Addition of negatively charged LUVs (DOPS or DOPG) to the NC-dTAR complex caused an increase of the fluorescence quantum yield and appearance of two separated N* and T* bands in the emission spectra (Figure 2.30). The obtained spectrum was of twice lower intensity than the one of the same concentration of NC added to DOPS, in the absence of dTAR, indicating that at least, some fraction of the NC-DNA complexes bind to the DOPS vesicles, without being dissociated. The lower quantum yield of NC-DNA/DOPS as compared to NC/DOPS suggests that the label may be more exposed to water, and thus more quenched in the former.

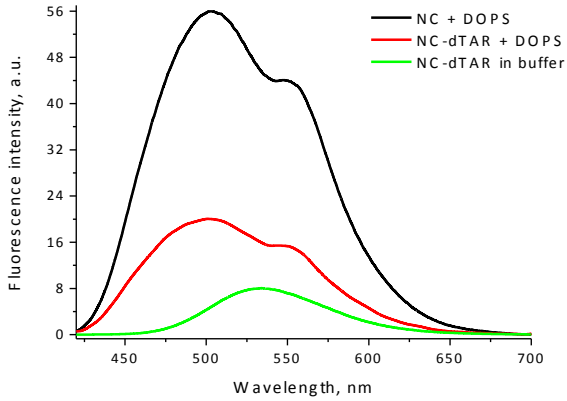


Figure 2.30. Fluorescence spectra of the complexes of MFL-NC alone (black) and MFL-NC/dTAR (red) with DOPS LUVs. The spectrum of MFL-NC/dTAR in the absence of vesicles is given as a reference (green). Experiments were performed in 20 mM phosphate buffer, pH 7.4 containing 150 mM of NaCl. Concentration of peptide and lipids was 0.3 and 240 μM, respectively. Excitation wavelength was 400 nm.

Further, we studied the localization of NC in complex with dTAR bound to the membrane. The relatively high N*/T* band ratio of MFL-NC/dTAR bound to LUVs points that the label is located in an environment of rather high polarity, suggesting a localization of the complex on the water-membrane interface. To compare the insertion depth of the N-terminus of NC alone and in

complex with DNA, we performed parallax quenching experiments in DOPS:DOPC (60:40) vesicles. By using the same procedure as described above, we found that the N-terminus of NC in complex with dTAR locates 18.4 Å from the center of the bilayer, which points about a more shallow location of the label in comparison to NC alone.

To study the possible destabilization of the lipid bilayer by the NC-dTAR complex, we used AFM and two-photon microscopy. By using AFM, we first characterised the morphology and size distribution of the complexes. Figure 2.31A represents typical NC-DNA complexes homogeneously distributed on an atomic flat freshly cleaved mica surface. Two zoomed particles are shown on figure 2.31B, from which the height of complex was observed as ~4 nm (Figure 2.31C). The same height value for the NC-dTAR complexes was obtained by the grain analysis.

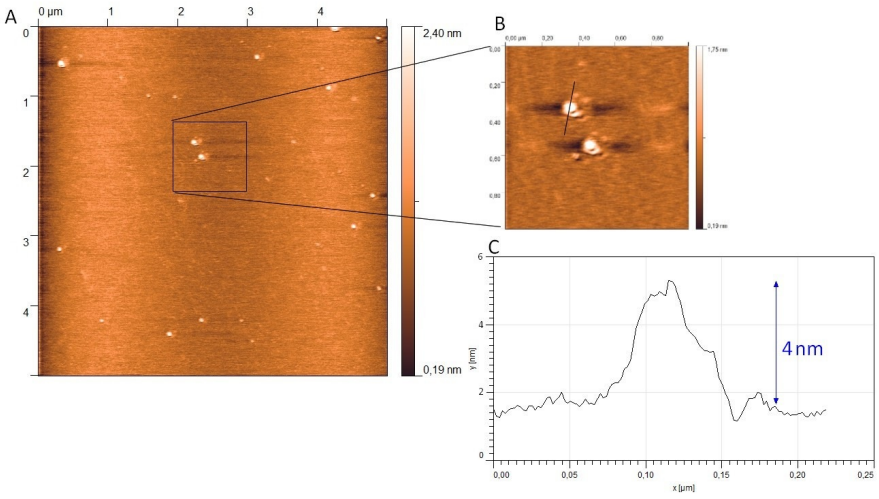


Figure 2.31. AFM image of NC-dTAR complex (3:1) on mica (A). Zoomed area of two particles of the image A (B). Profile of the section represented by the black line on the image B (C).

Then, AFM imaging of the NC-DNA complexes was performed on supported lipid bilayers (DOPS). The bilayer (Figure 2.32A) appeared as continuous, but with some islands of lipid bilayers on it. Addition of 0.1µM NC-DNA to the DOPS bilayer had a very limited effect on the bilayer. Figure 2.32B shows round-shaped particles of ~45 nm in diameter (Figure 2.32B) at the surface of the DOPS bilayer. These particles were not observed when NC was interacted with DOPS vesicles, suggesting that these particles correspond to NC-DNA complexes bound to the bilayer. Complexes are 2-3 nm in height on average, against 4 nm for the complexes adsorbed on mica. This decrease in height suggests that the complexes could penetrate to a depth of 1-2 nm in the bilayer or alternatively, are partly dissociated when they bind to the membrane. Comparison of the pictures (Figure 2.26C and 2.32B) of NC alone and NC-DNA complex at 0.1µM concentration showed their different behavior on the lipid bilayer. At these conditions no detectable destabilization of the bilayer was observed upon addition of complex in contrast to NC alone. At higher concentration (1µM), the effect of NC-DNA complex on the bilayer is much more pronounced. After addition of the complex, a deformation of the lipid islands was observed over time, with the appearance of several holes on the lipid bilayer (Figure 2.32C, D). The AFM images suggest that the NC-DNA

complexes locally destroy the bilayer, and this local destruction increases with time. A similar behavior was observed with DOPS:DOPC (1:1) bilayer (data not shown).

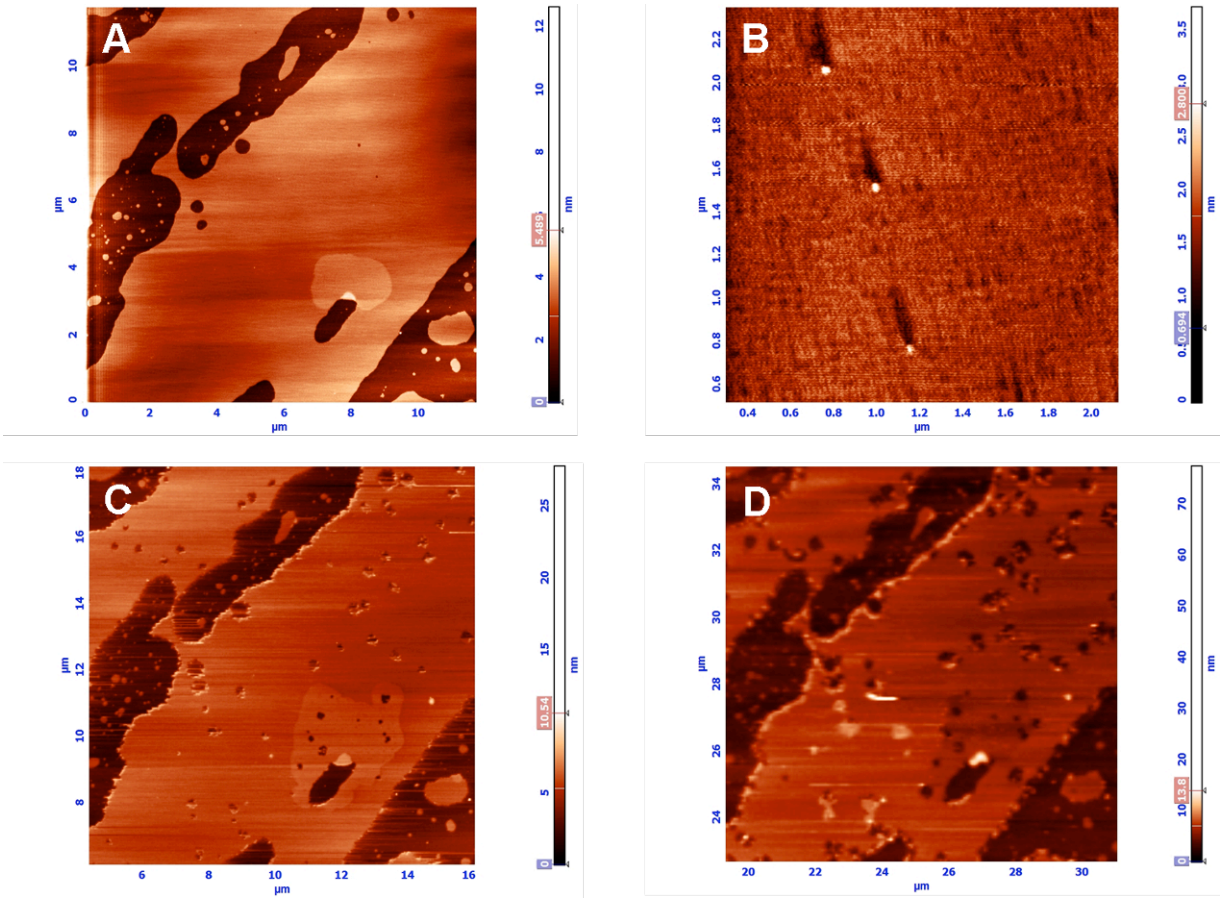


Figure 2.32. AFM images of DOPS bilayer recorded prior (A) and after addition of 0.1µM NC-DNA complex (B), 1µM NC-DNA complex after 5 min (C) and 20 min of incubation (D).

Afterwards, a DOPC bilayer was used to check the ability of NC-dTAR complexes to interact with neutral lipids. After addition of 0.1µM complex to DOPC bilayer, protein-DNA complexes were observed as round shaped particles adsorbed on the DOPC bilayer with a spherical diameter of 43 nm (Figure 2.33). This large size suggests that the NC-DNA complexes bind to the bilayer and probably form ternary complexes with the lipids. Interestingly, no change was observed in the morphology and surface structure of bilayer in these conditions. Similarly, addition of 1µM NC-dTAR complex to DOPC bilayer did not cause destabilization of bilayer. Thus, interaction of NC with DNA was found to strongly increase its ability to interact with neutral membranes.

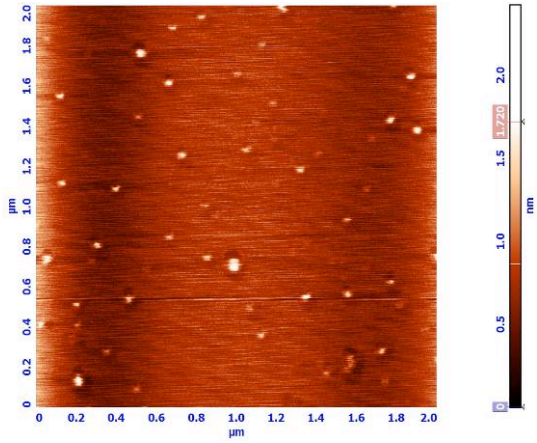


Figure 2.33. AFM images of 0.1μM NC-DNA complexes on supported lipid bilayers (DOPC).

By using two-photon microscopy, we found the behaviour of NC-dTAR complexes similar to that of the NC peptide when added to GUVs composed of DOPS:DOPC (60:40) lipids and labeled with F2N12SM probe (0.4 μM). Some GUVs appeared with small vesicles inside. NC-dTAR complexes were prepared with NC labeled with tetramethylrhodamine (Rh-NC) and unlabeled dTAR in a ratio 3:1. After addition of a small concentration of complex (up to 1 μM), binding of the complexes to GUVs was observed but no change in the lipid organization of GUVs was detected. Addition of 1μM complex resulted in aggregation of GUVs (Figure 2.34A). A further injection of NC-dTAR complexes up to a 2μM concentration induced changes in the shape and size of the GUVs (Figure 2.34B). Moreover, visible membrane disruption started after 40min of incubation of the complexes with GUVs (Figure 2.34C). Thus, AFM and two-photon imaging of GUVs further indicated that NC bound to ODNs can destabilize and disrupt in a concentration- and time-dependent manner the lipid bilayer.

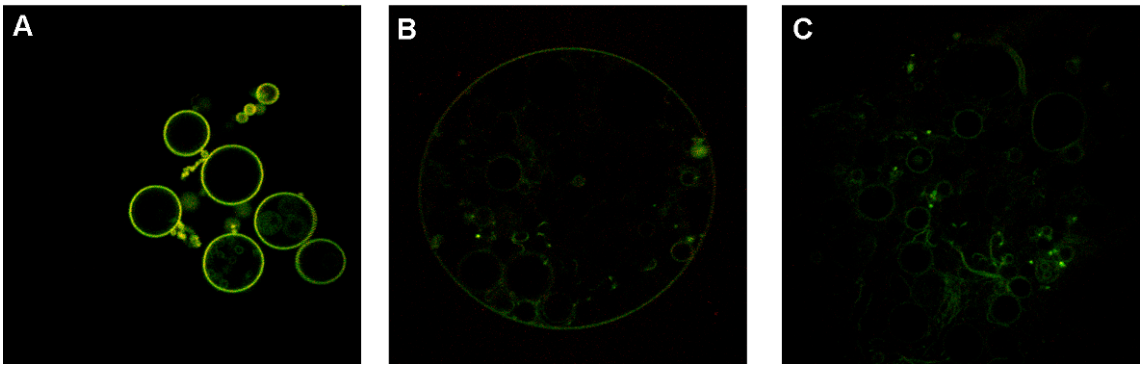


Figure 2.34. Two-photon fluorescence microscopy imaging of Rh-NC/dTAR complexes bound to DOPS:DOPC (6:4) GUVs labeled with 0.4 μM F2N12SM. Different concentrations of complex were used: 1μM Rh-NC + 0.33 μM dTAR after 30 min of incubation (A) and 2μM Rh-NC + 0.66 μM dTAR after 25 min (B) and 40 min (C) of incubation. Two-photon excitation wavelength was at 830 nm. Sizes of the images were 70 μm x 70 μm.

In order to characterize the diffusing properties of NC bound to supported lipid bilayers deposited on mica, single particle tracking was used. To monitor the diffusion of the protein, a NC N-terminally labeled with tetramethylrhodamine (Rh-NC) was used. The experiments were performed with DOPS:DOPC (1:1) supported lipid bilayers. In order to be able to track single particles, a concentration of 5 pM Rh-NC and 0.1 µM non-labeled NC was used. Ten trajectories were analyzed to extract the diffusion constant and the hydrodynamic radius of the diffusion particles. An average diffusion coefficient of 0.32 $\mu m^2 s^{-1}$ and a hydrodynamic radius of 13 nm were obtained. These observations show that the diffusing species are composed of several NC peptides. Thus, this confirms that NC molecules likely form lipoprotein complexes with lipids, and diffuse in this form in the membrane. Further, the same experiments were repeated in the presence of dTAR (NC:dTAR ratio as 3:1). In these conditions, the diffusion constant and the hydrodynamic radius were modified to values of 0.06 $\mu m^2 s^{-1}$ and 60 nm, respectively. Again, the obtained data support that NC-DNA complexes interact with membrane. Moreover, the strong increase of the hydrodynamic radius of the diffusing complexes is consistent with the AFM results, suggesting that particles are composed of multiple several NC-DNA complexes forming large particles with lipids.

Based on our data, NC was shown to bind with high affinity to lipid membranes containing negatively charged lipids. This interaction is mainly driven by electrostatic forces and locates the protein at the level of the lipid heads. At high concentrations, NC is able to destabilize the membrane. Moreover, our data further showed that NC-DNA complexes can also bind and destabilize the membrane similarly to free NC. All these results support the hypothesis that NC may play a role in the nuclear entry of PIC.

On the basis of this work, two hypotheses for the role of NC in the nuclear internalization of PIC could be proposed. In the first one, the peptide in its free form binds to the nuclear membrane and destabilizes it transiently, allowing the entry of the PIC into the nucleus (Figure 2.35). The hypothesis that NC exists in free form, being released by the PIC before nuclear entry is consistent with the work of Mirambeau (Mirambeau, Lyonnais et al. 2006), where it was shown that NC shows a lower affinity for the proviral DNA than for the viral RNA, resulting in the release of NC after the conversion of the reverse transcription complex in the PIC. In this respect, NC could cooperate together with Vpr to destabilize the nuclear membrane, since Vpr is also present in the PIC and also exhibits destabilizing properties (de Noronha, Sherman et al. 2001). Data on the interaction of NC-DNA complexes with lipid bilayers also suggest that NC could be released during the interaction with membranes and thus destabilize them. This release of NC from NC-DNA complexes may result from the higher affinity of NC for lipids than for DNA. The second hypothesis is based on the ability of NC-DNA complexes and thus the PIC to bind to the nuclear membrane (Figure 2.35). This interaction could be a triggering event initiated in cooperation with the matrix and Vpr proteins, which allows the entry of PIC through a nuclear pore (Shahin, Hafezi et al. 2006).

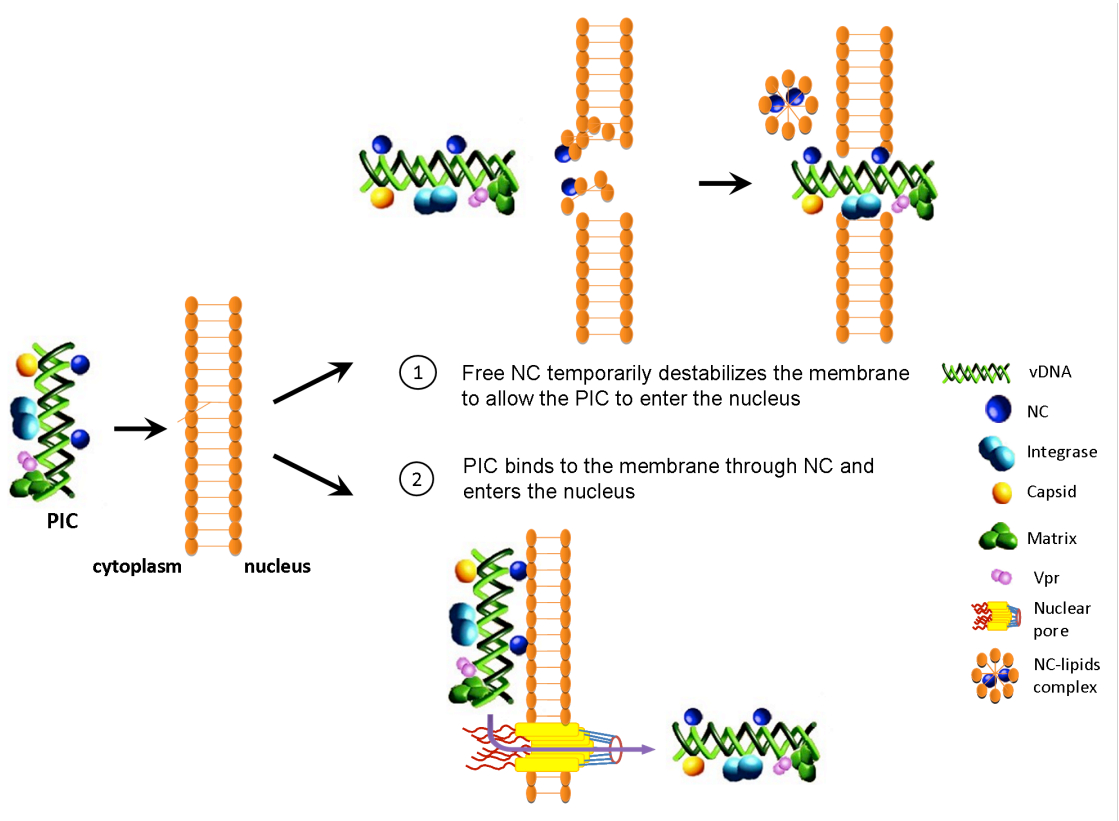


Figure 2.35. Scheme of the two hypotheses of the role of NC in the nuclear internalization of the PIC.

CHAPTER 3

CONCLUSIONS AND PERSPECTIVES

3. CONCLUSIONS AND PERSPECTIVES

Conclusions

The aim of this work was to develop a methodology for sensing interactions of proteins with oligonucleotides and membranes based on environment-sensitive 3-hydroxychromone (3HC) labels which due to their excited state intramolecular proton transfer (ESIPT) and charge transfer present a strong sensitivity of their dual emission to solvent polarity and hydration.

The first part of my work was focused on the development of two-color 3-hydroxychromone labels for monitoring peptide-oligonucleotide interactions. A fluorescent L-amino acid analogue based on 2-(2-furyl)-3-hydroxychromone dye was synthesized and inserted at three different positions of the NC(11-55) peptide in order to reveal the role of individual residues in peptide-ODN interactions. The peptides bearing the 3HC amino acid were found to keep their native folding and activity. Interaction of the labeled peptides with oligonucleotides, whose structures in complex with NC have been solved, changed dramatically the dual emission of the introduced amino acid. Moreover, we found that the ratiometric response of the fluorescent amino acid analogue correlates well with its proximity to the closest base in the complexes. Therefore, this amino acid analogue provides a new and universal approach for site-specific monitoring of peptide-nucleic acid proximity, opening a broad range of applications in the characterization of peptide-nucleic acid interactions.

In parallel, to overcome the drawbacks of the 3HC label, three optimized probes for N-terminal labeling have been synthesized. These labels were covalently linked to the N-terminus of the Tat(44-61) peptide and used for sensing the binding of ODNs of different size, composition and strandedness. All new labels showed larger sensitivity of their dual emission to peptide-ODN interactions compared to the parent label, which can be explained by their higher sensitivity to hydration. Thus, the new generation of hydration-sensitive probes appears as highly sensitive ratiometric tools to site-selectively monitor the binding of peptides to ODNs. Moreover, we developed a new approach for estimating the local hydrogen bond donor concentration using the developed 3-hydroxychromone labels.

In collaboration with the team of A. Burger (Nice), new fluorescent nucleoside analogues were synthesized where the natural base is substituted by a 3HC dye. Fluorescence studies showed that the nucleoside analogue featuring a 3HC-thienylchromone displays high sensitivity of its dual emission to polarity changes and therefore is highly promising for nucleic acid labeling. As a next step, this nucleoside analogue was introduced in different positions of various ODNs. Afterwards, UV melting curves showed that the nucleoside analogues marginally affect the stability of the labeled ODNs, and thus, could be used as a universal base for ODN labeling. The absorption and fluorescence properties of 3HC-thienylchromone in the ODNs either in their single-strand or duplex states as well as after interaction with NC peptide were characterized. Taken together, our data suggest that the labeling of short ODNs at selected positions provided site-specific multi-channel information on local environment changes induced by the interaction with proteins of interest.

The second part of my work consisted in developing an approach for monitoring the binding of peptides to lipid membranes, using a fluorescent environment-sensitive label of the 3-hydroxyflavone (3HF) family. For this purpose, we selected the 4'-(dimethylamino)-3-

hydroxyflavone fluorophore adapted to aprotic apolar media. To validate its ability to sense peptide-membrane interactions, we labeled the N-terminus of three synthetic peptides, melittin, magainin 2 and poly-L-lysine capable to interact with lipid membranes. Binding of these labeled peptides to lipid vesicles induced a strong fluorescence increase, which enabled to quantify the peptide-membrane interaction. Moreover, the dual emission of the label in these peptides correlated well with the depth of its insertion measured by the parallax quenching method. We found that the N-terminus of melittin and magainin 2 is immersed into the bilayer, while that of poly-L-lysine is localized at the membrane surface. Fluorescence microscopy measurements in giant vesicles further revealed differences in the orientation of the label bound to these two peptides. Thus, this label constitutes an interesting new tool for monitoring membrane binding and insertion of peptides.

As a next step, we synthesized a NC peptide N-terminally labeled with the MFL dye for investigating the possible interaction of NC with the membrane. The properties of the labeled peptide were found to be close to those of the native one. The two-color fluorescence of the labeled NC was exquisitely sensitive to the binding to lipid membranes containing negatively charged lipids. In contrast, neutral LUVs do not affect the fluorescence properties of MFL-NC due to lack of interaction. Moreover, we studied the orientation of NC bound to the membrane by using the ratiometric response of the label and found that the protein locates at the level of the lipid heads, which was proved by parallax experiment. Interestingly, interaction of NC with DNA was found to strongly increase its ability to interact with neutral membranes (DOPC). Formation of protein-DNA-membrane complexes was confirmed by atomic force microscopy on supported lipid bilayers. Furthermore, AFM and two-photon imaging of giant unilamellar vesicles further indicated that NC in its free form or bound to ODNs can destabilize and disrupt in a concentration- and time-dependent manner the lipid bilayer. Taken together, our data indicate that both free NC and NC-nucleic acid complexes could interact and destabilize lipid membranes, suggesting that NC could participate in the nuclear import of the pre-integration complex. On the basis of these results, two hypotheses for the role of NC in the nuclear internalization of PIC have been proposed.

Taken together, our data show that fluorescent labels based on 3HC dyes constitute particularly interesting multi-channel tools for monitoring biomolecular interactions involving peptides or proteins.

Perspectives

The present work on the development and application of new methodologies based on 3-hydroxychromone dyes for monitoring biomolecular interactions could be extended in different ways. For instance, these probes could be further used to characterize the interaction of NC with its nucleic acid targets and thus, further characterize the nucleic acid chaperone properties of NC. Using NC labeled with the synthesized 3HCaa fluorescent amino acid analogue, its dynamic interaction with the HIV-1 primer binding site (PBS) and the transactivation response element (TAR) could be investigated in depth. This could be achieved by using NC peptides labeled at different positions and time-resolved techniques. Further important complementary information is expected by using PBS and cTAR labeled with 3HC-based nucleosides. These labeled sequences will allow monitoring site-selectively the structural changes provided by NC at the level of the ODNs. Comparison of the data at the level of the peptide and the ODN should allow us to further characterize the molecular mechanisms of NC nucleic acid chaperone activity.

The developed probes could also be used to further characterize NC-membrane interaction. Through the fluorescence lifetimes which are absolute parameters independent on the local concentration, it will be possible to demonstrate a direct interaction of NC with both model and cell membranes. By using FRET and FLIM-FRET techniques, it will also be possible to further determine the nature of the aggregates (lipoprotein particles) that appeared subsequently to the interaction of NC with the supported lipid bilayer. In addition, these techniques will evidence whether NC-DNA complexes dissociate or remain stable after binding to a lipid bilayer. Moreover, the AFM/TIRF technique developed in our laboratory that combines the sensitivity of fluorescence with the topographic information and resolution provided by AFM, will allow unambiguously monitoring the changes in the structure and the properties of the bilayer upon binding of NC or NC-DNA complex. Next, the interaction of NC alone and in complex with DNA with fragments of nuclear membrane (Oberleithner, Brinckmann et al. 1994) will be investigated to check the hypotheses made on the basis of the experiments performed with the supported lipid bilayers. Finally, collaboration with a virology group will be established to demonstrate the functional relevance of NC-membrane interactions in infected cells. These data will demonstrate for the first time a new property of NC that is likely important for its functions in the retroviral cycle and thus, initiate a new research axis on the protein.

Another important direction of this work will be to further develop the environment-sensitive labels for investigating the interaction of proteins with different targets.

Protein dynamics and folding as well as interactions between proteins, nucleic acids, membrane and other biomolecules could be demonstrated at the single-molecule level. To meet the strict requirements for single-molecule imaging, the dyes should present low photobleaching, high quantum yields and large molar extinction coefficients. Commonly used dyes frequently do not fit to all these requirements. An additional disadvantage of these dyes is their insensitivity to the environment. Thus, design of new environment-sensitive labels with improved properties is highly demanded for site-specific investigation of protein interactions at the single-molecule level. One prospective direction in this field could be use of fluorophores based on 3-methoxychromone, recently developed in our laboratory. In comparison to the parent 3-hydroxychromone derivatives, the new fluorophores showed improved photostability and spectroscopic properties, serving as

attractive building block for the development of environment-sensitive labels for the protein and oligonucleotide labeling.

Improvement of the labels could be also achieved through the synthesis of non-natural amino acid based on 4'-(dimethylamino)-3-hydroxyflavone to further investigate the peptide-membrane interactions. The peptides will be labeled at different positions in order to investigate the involvement of different residues in the interaction. Moreover, site-specific information on the protein conformation and its position in the membrane could be obtained. Introduction of the dye close to a putative interaction site should induce a strong change in its signal, allowing a site-selective characterization of its environment within the binding site.

CHAPTER 4

MATERIALS AND METHODS

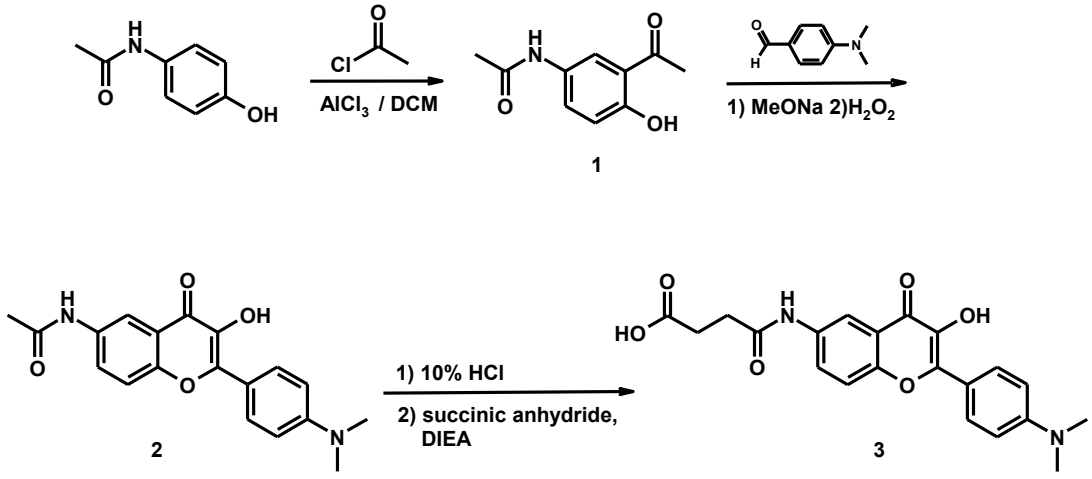
4. MATERIALS AND METHODS

4.1. Synthetic procedures

All the solvents and chemicals were purchased from Sigma-Aldrich. The solvents for spectroscopy were of analytical grade. NMR spectra were recorded on a Bruker Avance III, BBFO+ spectrometer at frequency 300-500 MHz at room temperature. Mass spectra were obtained on a Bruker HCT Ultra and Agilent Technologies Accurate-Mass Q-TOF LC/MS 6520 mass spectrometers.

4.1.1. Synthesis of fluorophores

The fluorescent amino acid analogue (3HCaa) was synthesized by Aleksandr Strizhak (synthesis was described in **Publication 1**). Synthesis of F6A, F6C and F4O labels was performed by Oleksandr Zamotaiev (**Publication 2**). Fluorescent nucleosides bearing 3-hydroxychromone nucleobases were synthesized by Marie Spadafora and Dmytro Dziuba from Université de Nice Sophia Antipolis (**Publication 4, 5**). Synthesis of MFL label was performed by using next procedures (Scheme 1):



Scheme 1. Scheme of the synthesis of the MFL label.

N-(3-acetyl-4-hydroxyphenyl)-acetamide (1). To a stirred ice-cooled solution of N-(4-hydroxyphenyl)-acetamide (15 g, 0.091 mol) in 130 mL of dry CH_2Cl_2 in Ar-atmosphere 16.15 mL (0.23 mol) of acetyl chloride was added. After 5 minutes of stirring 30.3 g (0.23 mol) of $AlCl_3$ was added in small portions during 30 min. Then slightly violet solution was boiled during 5 h at 50°C. After 1h of boiling, dark precipitate was formed. The reaction was monitored using TLC (CH_2Cl_2:MeOH, 95:5). Then, the reaction mixture was poured into ice-water and filtered. The slightly green crystals were obtained after recrystalization from ethanol, yield 11.88 g (67.8 %). ^{1}H NMR (200MHz, $CDCl_3$) δ 12.09 (s, 1H), 8.16 (d, 1H, J = 2.45 Hz), 7.32 (dd, 1H, J_1 = 8.81 Hz, J_2 = 2.45 Hz), 6.91 (d, 1H, J = 8.81 Hz), 2.6 (s, 3H), 2.16 (s, 3H).

6-acetamido-4'-(dimethylamino)-3-hydroxyflavone (2). N-(3-acetyl-4-hydroxyphenyl)-acetamide (1) (15.5 mmol, 3 g) and 4'-dimethylaminobenzaldehyde (17.1 mmol, 2.55 g) were dissolved in 20 mL of dry DMF. Finally, sodium methoxide (92.6 mmol, 5 g) was added in two portions. The mixture was stirred for 24 hours at 60^{o}C. Then, the reaction mixture was diluted with 100 mL of ethanol followed by addition of 10 mol excess of hydrogen peroxide (16 mL) and 15 mol excess of sodium methoxide (14 g). The reaction mixture was refluxed for 10 min. After cooling, the mixture was poured into water and neutralized with conc. HCl to pH = 6-7. The formed precipitate was filtered off, washed with water, and dried under vacuum. The yellow solid was purified by crystallization from methanol/ethanol mixture to give 6-acetamido-4'-(dimethylamino)-3-hydroxyflavone (2.8 g, 54 %) as yellow crystals. ^{1}H NMR (400 MHz, DMSO-d6) δ 10.22 (s, 1H), 9.12 (s, 1H), 8.39 (d, 1H, J = 2.51 Hz), 8.11 (d, 2H, J = 9.03 Hz), 7.87 (dd, 1H, J_1 = 9.03 Hz, J_2 = 2.51 Hz), 7,68 (d, 1H, J = 9.03 Hz), 6.85 (d, 2H, J = 9.03 Hz), 3.02 (s, 6H) , 2.10 (s, 3H); m/z ($M+H^+$) calculated for $C_{19}H_{19}N_2O_4$: 339.1; found: 339.0.

MFL label (3). 6-acetamido-4'-(dimethylamino)-3-hydroxyflavone (2) (3 mmol, 1 g) was refluxed in 10% HCl (30 mmol, 11 mL) for 7h. Then, the reaction mixture was concentrated in vacuo and dissolved in 15 ml dry THF. Then, 0.3 g (3 mmol) of succinic anhydride and 2.6 mL of diisopropylethylamine (15 mmol) were added to this solution. The reaction mixture was stirred overnight at 60^{o}C. Then the solvent was evaporated and the residue was treated with water. The precipitate was filtered off and washed with water. The product was recrystallized from propanol-2 to give 1 g (yield 84%) of the final label 3. ^{1}H NMR (300 MHz, DMSO-d6) δ 12.5-12.0 (1H, br s), 10.24 (s, 1H), 9.12 (s, 1H), 8.40 (d, 1H, J = 2.64 Hz), 8.09 (d, 2H, J = 9.04 Hz), 7.84 (dd, 1H, J_1 = 9.04 Hz, J_2 = 2.64 Hz), 7,67 (d, 1H, J = 9.04 Hz), 6.84 (d, 2H, J = 9.04 Hz), 3.01 (s, 6H) , 2.65-2.50 (m, 4H); m/z ($M+H^+$) calculated for $C_{21}H_{21}N_2O_6$: 397.1; found: 397.1.

4.2. Peptide synthesis

4.2.1. General methods

- *Synthesis*

Peptides were synthesized by solid phase peptide synthesis on a 433A synthesizer (ABI, Foster City, CA). The synthesis was performed at a 0.1 mmole scale using the standard side-chain protected fluorenylmethoxycarbonyl (Fmoc)-amino acids and HBTU/HOBt coupling protocol. LL-HMP resin (ABI, 0.44 mmol/g reactive group concentrations) was used as solid support. Deprotection steps were performed by piperidine and automatically controlled by UV absorbance. At the end of the synthesis, the peptidylresin was isolated, and twice washed by NMP.

- *Labeling*

Labeling procedures were performed in flasks. Two-four equivalents (0.2-0.4 mmol) of the label were dissolved in 1 mL of NMP mixed with two-four eq. of HBTU/HOBt coupling solution (in DMF) and added to Fmoc-deprotected peptidylresin (0.1 mmol) swelled in 1 mL of NMP. After a few minutes of shaking four eq. of DIEA solution were added. Then, the reaction mixture was stirred overnight at 37°C. The resin was filtrated, washed by NMP, methanol, DCM and dried.

- *Cleavage and deprotection*

Cleavage of the peptidylresin and its deprotection were performed for 2 h using a 10 mL trifluoroacetic acid (TFA) solution containing water (5%, v/v), phenol (1%, w/v), thioanisole (5%, v/v) and ethanedithiol (2.5%, v/v). The peptide was precipitated using ice-cold diethyl ether and pelleted by centrifugation. The pellet was air-dried for approx. 15 minutes, solubilized with aqueous TFA (0.05 %, v/v) and lyophilized.

- *Purification*

Before purification, the peptides were solubilized with aqueous TFA (0.05 %, v/v). HPLC purification was carried out on a C8 column (uptisphere 300A, 5µm; 250X10, Interchim, France) in water/acetonitrile mixture containing 0.05% TFA with linear gradients depending on the peptide (typically 10 to 50% of acetonitrile for 30 min) and monitored at 210 nm (detection of all peptides including non-labeled) and 370 nm (detection of labeled peptides only). After purification, the fractions containing pure peptide were combined and lyophilized.

All peptides were characterized by ion spray mass spectrometry and the expected molecular masses were found. Prior to use, the peptides were dissolved in distilled water, aliquoted and stored at -20°C.

4.2.2. Synthesized peptides

Table 7. Synthesized peptides, their sequences and molecular masses (MS)

Peptide	Sequence	MS
Tat(44-61)	GISYGRKKRRQRRRPPQG	2197
FC-Tat	**FC**-GISYGRKKRRQRRRPPQG	2523
F6C-Tat	**F6C**-GISYGRKKRRQRRRPPQG	2491
F6A-Tat	**F6A**-GISYGRKKRRQRRRPPQG	2562
F4O-Tat	**F4O**-GISYGRKKRRQRRRPPQG	2491
MFL-melittin	**MFL**-GIGAVLKVLTTGLPALISWIKRKRQQ	3225
MFL-magainin 2	**MFL**-GIGKFLHSAKKFGKAFVGEIMNS	2844
MFL-poly-L-lysine	**MFL**-KKKKKKKKKKKKKKKKKKKKK	3017
NC(1-55)	MQRGNFRNQRKNVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	6427
MFL-NC(1-55)	**MFL**-MQRGNFRNQRKNVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	6805
Rh-NC(1-55)	**Rh**-MQRGNFRNQRKNVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	6839
NC(11-55)	KNVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	5138
MFL-NC(11-55)	**MFL**-KNVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	5516
NC(1-55)-SSHS	MQRGNFQRKNVK**S**FN**S**GKEGHTARN**S**RAPRKKG**S**WK**S**GKEGHQMKD**S**TERQAN	6060
MFL-NC(1-55)-SSHS	**MFL**-MQRGNFQRKNVK**S**FN**S**GKEGHTARN**S**RAPRKKG**S**WK**S**GKEGHQMKD**S**TERQAN	6438
NC(1-55)-H23C	MQRGNFRNQRKNVKCFNCGKEG**C**TARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	6392
MFL-NC(1-55)-H23C	**MFL**-MQRGNFRNQRKNVKCFNCGKEG**C**TARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	6770
NC(11-55)-F16-3HCaa	KNVKC(**3HCaa**)NCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQAN	5287
NC(11-55)-A30-3HCaa	KNVKCFNCGKEGHTARNCR(**3HCaa**)PRKKGCWKCGKEGHQMKDCTERQAN	5360
NC(11-55)-W37-3HCaa	KNVKCFNCGKEGHTARNCRAPRKKGC(**3HCaa**)KCGKEGHQMKDCTERQAN	5245

4.2.3. Preparation of Zn-bound peptides

The complexation of zinc with NC was achieved by adding a concentrated solution of $ZnSO_4$. The lyophilized peptides were dissolved in water (~ 0.5 mg in 500 µL). Then, about 10 µL of this solution was used to determine the peptide concentration using an appropriate extinction coefficient. Next, 2.2 molar equivalents of $ZnSO_4$ were added to the peptide and pH was raised to its final value by adding buffer. This last step was done only at the end of the preparation to prevent peptide oxidization. Noticeably, in case of labels based on 3-hydroxychromone, a large excess of Zn^{2+} ions should be avoided since this ion could affect the 3HC fluorescence.

4.3. Oligonucleotides

All unlabeled oligonucleotides were synthesized and HPLC-purified by IBA GmbH (Germany). Oligonucleotides labeled with 3-hydroxychromone were provided by Alain Burger (Université de Nice Sophia Antipolis). Concentrations of the ODNs were calculated from their absorbance using the molar extinction coefficients at 260 nm specified by the supplier.

4.3.1. Preparation of the double stranded ODN

The double stranded ODNs were prepared by mixing stock solutions of both ODN strands (32μM of each strand) with 50 μL of buffer (10 mM cacodylate, 150 mM NaCl, pH 7), heating for 5 min at 80 ºC and cooling down slowly to room temperature. Then, the resulting solutions were diluted with buffer to a final volume of 800 μL to get the 2 μM concentration needed for further experiments.

4.4. Lipids and vesicles preparation

4.4.1. Lipids

Dioleoylphosphatidylcholine (DOPC), dioleoylphosphatidylethanolamine (DOPE), dioleoylphosphatidylserine (DOPS), dioleoylphosphatidylglycerol (DOPG), phosphatidylinositol (PI) and phosphatidylinositol-4,5-bisphosphate (PIP_2) were from Sigma-Aldrich. 1,2-dipalmitoyl-**sn**-glycero-3-phospho(tempo)choline (TempoPC), 1-palmitoyl-2-stearoyl-(5- and 12-doxyl)-*sn*-glycero-3-phosphocholine (5- and 12-SLPC) were from Avanti Polar Lipids (Alabaster, AL, USA). The concentration of phospholipid stock solutions in chloroform was determined by dry weight. The nitroxide content of nitroxide-labeled lipids was calculated using the electron spin resonance integrated spectra of the corresponding diluted stock solutions in chloroform by comparison with a tempocholine reference solution in the same solvent.

4.4.2. Preparation of large unilamellar vesicles by extrusion

To prepare liposomes composed of pure or mixed lipids, the lipids were first dissolved in an organic solvent (chloroform or chloroform/methanol mixtures). The organic solvent was then removed by rotary evaporation, yielding a thin lipid film on the wall of a round bottom flask. The lipid film was thoroughly dried to remove the residual organic solvent by continuing evaporation with additional heating during 30 minutes. Hydration of the dry lipid film was done by adding corresponding buffer to the flask. This step of hydration was maintained approximately 30 minutes. Typically lipid suspensions were prepared in order to obtain a 0.2 mM final concentration, although higher concentrations could be used. After hydration, the solution was thoroughly vortexed during 2

minutes, giving rise to a suspension of multilamellar vesicles (MLV) with a heterogeneous size distribution.

Once a suspension has been produced, the vesicles were downsized by extrusion. Lipid extrusion is a technique in which a lipid suspension is forced through a polycarbonate filter with a defined pore size to yield vesicles having a diameter near the pore size. An extruder (Lipex Biomembranes Inc) with polycarbonate filters of calibrated pores (Nucleopore) is used. Prior to extrusion through the final pore size, the lipid suspension was first downsized by passing through a large pore size (0.2 μm) filter seven times. Then the suspension was passed through the filter with the final pore size (0.1 μm) ten times. This final extrusion through a filter with 0.1 μm pores yields large unilamellar vesicles (LUV) with a mean diameter of 110-120 nm (according to dynamic light scattering technique). This method ensures a homogeneous size distribution of the final suspension. All preparation steps including hydration, vortexing and extrusion were done at a temperature that is higher than the temperature of the gel-liquid crystalline phase transition (Tm) of the lipid. For instance, heating at 55°C was used in the case of lipids that bear saturated alkyl chains.

4.4.3. Preparation of large unilamellar vesicles by ethanol dilution

Preparation of large unilamellar vesicles by ethanol dilution was used for quenching experiments by the parallax method. Lipid vesicles were prepared according to Kachel et al. (Kachel, Asuncion-Punzalan et al. 1998) with small modifications. DOPC (85%) and nitroxide-labeled PCs (15%) were mixed in chloroform. The mixtures were dried under vacuum for 30 min, and then resuspended in ethanol (120 μL) by continuous rotation using a rotary evaporator for 30 min. Finally, 6 ml of buffer was added to the lipid solution and the mixture was vortexed briefly. The sizes of vesicles were determined by light scattering using a N4SD Coultronics Nanosizer. The vesicles obtained by the ethanol dilution method were of similar size as those obtained by extrusion (0.12 μm).

4.4.4. Preparation of giant unilamellar vesicles

Giant unilamellar vesicles (GUVs) were generated by electroformation in a home-built liquid cell (University of Odense, Denmark), using previously described procedures (Angelova and Dimitrov 1986; Fidorra, Duelund et al. 2006; Klymchenko, Oncul et al. 2009). 1 mM solution of lipids in chloroform was deposited on the platinum wires of the chamber, and the solvent was evaporated under vacuum for 30 min. The chamber was filled with 500 μL of 300 mM sucrose solution, and a 2-V, 10-Hz alternating electric current was applied to this capacitor-like configuration for *ca* 1.5 h. Then, a 50 μL aliquot of the obtained stock solution of GUVs in sucrose (cooled down to room temperature) was added to 200 μL of 300 mM glucose solution to give the final suspension of GUVs used in microscopy experiments. The staining of GUVs was performed by addition of an aliquot of the label or labeled peptide solutions to obtain a desired concentration.

4.4.5. Preparation of supported lipid bilayers

The fusion of lipid vesicles on solid supports is the most simple and popular method for preparing supported lipid bilayers (Horn 1984; Richter and Brisson 2005). The fusion step is achieved by depositing the LUVs suspension onto freshly cleaved mica for 45–60 min at room temperature. The speed of the bilayer formation and its overall morphology strongly depend on the lipid composition and concentration, temperature, electrostatic effects, solid support and presence of divalent cations (particularly Ca^{2+}) (Richter and Brisson 2005; Richter, Berat et al. 2006).

Mica was used as a solid support for supported lipid bilayers. The two sides of a piece of mica, around 7 × 7 mm, were cleaved with an adhesive tape to obtain a flat surface. The mica was immersed in an eppendorf containing 300 µL of 200 µM LUVs solution and 300 µL of buffer (20 mM Hepes buffer, 150 mM NaCl at pH 7.4) with 4 mM $CaCl_2$. The Ca^{2+} was added to induce adsorption and rupture of the vesicles on the mica. After ~1 hour of incubation, the mica was rinsed abundantly with the buffer without $CaCl_2$ to remove unfused vesicles.

4.4.6. Sample preparation

All measurements were done in 20 mM phosphate or Hepes buffer containing 150 mM NaCl (pH 7.4) at 20 ºC. To incorporate the peptides into the model membranes, an aliquot of a peptide stock solution in buffer was added to a suspension of lipid vesicles. Then, the measurements were performed after 5 min of incubation.

4.4.7. Calculation of the vesicle concentration

To calculate the vesicle concentration, the external radius of the vesicles (R) was considered to be 535 Å, as determined by DLS measurement. The thickness of the lipid bilayer (t) and the average lipid density (d) were assumed to be 40 Å and 70 $Å^2$/lipid, respectively (Lewis and Engelman 1983). The number of lipids per vesicle was thus determined as

$$n = 4\pi*(R^2 + (R - t)^2)/d = 9.54\times10^4 \ lipids \ /vesicle \qquad (1)$$

The vesicles concentration can thus be obtained by using equation (2):

$$C(vesicles) = C \ (lipids)/n \qquad (2)$$

4.5. Physical measurements

4.5.1. Fluorescence spectroscopy

Absorption spectra were recorded on a Cary 4 spectrophotometer (Varian) and fluorescence spectra measurements were performed on a spectrofluorometer FluoroMax 3.0 or Fluorolog (Jobin Yvon, Horiba) equipped with a thermostated cuvette holder. Fluorescence emission spectra were systematically recorded at 20°C. Excitation wavelength was at the absorption maximum for each label. All the spectra were corrected from the fluorescence of the corresponding blank solution (neat solvent, lipid vesicles). Equation (3) was used for calculation of fluorescence quantum yield (QY):

$$QY_x = QY_r \frac{I_x A_r}{I_r A_x} \frac{n_x^2}{n_r^2} \quad (3)$$

where QY_x is the quantum yield of the dye; QY_r is the known quantum yield of the reference dye; I_x and I_r are their respective fluorescence intensities as measured by integration of the surface under the emission spectrum corrected for the photomultiplier response; A_x and A_r are the absorbance of the dye and the reference at an appropriate excitation wavelength, respectively; n_x and n_r are the refractive indexes of their respective solvents.

Quinine sulphate in 0.5 M sulphuric acid (QY = 0.577 (Eastman 1967)) and 4'-(dialkylamino)-3-hydroxyflavone in ethanol (QY = 0.52 (Chou, Martinez et al. 1993)) were used as references for the dyes absorbing below and above 370 nm, respectively.

4.5.2. Thermal denaturation measurements

Melting curves were recorded by following the temperature-dependence of the absorbance changes of a sample with 2 μM concentration of each of complementary strands in sample. Absorption spectra were recorded in a Peltier thermostated cell holder on a Cary 4 spectrophotometer (Varian). Wavelength for detection was 260 nm. Length of cell was 1 cm. Temperature range for denaturation measurements was 15 – 80°C. Speed of heating was 0.5°C/min. 10 mM cacodylate buffer, 150mM NaCl, pH 7.0 was used. The melting curves were converted into a plot of α versus temperature, where α represents the fraction of single-strands in the duplex state. The melting temperatures were extracted from these curves after differentiation, as described elsewhere (Breslauer 1995).

4.5.3. Photodegradation experiments

In photodegradation assays, a 2 μM solution of a given dye in a quartz micro-cuvette (50 μM volume) was illuminated by the 360 nm light of the Xenon lamp of a FluoroLog spectrofluorometer (slits were open to 8 nm). During the time of illumination (5000 seconds), the fluorescence at the maximum was recorded as a function of time.

4.5.4. Determination of the binding affinity of peptide/lipid and peptide/oligonucleotide complexes

To determine the affinity of the labeled peptides to vesicles, a fixed amount of the peptide was titrated with lipids by monitoring the two-band fluorescence of the labeled peptides. The fluorescence intensity versus the vesicle concentration was plotted and the affinity constant was determined from direct fitting of the curve by the following equation (Beltz, Piemont et al. 2004):

$$I = I_0 - \frac{(I_0 - I_t)}{P_t} x \frac{(1 + (P_t + nN_t)K_a) - \sqrt{(1 + (P_t + nN_t)K_a)^2 - 4P_t nN_t K_a^2}}{2K_a} \quad (4)$$

where I and I_t are the integrated intensity of the whole emission spectrum at a given and a saturating vesicle concentration, respectively, I_0 is the corresponding intensity in the absence of vesicles. N_t is the total vesicle concentration, P_t is the total concentration of peptide, K_a is the apparent affinity constant, and n is the number of peptides per vesicle.

To determine the affinity of the labeled peptides for the ODNs, fixed amounts of the ODN were titrated with peptides by monitoring the two-band fluorescence of the labeled peptides. For each data point, the emission of the same concentration of labeled peptide in buffer was subtracted from the signal measured in the presence of the ODN. Affinity constants were determined from direct fitting of the corrected signal to the rewritten Scatchard equation:

$$I = I_0 - \frac{(I_0 - I_t)}{N_t} x \frac{(1 + (P_t + nN_t)K_a) - \sqrt{(1 + (P_t + nN_t)K_a)^2 - 4P_t nN_t K_a^2}}{2nK_a} \quad (5)$$

where I and I_t are the signal at a given and a saturating peptide concentration, respectively, I_0 is the signal in the absence of peptide, N_t is the total ODN concentration, P_t is the total concentration of peptide, K_a is the apparent affinity constant, and n is the number of binding sites. The parameters were recovered from non-linear fits of equation (5) to experimental datasets using the Microcal Origin™ 6.0 software.

4.5.5. Analysis of the chaperone activity of peptides

To check the influence of the labels on Tat(44-61) or NC chaperone properties, the ability of the labeled peptides to promote the annealing of cTAR with its complementary dTAR sequence was compared to that of the unlabeled peptide (Godet, De Rocquigny et al. 2006; Kuciak, Gabus et al. 2008; Boudier, Storchak et al. 2010). The kinetic measurements were performed under pseudo-first-order conditions by using unlabeled dTAR at a concentration which was 30-fold higher than the concentration of cTAR labeled with carboxytetramethylrhodamine (TMR) at the 5′ end and with 5/6-carboxyfluorescein (Fl) at the 3′ end (Godet, De Rocquigny et al. 2006). Excitation and emission wavelengths were 480 and 520 nm for monitoring the Fl fluorescence, respectively. All reported concentrations correspond to those after mixing. To avoid high local concentrations during

mixing, both reactants were prepared at the same volume. Peptides were added to each reactant separately, and then the reaction was initiated by mixing the peptide-coated ODNs together. Experiments were performed in 25 mM Tris (pH 7.5) containing 30 mM NaCl and 0.2 mM $MgCl_2$ at 20ºC.

4.5.6. Parallax quenching method

The fluorescence intensity of labeled vesicles, either DOPC or DOPC with 15% nitroxide lipids, was measured in a 1 cm semi-micro quartz cuvette. Using the corrected F/F_0 values, the distance of the fluorophores from the center of the bilayer was calculated using the parallax equation (6) originally proposed by London and collaborators (Chattopadhyay and London 1987; Abrams and London 1993; Kaiser and London 1998):

$$Z_{cf} = L_{cl} + [-\ln(F_1/F_2)/\pi C - L_{21}^2\,]/2L_{21} \qquad (6)$$

where Z_{cf} is the distance of the fluorophore from the center of the bilayer; F_1 and F_2 are the fluorescence intensities in the presence of the shallow quencher (quencher 1) or the deeper quencher (quencher 2), respectively; L_{cl} is the distance of the shallow quencher from the center of the bilayer, L_{21} is the distance between the shallow and deep quenchers, and C the concentration of quencher in molecules/$Å^2$ (equals the mole fraction of nitroxide-labeled phospholipid divided by area per phospholipid; presently C = 0.15/70 $Å^2$ (Tieleman, Marrink et al. 1997)). The quenching by the two most efficient quenchers (TempoPC/5-SLPC or 5-SLPC/12-SLPC) is used to calculate Z_{cf} (Chattopadhyay and London 1987; Abrams and London 1993; Kaiser and London 1998). The values used for the distances of the nitroxide group from the bilayer center were 5.85 Å for 12-SLPC, 12.15 Å for 5-SLPC, and 19.5 Å for TempoPC (Chattopadhyay and London 1987; Abrams and London 1993; Kaiser and London 1998).

4.5.7. Deconvolution of fluorescence spectra

Deconvolution of the fluorescence spectra into three bands, corresponding to the normal (N*), H-bonded normal (H-N*) and tautomer (T*) forms, was performed using the Siano software kindly provided by Dr. A.O. Doroshenko (Kharkov, Ukraine), as previously described (Klymchenko, Duportail et al. 2004; Klymchenko, Mely et al. 2004). The program is based on an iterative nonlinear least-squares method, where the individual emission bands were approximated by a log-normal function accounting for several parameters: maximal amplitude, I_{max}, spectral maximum position, ν_{max}, and position of half-maximum amplitudes, ν_1 and ν_2, for the blue and red parts of the band, respectively. These parameters determine the shape parameters of the log-normal function, namely the full width at the half-maximum, FWHM = $\nu_1-\nu_2$, and the band asymmetry, P = $(\nu_1-\nu_{max})/(\nu_{max}-\nu_2)$. For the iteration process in case of MFL-labeled peptides, the FWHM of the two short-wavelength bands (N* and H-N*) were fixed at 3000 cm^{-1}. For the H-N* band, the asymmetry and the band position were fixed at 0.9 and 19000 cm^{-1}, respectively. The other parameters, i.e. asymmetry of N* and T* bands, the band width of the T* band and the relative intensities of the bands, were allowed to vary in the iteration process. The resulting fluorescence

intensities of the separated N*, H-N* and T* bands (I_{N*}, $I_{H\text{-}N*}$ and I_{T*}) were used for calculation of the hydration parameter, which was expressed as the ratio of the peak emission intensity of the hydrated (H-N*) form to the summed intensities of the non-hydrated (N* and T*) forms. Taking into account that the FWHM for the T* band is ca 2-fold narrower than for the N* and H-N* bands, the hydration was estimated as $I_{H\text{-}N*}/(I_{N*}+0.5\times I_{T*})$. The "polarity" parameter was expressed as the I_{N*}/I_{T*} ratio (Klymchenko, Duportail et al. 2004; Klymchenko, Mely et al. 2004).

4.5.8. Dynamic light scattering

Dynamic Light Scattering (DLS) is a technique used for measuring the particle size typically in the sub-micron domain. It is based on the principle of Brownian motion and relates this to the size of the particles. Brownian motion is the random movement of particles due to their collision with the solvent molecules that surround them. Normally DLS measures particles suspended within a liquid. The larger the particle, the slower the Brownian motion will be. Smaller particles are "kicked" further by the solvent molecules and move more rapidly.

The diameter measured in DLS refers to how a particle diffuses within a fluid so it is referred to as a hydrodynamic diameter. The diameter obtained by this technique is the diameter of a sphere that has the same translational diffusion coefficient as the particles. The size of the spherical particle is calculated from the translational diffusion coefficient by using the Stokes-Einstein equation (7):

$$d(H) = \frac{kT}{3\pi\eta D} \qquad (7)$$

where d (H) = hydrodynamic diameter; D = translational diffusion coefficient; k = Boltzmann's constant; T = absolute temperature; η = viscosity

In a dynamic light scattering instrument, a laser is used as the light source to illuminate the sample cells. For dilute concentrations, most of the laser beam passes through the sample, but some get scattered by the particles within the sample at all angles. In common device, a detector measures the scattered light at 90°. The scattering intensity signal from the detector is processed by a correlator and then analysed by a computer to derive the size information.

The average size of the vesicles and their complexes with peptides was determined with a Zetasizer Nano ZS (Malvern Instruments, Paris, France). All experiments were performed using 800 µl of the final volume after 5 min incubation at room temperature.

4.5.9. Atomic force microscopy (AFM)

Atomic force microscope (AFM) (Figure 4.1) provides a 3D profile of the surface on a nanoscale by measuring the forces between a sharp probe (<10 nm) and the surface at very short distance (0.2-10 nm probe-sample separation). The probe is supported on a flexible cantilever. The

AFM tip "gently" touches the surface and records the small force between the probe and the surface. Attractive or repulsive forces resulting from interactions between the tip and the surface show a positive or negative bending of the cantilever. The bending is detected by means of a laser beam, which is reflected from the back side of the cantilever. The advantages of AFM compared to, for example, electron microscopy are the ease of sample preparation and the possibility to image samples in native conditions in buffer.

AFM measurements were performed using a Solver Pro M (NT-MDT) instrument. The measurements were performed in liquid phase (Hepes buffer pH 7.4) by using the tapping mode. The cantilevers used were NSG03 type (NT-MDT) with a typical spring constant of 1.7 N/m, a resonance frequency of 32 kHz in liquid and a tip curvature radius of 10 nm. Images were acquired with a resolution of 512x512 points and a scan rate of 2 Hz.

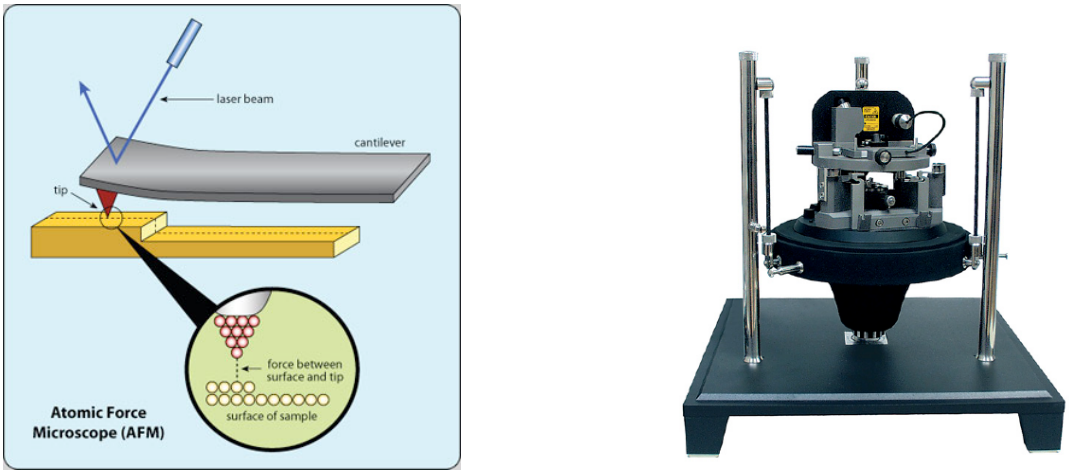


Figure 4.1. Schematic presentation of AFM and a photo of the Solver Pro M (NT-MDT) AFM instrument.

4.5.10. Fluorescence microscopy

Fluorescence microscopy experiments were performed by using a home-built two-photon laser scanning setup based on an Olympus IX70 inverted microscope with an Olympus 60x 1.2NA water immersion objective (Azoulay, Clamme et al. 2003; Clamme, Azoulay et al. 2003). Two-photon excitation was provided by a titanium-sapphire laser (Tsunami, Spectra Physics), and photons were detected with Avalanche Photodiodes (APD SPCM-AQR-14-FC, Perkin-Elmer) connected to a counter/timer PCI board (PCI6602, National Instrument). Imaging was carried out using two fast galvo-mirrors in the descanned fluorescence collection mode. Typical acquisition time was 5 s with an excitation power around 2.5 mW (830 nm) at the sample. Images corresponding to the blue and red channels were recorded simultaneously using a dichroic mirror (Beamsplitter 585 DCXR) and two APDs. The images were processed with a home-made program under LabView that generates a ratiometric image by dividing the image of the blue channel by that of the red channel. For each pixel, a pseudocolor scale is used for coding the ratio, while the

intensity is defined by the integrated intensity recorded for both channels at the corresponding pixel (Klymchenko, Oncul et al. 2009).

4.5.11. Total internal reflection fluorescence microscopy

Single particle tracking is based on the use of a wide field microscope with an evanescent wave as an excitation source which can selectively excite fluorescent molecules in the liquid near the interface with very low background fluorescence. By recording image sequences over time, it is possible to reconstruct the trajectory of fluorescently labeled molecules from which the diffusion constant can be obtained. In the case of a pure Brownian diffusion, the mean square displacement (MSD) is proportional to the elapsed time (t) and the diffusion constant (D):

$$MSD = 4Dt \qquad (8)$$

In the frame of the Stockes-Einstein model, the hydrodynamic radius of the particle can be obtained:

$$r_H = \frac{kT}{6\pi\eta D} \qquad (9)$$

where k is the Boltzmann constant, T the absolute temperature, η the viscosity and D the diffusion constant of the particle.

Single particle tracking experiments were performed on a widefield microscope based on an Olympus IX-71 setup coupled to several CW lasers (405, 488, 532 and 635 nm) and a highly sensitive Electron Multiplying Charge Coupled Device (EMCCD, Hamamatsu). Total Internal Reflection (TIRF) excitation geometry was added to this setup to generate an evanescent wave used to excite fluorophores within the first 150 nm of the sample. The rhodamine dye was excited at 532 nm and a bandpass filter (Semrock 585+/-10 nm) was used in front of the EMCCD. ImageJ and Igor softwares were used to calculate the diffusion coefficients of the particles in lipid bilayers.

CHAPTER 5

REFERENCES

5. REFERENCES

Abrams, F. S. and E. London (1993). "Extension of the parallax analysis of membrane penetration depth to the polar region of model membranes: use of fluorescence quenching by a spin-label attached to the phospholipid polar headgroup." *Biochemistry* **32**(40): 10826-10831.

Acuna, A. U., F. Toribio, F. Amat-Guerri and J. Catalan (1985). "Exited state proton transfer: a new feature in the fluorescence of methyl 5-chlorosalicylate and methyl 5-methoxysalicylate." *J. Photochem.* **30**: 339-352.

Adamson, C. S. and E. O. Freed (2007). "Human immunodeficiency virus type 1 assembly, release, and maturation." *Adv. Pharmacol.* **55**: 347-87.

Agard, N. J., J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi (2006). "A comparative study of bioorthogonal reactions with azides." *ACS Chem. Biol.* **1**(10): 644-648.

Aggarwal, A. K., D. W. Rodgers, M. Drottar, M. Ptashne and S. C. Harrison (1988). "Recognition of a DNA operator by the repressor of phage-434 - a view at high-resolution." *Science* **242**(4880): 899-907.

Aldovini, A. and R. A. Young (1990). "Mutations of RNA and protein sequences involved in Human-Immunodeficiency-Virus type-1 packaging result in production of non-infectious virus." *J. Virol.* **64**(5): 1920-1926.

Alessandrini, A. and P. Facci (2011). "Unraveling lipid/protein interaction in model lipid bilayers by Atomic Force Microscopy." *J. Mol. Recognit.* **24**(3): 387-396.

Allende, D., S. A. Simon and T. J. McIntosh (2005). "Melittin-induced bilayer leakage depends on lipid material properties: evidence for toroidal pores." *Biophys. J.* **88**(3): 1828-1837.

Almeida, P. F. and A. Pokorny (2009). "Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: from kinetics to thermodynamics." *Biochemistry* **48**(34): 8083-8093.

Alves, I. D., I. Correia, C. Y. Jiao, E. Sachon, S. Sagan, S. Lavielle, G. Tollin and G. Chassaing (2009). "The interaction of cell-penetrating peptides with lipid model systems and subsequent lipid reorganization: thermodynamic and structural characterization." *J. Pept. Sci.* **15**(3): 200-209.

Alves, I. D., N. Goasdoue, I. Correia, S. Aubry, C. Galanth, S. Sagan, S. Lavielle and G. Chassaing (2008). "Membrane interaction and perturbation mechanisms induced by two cationic cell penetrating peptides with distinct charge distribution." *Biochim. Biophys. Acta-Gen. Subj.* **1780**(7-8): 948-959.

Amarasinghe, G. K., R. N. De Guzman, R. B. Turner, K. J. Chancellor, Z. R. Wu and M. F. Summers (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**(2): 491-511.

Amarasinghe, G. K., J. Zhou, M. Miskimon, K. J. Chancellor, J. A. McDonald, A. G. Matthews, R. R. Miller, M. D. Rouse and M. F. Summers (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**(5): 961-970.

Andreatta, D., S. Sen, J. L. P. Lustres, S. A. Kovalenko, N. P. Ernsting, C. J. Murphy, R. S. Coleman and M. A. Berg (2006). "Ultrafast dynamics in DNA: "fraying" at the end of the helix." *J. Am. Chem. Soc.* **128**(21): 6885-6892.

Angell, Y. L. and K. Burgess (2007). "Peptidomimetics via copper-catalyzed azide-alkyne cycloadditions." *Chem. Soc. Rev.* **36**(10): 1674-1689.

Angelova, M. I. and D. S. Dimitrov (1986). "Liposome electroformation." *Faraday Discuss. Chem. S.* **81**: 303-311.

Ao, Z. J., G. Y. Huang, H. Yao, Z. K. Xu, M. Labine, A. W. Cochrane and X. J. Yao (2007). "Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication." *J. Biol. Chem.* **282**(18): 13456-13467.

Arhel, N., S. Munier, P. Souque, K. Mollier and P. Charneau (2006). "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.* **80**(20): 10262-10269.

Arhel, N. J., S. Souquere-Besse, S. Munier, P. Souque, S. Guadagnini, S. Rutherford, M. C. Prevost, T. D. Allen and P. Chameau (2007). "HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore." *Embo J.* **26**(12): 3025-3037.

Arya, S. K., C. Guo, S. F. Josephs and F. Wongstaal (1985). "Trans-activator gene of human T-lymphotropic virus type-III (HTLV-III)." *Science* **229**(4708): 69-73.

Ashe, M. P., L. H. Pearson and N. J. Proudfoot (1997). "The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site." *Embo J.* **16**(18): 5752-5763.

Ashkin, A. (1997). "Optical trapping and manipulation of neutral particles using lasers." *Proc. Natl. Acad. Sci. U. S. A.* **94**(10): 4853-4860.

Avery, O. T., C. M. Macleod and M. McCarty (1944). "Studies on the chemical nature of the substance inducing transformation of pneumococcal types - induction of transformation by a deoxyribonucleic-acid fraction isolated from pneumococcus type-III." *J. Exp. Med.* **79**(4): 137-158.

Azoulay, J., J. P. Clamme, J. L. Darlix, B. P. Roques and Y. Mely (2003). "Destabilization of the HIV-1 complementary sequence of TAR by the nucleocapsid protein through activation of conformational fluctuations." *J. Mol. Biol.* **326**(3): 691-700.

Bachand, F., X. J. Yao, M. Hrimech, N. Rougeau and E. A. Cohen (1999). "Incorporation of Vpr into human immunodeficiency virus type 1 requires a direct interaction with the p6 domain of the p55 gag precursor." *J. Biol. Chem.* **274**(13): 9083-9091.

Bagatolli, L. A. (2006). "To see or not to see: Lateral organization of biological membranes and fluorescence microscopy." *Biochim. Biophys. Acta* **1758**(10): 1541-1556.

Bak, M., R. P. Bywater, M. Hohwy, J. K. Thomsen, K. Adelhorst, H. J. Jakobsen, O. W. Sorensen and N. C. Nielsen (2001). "Conformation of alamethicin in oriented phospholipid bilayers determined by N-15 solid-state nuclear magnetic resonance." *Biophys. J.* **81**(3): 1684-1698.

Barillari, G., R. Gendelman, R. C. Gallo and B. Ensoli (1993). "The Tat protein of human immunodeficiency virus type 1, a growth factor for AIDS Kaposi sarcoma and cytokine-activated vascular cells, induces adhesion of the same cell types by using integrin receptors recognizing the RGD amino acid sequence." *Proc. Natl. Acad. Sci. U. S. A.* **90**(17): 7941-7945.

Barone, A. D., C. Chen, G. H. McGall, K. Rafii, P. R. Buzby and J. J. Dimeo (2001). "Novel nucleoside triphosphate analogs for the enzymatic labeling of nucleic acids." *Nucleosides Nucleotides & Nucleic Acids* **20**(4-7): 1141-1145.

Barresinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axlerblin, F. Vezinetbrun, C. Rouzioux, W. Rozenbaum and L. Montagnier (1983). "Isolation of a T-lymphotropic retrovirus from a patient at risk for Acquired Immune-Deficiency Syndrome (AIDS)." *Science* **220**(4599): 868-871.

Bartels, F. W., B. Baumgarth, D. Anselmetti, R. Ros and A. Becker (2003). "Specific binding of the regulatory protein ExpG to promoter regions of the galactoglucan biosynthesis gene cluster of Sinorhizobium meliloti - a combined molecular biology and force spectroscopy investigation." *J. Struct. Biol.* **143**(2): 145-152.

Baskin, J. M., J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi (2007). "Copper-free click chemistry for dynamic in vivo imaging." *Proc. Natl. Acad. Sci. U. S. A.* **104**(43): 16793-16797.

Basu, V. P., M. Song, L. Gao, S. T. Rigby, M. N. Hanson and R. A. Bambara (2008). "Strand transfer events during HIV-1 reverse transcription." *Virus Res.* **134**(1-2): 19-38.

Baudier, J., N. Glasser and G. Duportail (1986). "Bimane- and acrylodan-labeled S100 proteins. Role of cysteines-85 alpha and -84 beta in the conformation and calcium binding properties of S100 alpha alpha and S100b (beta beta) proteins." *Biochemistry* **25**(22): 6934-6941.

Bayer, P., M. Kraft, A. Ejchart, M. Westendorp, R. Frank and P. Rosch (1995). "Structural studies of HIV-1 Tat protein." *J. Mol. Biol.* **247**(4): 529-535.

Bazzi, A., L. Zargarian, F. Chaminade, C. Boudier, H. de Rocquigny, B. Rene, Y. Mely, P. Fosse and O. Mauffret (2011). "Structural insights into the cTAR DNA recognition by the HIV-1 nucleocapsid protein: role of sugar deoxyriboses in the binding polarity of NC." *Nucleic Acids Res.* **39**(9): 3903-3916.

Beaucage, S. L. and R. P. Iyer (1992). "Advances in the synthesis of oligonucleotides by the phosphoramidite approach." *Tetrahedron* **48**(12): 2223-2311.

Bechinger, B. (1997). "Structure and functions of channel-forming peptides: magainins, cecropins, melittin and alamethicin." *J. Membr. Biol.* **156**(3): 197-211.

Bechinger, B. (1999). "The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy." *Biochim. Biophys. Acta-Biomembr.* **1462**(1-2): 157-183.

Bechinger, B. (2005). "Detergent-like properties of magainin antibiotic peptides: a P-31 solid-state NMR spectroscopy study." *Biochim. Biophys. Acta-Biomembr.* **1712**(1): 101-108.

Bechinger, B., R. Kinder, M. Helmle, T. C. B. Vogt, U. Harzer and S. Schinzel (1999). "Peptide structural analysis by solid-state NMR spectroscopy." *Biopolymers* **51**(3): 174-190.

Bechinger, B., D. A. Skladnev, A. Ogrel, X. Li, E. V. Rogozhkina, T. V. Ovchinnikova, J. D. J. O'Neil and J. Raap (2001). "N-15 and P-31 solid-state NMR investigations on the orientation of zervamicin II and alamethicin in phosphatidylcholine membranes." *Biochemistry* **40**(31): 9428-9437.

Bechinger, B., M. Zasloff and S. J. Opella (1992). "Structure and interactions of magainin antibiotic peptides in lipid bilayers - a solid-state nuclear magnetic resonance investigation." *Biophys. J.* **62**(1): 12-14.

Becker, C. F. W., C. L. Hunter, R. Seidel, S. B. H. Kent, R. S. Goody and M. Engelhard (2003). "Total chemical synthesis of a functional interacting protein pair: the protooncogene H-Ras and the Ras-binding domain of its effector c-Raf1." *Proc. Natl. Acad. Sci. U. S. A.* **100**(9): 5075-5080.

Belliard, G., A. Romieu, J. F. Zagury, H. Dali, O. Chaloin, R. Le Grand, E. Loret, J. P. Briand, B. Roques, C. Desgranges and S. Muller (2003). "Specificity and effect on apoptosis of Tat antibodies from vaccinated and SHIV-infected rhesus macaques and HIV-infected individuals." *Vaccine* **21**(23): 3186-3199.

Beltz, H., J. Azoulay, S. Bernacchi, J. P. Clamme, D. Ficheux, B. Roques, J. L. Darlix and Y. Mely (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**(1): 95-108.

Beltz, H., C. Clauss, E. Piemont, D. Ficheux, R. J. Gorelick, B. Roques, C. Gabus, J. L. Darlix, H. de Rocquigny and Y. Mely (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**(5): 1113-1126.

Beltz, H., E. Piemont, E. Schaub, D. Ficheux, B. Roques, J. L. Darlix and Y. Mely (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**(4): 711-723.

Benachir, T., M. Monette, J. Grenier and M. Lafleur (1997). "Melittin-induced leakage from phosphatidylcholine vesicles is modulated by cholesterol: a property used for membrane targeting." *Eur. Biophys. J. Biophys. Lett.* **25**(3): 201-210.

Benner, S. A. (2004). "Understanding nucleic acids using synthetic chemistry." *Acc. Chem. Res.* **37**(10): 784-797.

Berg, J. M. (1986). "Potential metal-binding domains in nucleic acid binding-proteins." *Science* **232**(4749): 485-487.

Berg, O. G., R. B. Winter and P. H. von Hippel (1981). "Diffusion-driven mechanisms of protein translocation on nucleic acids. Models and theory." *Biochemistry* **20**(24): 6929-6948.

Berger, R., M. R. Duncan and B. Berman (1993). "Nonradioactive gel mobility shift assay using chemiluminescent detection." *Biotechniques* **15**(4): 650-654.

Berkhout, B. (1996). Structure and function of the human immunodeficiency virus leader RNA. Progress In Nucleic Acid Research And Molecular Biology. San Diego, Academic Press Inc. **54:** 1-34.

Bernacchi, S., S. Henriet, P. Dumas, J. C. Paillart and R. Marquet (2007). "RNA and DNA binding properties of HIV-1 Vif protein: a fluorescence study." *J. Biol. Chem.* **282**(36): 26361-26368.

Bernacchi, S., S. Stoylov, E. Piemont, D. Ficheux, B. P. Roques, J. L. Darlix and Y. Mely (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**(3): 385-399.

Berneche, S., M. Nina and B. Roux (1998). "Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane." *Biophys. J.* **75**(4): 1603-1618.

Berthet-Colominas, C., S. Monaco, A. Novelli, G. Sibai, F. Mallet and S. Cusack (1999). "Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab." *Embo J.* **18**(5): 1124-36.

Berthoux, L., C. Pechoux, M. Ottmann, G. Morel and J. L. Darlix (1997). "Mutations in the N-terminal domain of human immunodeficiency virus type 1 nucleocapsid protein affect virion core structure and proviral DNA synthesis." *J. Virol.* **71**(9): 6973-81.

Billeter, M., Y. Qian, G. Otting, M. Muller, W. J. Gehring and K. Wuthrich (1990). "Determination of the 3-dimensional structure of the antennapedia homeodomain from Drosophila in solution by H-1 nuclear magnetic resonance spectroscopy." *J. Mol. Biol.* **214**(1): 183-197.

Binder, H. and G. Lindblom (2003). "Charge-dependent translocation of the Trojan peptide penetratin across lipid membranes." *Biophys. J.* **85**(2): 982-995.

Binnig, G., C. F. Quate and C. Gerber (1986). "Atomic Force Microscope." *Phys. Rev. Lett.* **56**(9): 930-933.

Blackburn, G. M., M. J. Gait, D. Loakes and D. M. Williams (2006). Nucleic acids in chemistry and biology. Cambridge, Royal Society of Chemistry.

Blainey, P. C., G. Luo, S. C. Kou, W. F. Mangel, G. L. Verdine, B. Bagchi and X. S. Xie (2009). "Nonspecifically bound proteins spin while diffusing along DNA." *Nat. Struct. Mol. Biol.* **16**(12): 1224-1234.

Blainey, P. C., A. M. van Oijent, A. Banerjee, G. L. Verdine and X. S. Xie (2006). "A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA." *Proc. Natl. Acad. Sci. U. S. A.* **103**(15): 5752-5757.

Bloomfield, V. A., D. M. Crothers and J. Tinoco (2000). Nucleic acids: structures, properties, and functions. New York, University Science Books.

Bombarda, E., E. Grell, B. P. Roques and Y. Mely (2007). "Molecular mechanism of the Zn2+-induced folding of the distal CCHC finger motif of the HIV-1 nucleocapsid protein." *Biophys. J.* **93**(1): 208-217.

Bonnet, I., A. Biebricher, P. Porte, C. Loverdo, O. Benichou, R. Voituriez, C. Escude, W. Wende, A. Pingoud and P. Desbiolles (2008). "Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA." *Nucleic Acids Res.* **36**(12): 4118-4127.

Boudier, C., R. Storchak, K. K. Sharma, P. Didier, A. Follenius-Wund, S. Muller, J. L. Darlix and Y. Mely (2010). "The mechanism of HIV-1 Tat-directed nucleic acid annealing supports its role in reverse transcription." *J. Mol. Biol.* **400**(3): 487-501.

Bourbigot, S., H. Beltz, J. Denis, N. Morellet, B. P. Roques, Y. Mely and S. Bouaziz (2005). "The C-terminal domain of the HIV-1 regulatory protein Vpr adopts an antiparallel dimeric structure in solution via its leucine-zipper-like domain." *Biochem. J.* **387**(Pt 2): 333-341.

Bourbigot, S., N. Ramalanjaona, C. Boudier, G. F. Salgado, B. P. Roques, Y. Mely, S. Bouaziz and N. Morellet (2008). "How the HIV-1 nucleocapsid protein binds and destabilises the (-)Primer Binding Site during reverse transcription." *J. Mol. Biol.* **383**(5): 1112-1128.

Brandt, F., L. A. Carlson, F. U. Hartl, W. Baumeister and K. Grunewald (2010). "The Three-Dimensional Organization of Polyribosomes in Intact Human Cells." *Mol. Cell* **39**(4): 560-569.

Brauns, E. B., C. J. Murphy and M. A. Berg (1998). "Local dynamics in DNA by temperature-dependent Stokes shifts of an intercalated dye." *J. Am. Chem. Soc.* **120**(10): 2449-2456.

Brenowitz, M., D. F. Senear, M. A. Shea and G. K. Ackers (1986). "Quantitative DNase footprint titration - a method for studying protein-DNA interactions." *Meth.Enzymol.* **130**: 132-181.

Breslauer, K. J. (1995). "Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions." *Methods Enzymol.* **259**: 221-242.

Brogden, K. A. (2005). "Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?" *Nat. Rev. Microbiol.* **3**(3): 238-250.

Brown, R. S., C. Sander and P. Argos (1985). "The Primary Structure Of Transcription Factor TFIIIA Has 12 Consecutive Repeats." *FEBS Lett.* **186**(2): 271-274.

Bukrinskaya, A. (2007). "HIV-1 matrix protein: A mysterious regulator of the viral life cycle." *Virus Res.* **124**(1-2): 1-11.

Bukrinsky, M. and A. Adzhubei (1999). "Viral protein R of HIV-1." *Rev. Med. Virol.* **9**(1): 39-49.

Bukrinsky, M. I., S. Haggerty, M. P. Dempsey, N. Sharova, A. Adzhubei, L. Spitz, P. Lewis, D. Goldfarb, M. Emerman and M. Stevenson (1993). "A nuclear-localization signal within HIV-1 matrix protein that governs infection of nondividing cells." *Nature* **365**(6447): 666-669.

Byeon, I. J. L., X. Meng, J. W. Jung, G. P. Zhao, R. F. Yang, J. W. Ahn, J. Shi, J. Concel, C. Aiken, P. J. Zhang and A. M. Gronenborn (2009). "Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function." *Cell* **139**(4): 780-790.

Byrne, R. D., M. Gamier-Lhomme, K. Han, M. Dowicki, N. Michael, N. Totty, V. Zhendre, A. Cho, T. R. Pettitt, M. J. Wakelam, D. L. Poccia and B. Larijani (2007). "PLC gamma is enriched on poly-phosphoinositide-rich vesicles to control nuclear envelope assembly." *Cell. Signal.* **19**(5): 913-922.

Campbell, G. R. and E. P. Loret (2009). "What does the structure-function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine?" *Retrovirology* **6**: 50.

Campbell, G. R., E. Pesquier, J. Watkins, V. Bourgarel-Rey, V. Peyrot, D. Esquieu, P. Barbier, J. de Mareuil, D. Braguer, P. Kaleebu, D. L. Yirrell and E. P. Loret (2004). "The glutamine-rich region of the HIV-1 Tat protein is involved in T-cell apoptosis." *J. Biol. Chem.* **279**(46): 48197-48204.

Carteau, S., S. C. Batson, L. Poljak, J. F. Mouscadet, H. DeRocquigny, J. L. Darlix, B. P. Roques, E. Kas and C. Auclair (1997). "Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg2+-dependent DNA integration in vitro." *J. Virol.* **71**(8): 6225-6229.

Carteau, S., R. J. Gorelick and F. D. Bushman (1999). "Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase in vitro: stimulation by the viral nucleocapsid protein." *J. Virol.* **73**(8): 6670-6679.

Caruthers, M. H., A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. Matteucci, Z. Stabinsky and J. Y. Tang (1987). "Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method." *Methods Enzymol.* **154**: 287-313.

Cary, P. D. and G. G. Kneale (2008). Circular dichroism for the analysis of protein-DNA interactions. DNA-Protein Interactions: Principles and Protocols, Third Edition. **543:** 613-624.

Catalan, J., P. Perez, J. Laynez and F. G. Blanco (1991). "Analysis of the solvent effect on the photophysics properties of 6-propionyl-2-(dimethylamino)naphthalene (PRODAN)." *J. Fluoresc.* **1**(4): 215-223.

Cech, T. R. (1987). "The chemistry of self-splicing RNA and RNA enzymes." *Science* **236**(4808): 1532-1539.

Cech, T. R. (1989). "RNA chemistry - ribozyme self-replication." *Nature* **339**(6225): 507-508.

Cech, T. R. (2000). "Structural biology - the ribosome is a ribozyme." *Science* **289**(5481): 878-879.

Cen, S., Y. E. Huang, A. Khorchid, J. L. Darlix, M. A. Wainberg and L. Kleiman (1999). "The role of Pr55(gag) in the annealing of tRNA44(3)(Lys) to human immunodeficiency virus type 1 genomic RNA." *J. Virol.* **73**(5): 4485-4488.

Cerezo, F. M., S. C. Rocafort, P. S. Sierra, F. Garcia-Blanco, C. D. Oliva and J. C. Sierra (2001). "Photophysical study of the probes acrylodan (1[6-(dimethylamino)naphthalene-2-yl]prop-2-en-1-one), ANS (8-anilino-naphthalene-1-sulfonate) and prodan (1-[6-(dymethylamino)naphthalene-2-yl]propan-1-1) in aqueous mixtures of various alcohols." *Helv. Chim. Acta* **84**: 3306-3312.

Chakrabarti, A., D. A. Kelkar and A. Chattopadhyay (2006). "Spectrin organization and dynamics: new insights." *Biosci. Rep.* **26**(6): 369-386.

Chang, H. C., F. Samaniego, B. C. Nair, L. Buonaguro and B. Ensoli (1997). "HIV-1 Tat protein exits from cells via a leaderless secretory pathway and binds to extracellular matrix-associated heparan sulfate proteoglycans through its basic region." *Aids* **11**(12): 1421-1431.

Chattopadhyay, A. (1990). "Chemistry and biology of N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-labeled lipids: Fluorescent probes of biological and model membranes." *Chem. Phys. Lipids* **53**(1): 1-15.

Chattopadhyay, A. and E. London (1987). "Parallax method for direct measurement of membrane penetration depth utilizing fluorescence quenching by spin-labeled phospholipids." *Biochemistry* **26**(1): 39-45.

Chattopadhyay, A. and H. Raghuraman (2004). "Application of fluorescence spectroscopy to membrane protein structure and dynamics." *Curr. Sci.* **87**(2): 175-180.

Chauhan, A., J. Turchan, C. Pocernich, A. Bruce-Keller, S. Roth, D. A. Butterfield, E. O. Major and A. Nath (2003). "Intracellular human immunodeficiency virus tat expression in astrocytes promotes astrocyte survival but induces potent neurotoxicity at distant sites via axonal transport." *J. Biol. Chem.* **278**(15): 13512-13519.

Chen, F. Y., M. T. Lee and H. W. Huang (2003). "Evidence for membrane thinning effect as the mechanism for peptide-induced pore formation." *Biophys. J.* **84**(6): 3751-3758.

Chen, H., N. N. Chung, C. Lemieux, B. Zelent, J. M. Vanderkooi, I. Gryczynski, B. C. Wilkes and P. W. Schiller (2005). "[Aladan3]TIPP: a fluorescent delta-opioid antagonist with high delta-receptor binding affinity and delta selectivity." *Biopolymers* **80**(2-3): 325-331.

Chen, J. and W. R. Lariviere (2010). "The nociceptive and anti-nociceptive effects of bee venom injection and therapy: A double-edged sword." *Prog. Neurobiol.* **92**(2): 151-183.

Chen, S. F., A. Gunasekera, X. P. Zhang, T. A. Kunkel, R. H. Ebright and H. M. Berman (2001). "Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: alteration of DNA binding specificity through alteration of DNA kinking." *J. Mol. Biol.* **314**(1): 75-82.

Chou, I.-T., M. L. Martinez and H. Clements (1993). "Reversal of excitation behavior of proton-transfer vs. charge-transfer by dielectric pertirbation of electronic manifolds." *J. Phys. Chem.* **97**: 2618-2622.

Chou, P. T., S. C. Pu, Y. M. Cheng, W. S. Yu, Y. C. Yu, F. T. Hung and W. P. Hu (2005). "Femtosecond dynamics on excited-state proton/charge-transfer reaction in 4'-N,N-diethylamino-3-hydroxyflavone. The role of dipolar vectors in constructing a rational mechanism." *J. Phys. Chem. A* **109**(17): 3777-3787.

Chow, C. S., S. K. Mahto and T. N. Lamichhane (2008). "Combined approaches to site-specific modification of RNA." *ACS Chem. Biol.* **3**(1): 30-37.

Cimarelli, A., S. Sandin, S. Hoglund and J. Luban (2000). "Basic residues in human immunodeficiency virus type 1 nucleocapsid promote virion assembly via interaction with RNA." *J. Virol.* **74**(7): 3046-3057.

Clamme, J. P., J. Azoulay and Y. Mely (2003). "Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy." *Biophys. J.* **84**(3): 1960-1968.

Clavel, F., M. Guyader, D. Guetard, M. Salle, L. Montagnier and M. Alizon (1986). "Molecular-cloning and polymorphism of the Human Immune-Deficiency Virus type-2." *Nature* **324**(6098): 691-695.

Clever, J., C. Sassetti and T. G. Parslow (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**(4): 2101-2109.

Clever, J. L., D. A. Eckstein and T. G. Parslow (1999). "Genetic dissociation of the encapsidation and reverse transcription functions in the 5' R region of human immunodeficiency virus type 1." *J. Virol.* **73**(1): 101-9.

Clore, G. M. and J. Iwahara (2009). "Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes." *Chem. Rev.* **109**(9): 4108-4139.

Coeytaux, E., D. Coulaud, E. Le Cam, O. Danos and A. Kichler (2003). "The cationic amphipathic alpha-helix of HIV-1 viral protein R (Vpr) binds to nucleic acids, permeabilizes membranes, and efficiently transfects cells." *J. Biol. Chem.* **278**(20): 18110-18116.

Cohen, B. E., T. B. McAnaney, E. S. Park, Y. N. Jan, S. G. Boxer and L. Y. Jan (2002). "Probing protein electrostatics with a synthetic fluorescent amino acid." *Science* **296**(5573): 1700-1703.

Cohen, B. E., A. Pralle, X. Yao, G. Swaminath, C. S. Gandhi, Y. N. Jan, B. K. Kobilka, E. Y. Isacoff and L. Y. Jan (2005). "A fluorescent probe designed for studying protein conformational change." *Proc. Natl. Acad. Sci. U. S. A.* **102**(4): 965-970.

Coleman, R. S., M. A. Berg and C. J. Murphy (2007). "Coumarin base-pair replacement as a fluorescent probe of ultrafast DNA dynamics." *Tetrahedron* **63**(17): 3450-3456.

Coleman, R. S. and M. L. Madaras (1998). "Synthesis of a novel coumarin C-riboside as a photophysical probe of oligonucleotide dynamics." *J. Org. Chem.* **63**(16): 5700-5703.

Conner, S. D. and S. L. Schmid (2003). "Regulated portals of entry into the cell." *Nature* **422**(6927): 37-44.

Corbin, A., B. Grigorov, P. Roingeard, J. L. Darlix and D. Muriaux (2008). "Revisiting HIV-1 assembly." *M S-Med. Sci.* **24**(1): 49-55.

Coren, L. V., J. A. Thomas, E. Chertova, R. C. Sowder, T. D. Gagliardi, R. J. Gorelick and D. E. Ott (2007). "Mutational analysis of the C-terminal Gag cleavage sites in human immunodeficiency virus type 1." *J. Virol.* **81**(18): 10047-10054.

Cosson, P. (1996). "Direct interaction between the envelope and matrix proteins of HIV-1." *Embo J.* **15**(21): 5783-5788.

Cox, W. G. and V. L. Singer (2004). "Fluorescent DNA hybridization probe preparation using amine modification and reactive dye coupling." *BioTechniques* **36**(1): 114-122.

D'Souza, V. and M. F. Summers (2005). "How retroviruses select their genomes." *Nat. Rev. Microbiol.* **3**(8): 643-655.

Da Costa, C. P., M. J. Fedor and L. G. Scott (2007). "8-Azaguanine reporter of purine ionization states in structured RNAs." *J. Am. Chem. Soc.* **129**(11): 3426-3432.

Dabrowska, A., N. Kim and A. Aldovini (2008). "Tat-induced FOXO3a is a key mediator of apoptosis in HIV-1 infected human CD4(+) T-lymphocytes." *J. Immunol.* **181**(12): 8460-8477.

Daniels, M. and W. Hauswirt (1971). "Fluorescence of purine and pyrimidine bases of nucleic acids in neutral aqueous Solution at 300 degrees K." *Science* **171**(3972): 675-677.

Darlix, J. L., J. Godet, R. Ivanyi-Nagy, P. Fosse, O. Mauffret and Y. Mely (2011). "Flexible nature and specific functions of the HIV-1 nucleocapsid protein." *J. Mol. Biol.* **410**(4): 565-581.

Darlix, J. L., M. Lapadattapolsky, H. Derocquigny and B. P. Roques (1995). "First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses." *J. Mol. Biol.* **254**(4): 523-537.

Dave, P. C., E. Billington, Y. L. Pan and S. K. Straus (2005). "Interaction of alamethicin with ether-linked phospholipid bilayers: Oriented circular dichroism, P-31 solid-state NMR, and differential scanning calorimetry studies." *Biophys. J.* **89**(4): 2434-2442.

Dawson, P. E. and S. B. H. Kent (2000). "Synthesis of native proteins by chemical ligation." *Annu. Rev. Biochem.* **69**: 923-960.

Dawson, P. E., T. W. Muir, I. Clarklewis and S. B. H. Kent (1994). "Synthesis of proteins by native chemical ligation." *Science* **266**(5186): 776-779.

de Graaf, A. J., M. Kooijman, W. E. Hennink and E. Mastrobattista (2009). "Nonnatural amino acids for site-specific protein conjugation." *Bioconjugate Chem.* **20**(7): 1281-1295.

de Guzman, R. N., Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer and M. F. Summers (1998). "Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element." *Science* **279**(5349): 384-388.

de Mareuil, J., M. Carre, P. Barbier, G. R. Campbell, S. Lancelot, S. Opi, D. Esquieu, J. D. Watkins, C. Prevot, D. Braguer, V. Peyrot and E. P. Loret (2005). "HIV-1 tat protein enhances microtubule polymerization." *Retrovirology* **2**: 5.

de Noronha, C. M., M. P. Sherman, H. W. Lin, M. V. Cavrois, R. D. Moir, R. D. Goldman and W. C. Greene (2001). "Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr." *Science* **294**(5544): 1105-1108.

de Rijck, J. and Z. Debyser (2006). "The central DNA flap of the human immunodeficiency virus type 1 is important for viral replication." *Biochem. Biophys. Res. Commun.* **349**(3): 1100-1110.

de Rocquigny, H., D. Ficheux, C. Gabus, M. C. Fournie-Zaluski, J. L. Darlix and B. P. Roques (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**(2): 1010-1018.

de Rocquigny, H., D. Ficheux, C. Gabus, M. C. Fourniezaluski, J. L. Darlix and B. P. Roques (1991). "1st large-scale chemical synthesis of the 72 amino-acid HIV-1 nucleocapsid protein NCp7 in an active form." *Biochemical And Biophysical Research Communications* **180**(2): 1010-1018.

Delelis, O., K. Carayon, A. Saib, E. Deprez and J. F. Mouscadet (2008). "Integrase and integration: biochemical activities of HIV-1 integrase." *Retrovirology* **5**: 114.

Demchenko, A. P. (1986). Ultraviolet Spectroscopy of Proteins. Berlin, Springer-Verlag.

Demchenko, A. P., Y. Mely, G. Duportail and A. S. Klymchenko (2009). "Monitoring biophysical properties of lipid membranes by environment-sensitive fluorescent probes." *Biophys. J.* **96**(9): 3461-3470.

Demene, H., C. Z. Dong, M. Ottmann, M. C. Rouyez, N. Jullian, N. Morellet, Y. Mely, J. L. Darlix, M. C. Fourniezaluski, S. Saragosti and B. P. Roques (1994). "H-1-NMR structure and biological studies of the His(23)->Cys mutant nucleocapsid protein of HIV-1 Indicate that the conformation of the first zinc-finger is critical for virus infectivity." *Biochemistry* **33**(39): 11707-11716.

Dempsey, C. E. (1990). "The actions of melittin on membranes." *Biochim. Biophys. Acta - Rev. Biomembr.* **1031**(2): 143-161.

Denek, M., A. Pelchen-Matthews, R. Byland, E. Ruiz-Mateos and M. Marsh (2007). "In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspanins CD81, CD9, and CD53." *J. Cell Biol.* **177**(2): 329-341.

Derossi, D., G. Chassaing and A. Prochiantz (1998). "Trojan peptides: the penetratin system for intracellular delivery." *Trends Cell Biol.* **8**(2): 84-87.

Derossi, D., A. H. Joliot, G. Chassaing and A. Prochiantz (1994). "The third helix of the antennapedia homeodomain translocates through biological membranes." *J. Biol. Chem.* **269**(14): 10444-10450.

Diaz, G., M. Melis, B. Batetta, F. Angius and A. Falchi (2008). "Hydrophobic characterization of intracellular lipids in situ by Nile Red red/yellow emission ratio." *MICRON* **39**(7): 819-824.

Dietz, G. P. H., P. C. Valbuena, B. Dietz, K. Meuer, P. Muller, J. H. Weishaupt and M. Bahr (2006). "Application of a blood-brain-barrier-penetrating form of GDNF in a mouse model for Parkinson's disease." *Brain Res.* **1082**: 61-66.

Dominguez, C., M. Schubert, O. Duss, S. Ravindranathan and F. H. T. Allain (2011). "Structure determination and dynamics of protein-RNA complexes by NMR spectroscopy." *Prog. Nucl. Mag. Res. Spec.* **58**(1-2): 1-61.

Dorfman, T., J. Luban, S. P. Goff, W. A. Haseltine and H. G. Gottlinger (1993). "Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type-1 nucleocapsid protein." *J. Virol.* **67**(10): 6159-6169.

Doucleff, M. and G. M. Clore (2008). "Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1." *Proc. Natl. Acad. Sci. U. S. A.* **105**(37): 13871-13876.

Dreyfuss, G., V. N. Kim and N. Kataoka (2002). "Messenger-RNA-binding proteins and the messages they carry." *Nat. Rev. Mol. Cell Biol.* **3**(3): 195-205.

Drummond, J. E., P. Mounts, R. J. Gorelick, J. R. CasasFinet, W. J. Bosche, L. E. Henderson, D. J. Waters and L. O. Arthur (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**(7): 533-543.

Duchardt, F., M. Fotin-Mleczek, H. Schwarz, R. Fischer and R. Brock (2007). "A comprehensive model for the cellular uptake of cationic cell-penetrating peptides." *Traffic* **8**(7): 848-866.

Dufau, I. and H. Mazarguil (2000). "Design of a fluorescent amino acid derivative usable in peptide synthesis." *Tetrahedron Lett.* **41**(32): 6063-6066.

Dvorin, J. D., P. Bell, G. G. Maul, M. Yamashita, M. Emerman and M. H. Malim (2002). "Reassessment of the roles of integrase and the central DNA flap in human immunodeficiency virus type 1 nuclear import." *J. Virol.* **76**(23): 12087-12096.

Eastman, J. W. (1967). "Quantitative spectrofluorimetry the fluorescence quantum yield of quinine sulfate." *Photochem. Photobiol.* **6**(1): 55-72.

Efcavitch, J. W. and J. F. Thompson (2010). Single-molecule DNA analysis. Ann. Rev. Anal. Chem. **3:** 109-128.

Egele, C., E. Schaub, N. Ramalanjaona, E. Piemont, D. Ficheux, B. Roques, J. L. Darlix and Y. Mely (2004). "HIV-1 nucleocapsid protein binds to the viral DNA initiation sequences and chaperones their kissing interactions." *J. Mol. Biol.* **342**(2): 453-466.

Eggeling, C., L. Brand and C. Seidel (1997). "Laser-induced fluorescence of coumarin derivatives in aqueous solution: photochemical aspects for single molecule detection." *Bioimaging* **5**(3): 105-115.

Ehrenstein, G. and H. Lecar (1977). "Electrically gated ionic channels in lipid bilayers." *Q. Rev. Biophys.* **10**(1): 1-34.

El Kirat, K., S. Morandat and Y. F. Dufrene (2010). "Nanoscale analysis of supported lipid bilayers using atomic force microscopy." *Biochim. Biophys. Acta-Biomembr.* **1798**(4): 750-765.

Ellenberger, T. E., C. J. Brandl, K. Struhl and S. C. Harrison (1992). "The GCN4 basic region leucine zipper binds DNZ as a dimer of uninterrupted alpha-helices - crystal-structure of the protein-DNA complex." *Cell* **71**(7): 1223-1237.

Ellman, J., D. Mendel, S. Anthonycahill, C. J. Noren and P. G. Schultz (1991). "Biosynthetic method for introducing unnatural amino acids site-specifically into proteins." *Method Enzymol.* **202**: 301-336.

Engel, A. and D. J. Muller (2000). "Observing single biomolecules at work with the atomic force microscope." *Nat. Struct. Biol.* **7**(9): 715-718.

Engelman, A., K. Mizuuchi and R. Craigie (1991). "HIV-1 DNA integration - mechanism of viral-DNA cleavage and DNA strand transfer." *Cell* **67**(6): 1211-1221.

Ensoli, B., L. Buonaguro, G. Barillari, V. Fiorelli, R. Gendelman, R. A. Morgan, P. Wingfield and R. C. Gallo (1993). "Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation." *J. Virol.* **67**(1): 277-287.

Ercelen, S., A. Klymchenko and A. Demchenko (2002). "An ultrasensitive fluorescent probe for hydrophobic range of solvent polarities." *Anal. Chim. Acta* **464**: 273-287.

Ercelen, S., A. D. Roshal, A. P. Demchenko and A. S. Klymchenko (2002). "Excited-state proton transfer reaction in a new benzofuryl 3-hydroxychromone derivative: the influence of low-polar solvents." *Polish J. Chem.* **76**(9): 1287-1299.

Fassati, A. and S. P. Goff (2001). "Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1." *J. Virol.* **75**(8): 3626-3635.

Fawell, S., J. Seery, Y. Daikh, C. Moore, L. L. Chen, B. Pepinsky and J. Barsoum (1994). "Tat-mediated delivery of heterologous proteins into cells." *Proc. Natl. Acad. Sci. U. S. A.* **91**(2): 664-668.

Feng, Y. X., S. Campbell, D. Harvin, B. Ehresmann, C. Ehresmann and A. Rein (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**(5): 4251-4256.

Feng, Y. X., T. D. Copeland, L. E. Henderson, R. J. Gorelick, W. J. Bosche, J. G. Levin and A. Rein (1996). "HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro." *Proc. Natl. Acad. Sci. U. S. A.* **93**(15): 7577-7581.

Fidorra, M., L. Duelund, C. Leidy, A. C. Simonsen and L. A. Bagatolli (2006). "Absence of fluid-ordered/fluid-disordered phase coexistence in ceramide/POPC mixtures containing cholesterol." *Biophys. J.* **90**(12): 4437-4451.

Finzi, A., A. Orthwein, J. Mercier and E. A. Cohen (2007). "Productive human immunodeficiency virus type 1 assembly takes place at the plasma membrane." *J. Virol.* **81**(14): 7476-7490.

Fiorentini, S., E. Marini, S. Caracciolo and A. Caruso (2006). "Functions of the HIV-1 matrix protein p17." *New Microbiol.* **29**(1): 1-10.

Fischer, W. B. (2003). "Vpu from HIV-1 on an atomic scale: experiments and computer simulations." *FEBS Lett.* **552**(1): 39-46.

Fisher, R. J., A. Rein, M. Fivash, M. A. Urbaneja, J. R. Casas-Finet, M. Medaglia and L. E. Henderson (1998). "Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides." *J. Virol.* **72**(3): 1902-1909.

Formosinho, S. J. and L. G. Arnaut (1993). "Excited-State Proton-Transfer Reactions.2. Intramolecular Reactions." *J. Photochem. Photobiol.* **75**(1): 21-48.

Foster, M. P., D. S. Wuttke, I. Radhakrishnan, D. A. Case, J. M. Gottesfeld and P. E. Wright (1997). "Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA." *Nat. Struct. Biol.* **4**(8): 605-608.

Foucault, M., K. Mayol, V. Receveur-Brechot, M. C. Bussat, C. Klinguer-Hamour, B. Verrier, A. Beck, R. Haser, P. Gouet and C. Guillon (2010). "UV and X-ray structural studies of a 101-residue long Tat protein from a HIV-1 primary isolate and of its mutated, detoxified, vaccine candidate." *Proteins-Structure Function And Bioinformatics* **78**(6): 1441-1456.

Fouchier, R. A. M., B. E. Meyer, J. H. M. Simon, U. Fischer, A. V. Albright, F. Gonzalez-Scarano and M. H. Malim (1998). "Interaction of the human immunodeficiency virus type 1 Vpr protein with the nuclear pore complex." *J. Virol.* **72**(7): 6004-6013.

Fouchier, R. A. M., B. E. Meyer, J. H. M. Simon, U. Fischer and M. H. Malim (1997). "HIV-1 infection of non-dividing cells: evidence that the amino-terminal basic region of the viral matrix protein is important for Gag processing but not for post-entry nuclear import." *Embo J.* **16**(15): 4531-4539.

Foulds, G. J. and F. A. Etzkorn (1998). "A capillary electrophoresis mobility shift assay for protein-DNA binding affinities free in solution." *Nucleic Acids Res.* **26**(18): 4304-4305.

Fox, C. B., J. R. Wayment, G. A. Myers, S. K. Endicott and J. M. Harris (2009). "Single-molecule fluorescence imaging of peptide binding to supported lipid bilayers." *Anal. Chem.* **81**(13): 5130-5138.

Frankel, A. D. and C. O. Pabo (1988). "Cellular uptake of the Tat protein from human immunodeficiency virus." *Cell* **55**(6): 1189-1193.

Frankel, A. D. and J. A. Young (1998). "HIV-1: fifteen proteins and an RNA." *Annu. Rev. Biochem.* **67**: 1-25.

Frederix, P., T. Akiyama, U. Staufer, C. Gerber, D. Fotiadis, D. J. Muller and A. Engel (2003). "Atomic force bio-analytics." *Curr. Opin. Chem. Biol.* **7**(5): 641-647.

Freed, E. O. (2002). "Viral late domains." *J. Virol.* **76**(10): 4679-4687.

Freed, E. O., D. J. Myers and R. Risser (1990). "Characterization of the fusion domain of the human immunodeficiency virus type-1 envelope glycoprotein Gp41." *Proc. Natl. Acad. Sci. U. S. A.* **87**(12): 4650-4654.

Fujinaga, K., T. P. Cujec, J. M. Peng, J. Garriga, D. H. Price, X. Grana and B. M. Peterlin (1998). "The ability of positive transcription elongation factor b to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat." *J. Virol.* **72**(9): 7154-7159.

Gaied, N. B., N. Glasser, N. Ramalanjaona, H. Beltz, P. Wolff, R. Marquet, A. Burger and Y. Mely (2005). "8-Vinyl-deoxyadenosine, an alternative fluorescent nucleoside analog to 2'-deoxyribosyl-2-aminopurine with improved properties." *Nucleic Acids Res.* **33**(3): 1031-1039.

Gait, M. J. (1984). Oligonucleotide Synthesis: A Practical Approach, Oxford University Press: Washington DC.

Galas, D. J. and A. Schmitz (1978). "DNAse footprinting - simple method for detection of protein-DNA binding specificity." *Nucleic Acids Res.* **5**(9): 3157-3170.

Gallay, P., T. Hope, D. Chin and D. Trono (1997). "HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway." *Proc. Natl. Acad. Sci. U. S. A.* **94**(18): 9825-9830.

Gallay, P., S. Swingler, J. P. Song, F. Bushman and D. Trono (1995). "HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase." *Cell* **83**(4): 569-576.

Gamble, T. R., F. F. Vajdos, S. H. Yoo, D. K. Worthylake, M. Houseweart, W. I. Sundquist and C. P. Hill (1996). "Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid." *Cell* **87**(7): 1285-1294.

Gamble, T. R., S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist and C. P. Hill (1997). "Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein." *Science* **278**(5339): 849-853.

Ganser-Pornillos, B. K., A. Cheng and M. Yeager (2007). "Structure of full-length HIV-1CA: a model for the mature capsid lattice." *Cell* **131**(1): 70-79.

Ganser-Pornillos, B. K., M. Yeager and W. I. Sundquist (2008). "The structural biology of HIV assembly." *Curr. Opin. Struct. Biol.* **18**(2): 203-217.

Gao, K., R. J. Gorelick, D. G. Johnson and F. Bushman (2003). "Cofactors for human immunodeficiency virus type 1 cDNA integration in vitro." *J. Virol.* **77**(2): 1598-1603.

Garner, M. M. and A. Revzin (1981). "A gel-electrophoresis method for quantifying the binding of proteins to specific DNA regions - application to components of the Escherichia-Coli lactose operon regulatory system." *Nucleic Acids Res.* **9**(13): 3047-3060.

Gatignol, A. (2007). "Transcription of HIV: Tat and cellular chromatin." *Adv. Pharmacol.* **55**: 137-59.

Gatignol, A., A. Kumar, A. Rabson and K. T. Jeang (1989). "Identification of cellular proteins that bind to the human immunodeficiency virus type 1 trans-activation-responsive TAR element RNA." *Proc. Natl. Acad. Sci. U. S. A.* **86**(20): 7828-7832.

Gell, C., D. Brockwell and A. Smith (2006). Handbook of Single Molecule Fluorescence Spectroscopy, Oxford University Press: New York.

Gheysen, D., E. Jacobs, F. Deforesta, C. Thiriart, M. Francotte, D. Thines and M. Dewilde (1989). "Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus infected insect cells." *Cell* **59**(1): 103-112.

Giacca, M. (2005). "HIV-1 Tat, apoptosis and the mitochondria: a tubulin link?" *Retrovirology* **2**.

Godde, F., J. J. Toulme and S. Moreau (1998). "Benzoquinazoline derivatives as substitutes for thymine in nucleic acid complexes. Use of fluorescence emission of benzo[g]quinazoline-2,4-(1H,3H)-dione in probing duplex and triplex formation." *Biochemistry* **37**(39): 13765-13775.

Godde, F., J. J. Toulme and S. Moreau (2000). "4-amino-1H-benzo[g]quinazoline-2-one: a fluorescent analog of cytosine to probe protonation sites in triplex forming oligonucleotides." *Nucleic Acids Res.* **28**(15): 2977-2985.

Godet, J. (2010). Etude du role de chaperon de la proteine NCp7 de la nucleocapside du virus VIH-1. Strasbourg, Université de Strasbourg.

Godet, J., H. de Rocquigny, C. Raja, N. Glasser, D. Ficheux, J. L. Darlix and Y. Mely (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**(5): 1180-1192.

Godet, J., N. Ramalanjaona, K. K. Sharma, L. Richert, H. de Rocquigny, J. L. Darlix, G. Duportail and Y. Mely (2011). "Specific implications of the HIV-1 nucleocapsid zinc fingers in the annealing of the primer binding site complementary sequences during the obligatory plus strand transfer." *Nucleic Acids Res.*: 1-13.

Goldschmidt, V. (2004). La rétrotranscription de HIV-1: étude du complexe d'initiation et des mécanismes de résistance aux inhibituers nucléosidiques. Strasbourg, Université Louis Pasteur.

Goncalves, M. S. (2009). "Fluorescent labeling of biomolecules with organic probes." *Chem. Rev.* **109**(1): 190-212.

Goodchild, J. (1990). "Conjugates of oligonucleotides and modified oligonucleotides: a review of their synthesis and properties." *Bioconjugate Chem.* **1**(3): 165-187.

Gordiyenko, Y. and C. V. Robinson (2008). "The emerging role of MS in structure elucidation of protein-nucleic acid complexes." *Biochem. Soc. Trans.* **36**: 723-731.

Gordon, J. P. (1973). "Radiation forces and momenta in dielectric media." *Phys. Rev., A* **8**: 14-21.

Gorelick, R. J., T. D. Gagliardi, W. J. Bosche, T. A. Wiltrout, L. V. Coren, D. J. Chabot, J. D. Lifson, L. E. Henderson and A. O. Arthur (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**(1): 92-104.

Gottlinger, H. G., T. Dorfman, E. A. Cohen and W. A. Haseltine (1993). "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.* **90**(15): 7381-7385.

Gousset, K., S. D. Ablan, L. V. Coren, A. Ono, F. Soheilian, K. Nagashima, D. E. Ott and E. O. Freed (2008). "Real-time visualization of HIV-1 GAG trafficking in infected macrophages." *PLoS Pathog.* **4**(3).

Green, M. and P. M. Loewenstein (1988). "Autonomous functional domains of chemically synthesized Human Immunodeficiency Virus Tat trans-activator protein." *Cell* **55**(6): 1179-1188.

Greenspan, P., E. P. Mayer and S. D. Fowler (1985). "Nile red: a selective fluorescent stain for intracellular lipid droplets." *J. Cell Biol.* **100**(3): 965-973.

Gregoire, C., J. M. Peloponese, D. Esquieu, S. Opi, G. Campbell, M. Solomiac, E. Lebrun, J. Lebreton and E. P. Loret (2001). "Homonuclear H-1-NMR assignment and structural characterization of Human Immunodeficiency Virus Type 1 Tat Mal protein." *Biopolymers* **62**(6): 324-335.

Griffin, B. A., S. R. Adams and R. Y. Tsien (1998). "Specific covalent labeling of recombinant protein molecules inside live cells." *Science* **281**(5374): 269-272.

Griffith, O. H., P. J. Dehlinger and S. P. Van (1974). "Shape of the hydrophobic barrier of phospholipid bilayers (evidence for water penetration in biological membranes)." *J. Membr. Biol.* **15**(2): 159-192.

Grigorov, B., F. Arcanger, P. Roingeard, J. L. Darlix and D. Muriaux (2006). "Assembly of infectious HIV-1 in human epithelial and T-lymphoblastic cell lines." *J. Mol. Biol.* **359**(4): 848-862.

Grigorov, B., V. Attuil-Audenis, F. Perugi, M. Nedelec, S. Watson, C. Pique, J. L. Darlix, H. Conjeaud and D. Muriaux (2009). "A role for CD81 on the late steps of HIV-1 replication in a chronically infected T cell line." *Retrovirology* **6**.

Grigorov, B., D. Decimo, F. Smagulova, C. Pechoux, M. Mougel, D. Muriaux and J. L. Darlix (2007). "Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers." *Retrovirology* **4**.

Guo, F., J. Saadatmand, M. J. Niu and L. Kleiman (2009). "Roles of Gag and NCp7 in facilitating tRNA(3)(Lys) annealing to viral RNA in Human Immunodeficiency Virus type 1." *J. Virol.* **83**(16): 8099-8107.

Guo, J. H., T. Y. Wu, J. Anderson, B. F. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson and J. G. Levin (2000). "Zinc finger structures in the human immunodeficiency virus type 1 nucleocapsid protein facilitate efficient minus- and plus-strand transfer." *J. Virol.* **74**(19): 8980-8988.

Guo, J. H., T. Y. Wu, B. F. Kane, D. G. Johnson, L. E. Henderson, R. J. Gorelick and J. G. Levin (2002). "Subtle alterations of the native zinc finger structures have dramatic effects on the nucleic acid chaperone activity of human immunodeficiency virus type 1 nucleocapsid protein." *J. Virol.* **76**(9): 4370-4378.

Guszczynski, T. and T. D. Copeland (1998). "A binding shift assay for the zinc-bound and zinc-free HIV-1 nucleocapsid protein by capillary electrophoresis." *Anal. Biochem.* **260**(2): 212-217.

Halford, S. E. and J. F. Marko (2004). "How do site-specific DNA-binding proteins find their targets?" *Nucleic Acids Res.* **32**(10): 3040-3052.

Hall, K. B. and J. K. Kranz (1999). "Nitrocellulose filter binding for determination of dissociation constants." *Methods in molecular biology* **118**: 105-114.

Hallbrink, M., A. Floren, A. Elmquist, M. Pooga, T. Bartfai and U. Langel (2001). "Cargo delivery kinetics of cell-penetrating peptides." *Biochim. Biophys. Acta-Biomembr.* **1515**(2): 101-109.

Harada, Y., T. Funatsu, K. Murakami, Y. Nonoyama, A. Ishihama and T. Yanagida (1999). "Single-molecule imaging of RNA polymerase-DNA interactions in real time." *Biophys. J.* **76**(2): 709-715.

Hargittai, M. R. S., R. J. Gorelick, L. Rouzina and K. Musier-Forsyth (2004). "Mechanistic insights into the kinetics of HIV-1 nucleocapsid protein-facilitated tRNA annealing to the primer binding site." *J. Mol. Biol.* **337**(4): 951-968.

Hargittai, M. R. S., A. T. Mangla, R. J. Gorelick and K. Musier-Forsyth (2001). "HIV-1 nucleocapsid protein zinc finger structures induce tRNA(Lys,3) structural changes but are not critical for primer/template annealing." *J. Mol. Biol.* **312**(5): 985-997.

Harikumar, K. G., D. I. Pinon, W. S. Wessels, F. G. Prendergast and L. J. Miller (2002). "Environment and mobility of a series of fluorescent reporters at the amino terminus of structurally related peptide agonists and antagonists bound to the cholecystokinin receptor." *J. Biol. Chem.* **277**(21): 18552-18560.

Harrich, D., C. Ulich and R. B. Gaynor (1996). "A critical role for the TAR element in promoting efficient human immunodeficiency virus type 1 reverse transcription." *J. Virol.* **70**(6): 4017-4027.

Haskard, C. A. and E. C. Y. Li-Chan (1998). "Hydrophobicity of bovine serum albumin and ovalbumin determined using uncharged (PRODAN) and anionic (ANS-) fluorescent probes." *J. Agricult. Food Chem.* **46**(7): 2671.

Haugland, R. P. (2005). The Handbook: A Guide to Fluorescent Probes and Labeling Technologies, Molecular Probes.

Hawkins, M. E., W. Pfleiderer, O. Jungmann and F. M. Balis (2001). "Synthesis and fluorescence characterization of pteridine adenosine nucleoside analogs for DNA incorporation." *Anal. Biochem.* **298**(2): 231-240.

Hayashi, T., T. Shioda, Y. Iwakura and H. Shibuta (1992). "RNA packaging signal of human immunodeficiency virus type 1." *Virology* **188**(2): 590-599.

Heaphy, S., J. T. Finch, M. J. Gait, J. Karn and M. Singh (1991). "Human Immunodeficiency Virus type-1 regulator of virion expression, Rev, forms nucleoprotein filaments after binding to a purine-rich bubble located within the Rev-responsive region of viral messenger-RNAs." *Proc. Natl. Acad. Sci. U. S. A.* **88**(16): 7366-7370.

Hearps, A. C. and D. A. Jans (2007). "Regulating the functions of the HIV-1 matrix protein." *Aids Res. Hum. Retrovir.* **23**(3): 341-346.

Hecht, S. M. (1996). Bioorganic Chemistry: Nucleic Acids. New York, Oxford University Press.

Helga-Maria, C., M. L. Hammarskjold and D. Rekosh (1999). "An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type 1 genomic RNA." *J Virol* **73**(5): 4127-4135.

Hellman, L. M. and M. G. Fried (2007). "Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions." *Nat. Protoc.* **2**(8): 1849-1861.

Helwa, R. and J. D. Hoheisel (2010). "Analysis of DNA-protein interactions: from nitrocellulose filter binding assays to microarray studies." *Anal. Bioanal. Chem.* **398**(6): 2551-2561.

Henderson, L. E., M. A. Bowers, R. C. Sowder, S. A. Serabyn, D. G. Johnson, J. W. Bess, L. O. Arthur, D. K. Bryant and C. Fenselau (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**(4): 1856-1865.

Henriet, S., L. Sinck, G. Bec, R. J. Gorelick, R. Marquet and J. C. Paillart (2007). "Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription." *Nucleic Acids Res.* **35**(15): 5141-5153.

Henriques, S. T., J. Costa and M. Castanho (2005). "Re-evaluating the role of strongly charged sequences in amphipathic cell-penetrating peptides: a fluorescence study using Pep-1." *FEBS Lett.* **579**(20): 4498-4502.

Henriques, S. T., M. N. Melo and M. Castanho (2006). "Cell-penetrating peptides and antimicrobial peptides: how different are they?" *Biochem. J.* **399**: 1-7.

Herek, J. L., S. Pedersen, L. Banares and A. H. Zewail (1992). "Femtosecond real-time probing of reactions. X. Hydrogen atom transfer." *J. Chem. Phys.* **97**: 9046-9061.

Herrmann, C. H. and A. P. Rice (1995). "Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor." *J. Virol.* **69**(3): 1612-1620.

Hershey, A. D. and M. Chase (1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage." *J. Gen. Physiol.* **36**(1): 39-56.

Hiratsuka, T. (1998). "Prodan fluorescence reflects differences in nucleotide-induced conformational states in the myosin head and allows continuous visualization of the ATPase reactions." *Biochemistry* **37**(20): 7167-7176.

Hiratsuka, T. (1999). "ATP-induced opposite changes in the local environments around Cys(697) (SH2) and Cys(707) (SH1) of the myosin motor domain revealed by the prodan fluorescence." *J. Biol. Chem.* **274**(41): 29156-29163.

Holmes-Farley, S. R. and G. M. Whitesides (1986). "Fluorescence properties of dansyl groups covalently bonded to the surface of oxidatively functionalized low-density polyethylene film." *Langmuir* **2**(3): 266-281.

Horn, R. G. (1984). "Direct measurement of the force between two lipid bilayers and observation of their fusion." *Biochim. Biophys. Acta* **778**(1): 224-228.

Horton, N. C., L. F. Dorner and J. J. Perona (2002). "Sequence selectivity and degeneracy of a restriction endonuclease mediated by DNA intercalation." *Nat. Struct. Biol.* **9**(1): 42-47.

Hristova, K., C. E. Dempsey and S. H. White (2001). "Structure, location, and lipid perturbations of melittin at the membrane interface." *Biophys. J.* **80**(2): 801-811.

Huang, C. C., M. Tang, M. Y. Zhang, S. Majeed, E. Montabana, R. L. Stanfield, D. S. Dimitrov, B. Korber, J. Sodroski, I. A. Wilson, R. Wyatt and P. D. Kwong (2005). "Structure of a V3-containing HIV-1 gp120 core." *Science* **310**(5750): 1025-1028.

Huang, H. F., R. Chopra, G. L. Verdine and S. C. Harrison (1998). "Structure of a covalently trapped catalytic complex of HIV-I reverse transcriptase: Implications for drug resistance." *Science* **282**(5394): 1669-1675.

Huang, H. W. (2006). "Molecular mechanism of antimicrobial peptides: The origin of cooperativity." *Biochim. Biophys. Acta-Biomembr.* **1758**(9): 1292-1302.

Huang, H. W. (2009). "Free Energies of Molecular Bound States in Lipid Bilayers: Lethal Concentrations of Antimicrobial Peptides." *Biophys. J.* **96**(8): 3263-3272.

Hudson, B. P., M. A. Martinez-Yamout, H. J. Dyson and P. E. Wright (2004). "Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d." *Nat. Struct. Mol. Biol.* **11**(3): 257-264.

Hultmark, D., H. Steiner, T. Rasmuson and H. G. Boman (1980). "Insect immunity - purification and properties of three inducible bactericidal proteins from hemolymph of immunized pupae of *Hyalophora-Cecropia*." *Eur. J. Biochem.* **106**(1): 7-16.

Hunter, H. N., W. G. Jing, D. J. Schibli, T. Trinh, I. Y. Park, S. C. Kim and H. J. Vogel (2005). "The interactions of antimicrobial peptides derived from lysozyme with model membrane systems." *Biochim. Biophys. Acta-Biomembr.* **1668**(2): 175-189.

Hwang, G. T., Y. J. Seo and B. H. Kim (2005). "Pyrene-labeled deoxyuridine and deoxyadenosine: fluorescent discriminating phenomena in their oligonucleotides." *Tetrahedron Lett.* **46**(9): 1475-1477.

Isel, C. and J. Karn (1999). "Direct evidence that MIV-1 tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation." *J. Mol. Biol.* **290**(5): 929-941.

Jackson, D. Y., J. Burnier, C. Quan, M. Stanley, J. Tom and J. A. Wells (1994). "A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues." *Science* **266**(5183): 243-247.

Jacotot, E., L. Ravagnan, M. Loeffler, K. F. Ferri, H. L. A. Vieira, N. Zamzami, P. Costantini, S. Druillennec, J. Hoebeke, J. P. Briand, T. Irinopoulou, E. Daugas, S. A. Susin, D. Cointe, Z. H. Xie, J. C. Reed, B. P. Roques and G. Kroemer (2000). "The HIV-1 viral protein R induces apoptosis via a direct effect on the mitochondrial permeability transition pore." *J. Exp. Med.* **191**(1): 33-45.

Jeang, K. T., H. Xiao and E. K. Rich (1999). "Multifaceted activities of the HIV-1 transactivator of transcription, Tat." *J. Biol. Chem.* **274**(41): 28837-28840.

Jenkins, Y., O. Pornillos, R. L. Rich, D. G. Myszka, W. I. Sundquist and M. H. Malim (2001). "Biochemical analyses of the interactions between human immunodeficiency virus type 1 Vpr and p6(Gag)." *J Virol* **75**(21): 10537-10542.

Ji, X. D., G. J. Klarmann and B. D. Preston (1996). "Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro." *Biochemistry* **35**(1): 132-143.

Jiao, G. S. and K. Burgess (2003). "Oligonucleotides with strongly fluorescent groups pi-conjugated to a nucleobase: syntheses, melting temperatures, and conformation." *Bioorg. Med. Chem. Lett.* **13**(16): 2785-2788.

Jimenez-Menendez, N., P. Fernandez-Millan, A. Rubio-Cosials, C. Arnan, J. Montoya, H. T. Jacobs, P. Bernado, M. Coll, I. Uson and M. Sola (2010). "Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat." *Nat. Struct. Mol. Biol.* **17**(7): 891-893.

Jing, D., J. Agnew, W. F. Patton, J. Hendrickson and J. M. Beechem (2003). "A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels." *Proteomics* **3**(7): 1172-1180.

Johnson, P. E., R. B. Turner, Z. R. Wu, L. Hairston, J. H. Guo, J. G. Levin and M. F. Summers (2000). "A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription." *Biochemistry* **39**(31): 9084-9091.

Joliot, A., C. Pernelle, H. Deagostinibazin and A. Prochiantz (1991). "Antennapedia homeobox peptide regulates neural morphogenesis." *Proc. Natl. Acad. Sci. U. S. A.* **88**(5): 1864-1868.

Jouvenet, N., P. D. Bieniasz and S. M. Simon (2008). "Imaging the biogenesis of individual HIV-1 virions in live cells." *Nature* **454**(7201): 236-240.

Jouvenet, N., S. J. D. Neil, C. Bess, M. C. Johnson, C. A. Virgen, S. M. Simon and P. D. Bieniasz (2006). "Plasma membrane is the site of productive HIV-1 particle assembly." *Plos Biology* **4**(12): 2296-2310.

Kachel, K., E. Asuncion-Punzalan and E. London (1998). "The location of fluorescence probes with charged groups in model membranes." *Biochim. Biophys. Acta* **1374**(1-2): 63-76.

Kaiser, R. D. and E. London (1998). "Location of diphenylhexatriene (DPH) and its derivatives within membranes: Comparison of different fluorescence quenching analyses of membrane depth." *Biochemistry* **37**(22): 8180-8190.

Kalodimos, C. G., N. Biris, A. Bonvin, M. M. Levandoski, M. Guennuegues, R. Boelens and R. Kaptein (2004). "Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes." *Science* **305**(5682): 386-389.

Kapanidis, A. N., Y. W. Ebright and R. H. Ebright (2001). "Site-specific incorporation of fluorescent probes into protein: hexahistidine-tag-mediated fluorescent labeling with (Ni(2+):nitrilotriacetic Acid (n)-fluorochrome conjugates." *J Am Chem Soc* **123**(48): 12123-5.

Kaplan, A. H., M. Manchester and R. Swanstrom (1994). "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.* **68**(10): 6782-6786.

Kaptein, R., E. R. P. Zuiderweg, R. M. Scheek, R. Boelens and W. F. Vangunsteren (1985). "A protein-structure from nuclear magnetic-resonance data - Lac repressor headpiece." *J. Mol. Biol.* **182**(1): 179-182.

Karn, J. (1999). "Tackling Tat." *J. Mol. Biol.* **293**(2): 235-254.

Karunatilaka, K. S., A. Solem, A. M. Pyle and D. Rueda (2010). "Single-molecule analysis of Mss116-mediated group II intron folding." *Nature* **467**(7318): 935-975.

Kaur, Y. and P. M. Horowitz (2004). "Prodan fluorescence mimics the GroEL folding cycle." *Protein J* **23**(7): 475-481.

Kawai, R., M. Kimoto, S. Ikeda, T. Mitsui, M. Endo, S. Yokoyama and L. Hirao (2005). "Site-specific fluorescent labeling of RNA molecules by specific transcription using unnatural base pairs." *J. Am. Chem. Soc.* **127**(49): 17286-17295.

Kenfack, C. A., A. Burger and Y. Mely (2006). "Excited-state properties and transitions of fluorescent 8-vinyl adenosine in DNA." *J. Phys. Chem. B* **110**(51): 26327-26336.

Khan, R. and D. P. Giedroc (1992). "Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA." *J Biol Chem* **267**(10): 6689-6695.

Khandelia, H., J. H. Ipsen and O. G. Mouritsen (2008). "The impact of peptides on lipid membranes." *Biochim. Biophys. Acta-Biomembr.* **1778**(7-8): 1528-1536.

Khorchid, A., H. Javanbakht, S. Wise, R. Halwani, M. A. Parniak, M. A. Wainberg and L. Kleiman (2000). "Sequences within Pr160gag-pol affecting the selective packaging of primer tRNA(Lys3) into HIV-1." *J. Mol. Biol.* **299**(1): 17-26.

Kim, H. J., K. A. Martemyanov and S. A. Thayer (2008). "Human immunodeficiency virus protein Tat induces synapse loss via a reversible process that is distinct from cell death." *J. Neurosci.* **28**(48): 12604-12613.

Kim, J. L., D. B. Nikolov and S. K. Burley (1993). "Co-crystal structure of TBP recognizing the minor-groove of A TATA element." *Nature* **365**(6446): 520-527.

Kim, T. A., H. K. Avraham, Y. H. Koh, S. X. Jiang, I. W. Park and S. Avraham (2003). "HIV-1 Tat-mediated apoptosis in human brain microvascular endothelial cells." *J. Immunol.* **170**(5): 2629-2637.

Klasens, B. I. F., H. T. Huthoff, A. T. Das, R. E. Jeeninga and B. Berkhout (1999). "The effect of template RNA structure on elongation by HIV-1 reverse transcriptase." *Biochim. Biophys. Acta* **1444**(3): 355-370.

Klasens, B. I. F., M. Thiesen, A. Virtanen and B. Berkhout (1999). "The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure." *Nucleic Acids Res.* **27**(2): 446-454.

Kleiman, L. (2002). "tRNA(Lys3): The primer tRNA for reverse transcription in HIV-1." *IUBMB Life* **53**(2): 107-114.

Klymchenko, A., T. Ozturk, V. G. Pivovarenko and A. Demchenko (2002). "Synthesis of furanochromones: a new step in improvement of fluorescence properties." *Tetrahedron Lett* **43**: 7079-7082.

Klymchenko, A., V. G. Pivovarenko and A. Demchenko (2003). "Elimination of the hydrogen bonding effect on the solvatochromism of 3-hydroxyflavones." *J. Phys. Chem.* **107**: 4211-4216.

Klymchenko, A. S. and A. P. Demchenko (2003). "Multiparametric probing of intermolecular interactions with fluorescent dye exhibiting excited state intramolecular proton transfer." *Phys. Chem. Chem. Phys.* **5**: 461-468.

Klymchenko, A. S., G. Duportail, A. P. Demchenko and Y. Mely (2004). "Bimodal distribution and fluorescence response of environment-sensitive probes in lipid bilayers." *Biophys. J.* **86**(5): 2929-2941.

Klymchenko, A. S., Y. Mely, A. P. Demchenko and G. Duportail (2004). "Simultaneous probing of hydration and polarity of lipid bilayers with 3-hydroxyflavone fluorescent dyes." *Biochim. Biophys. Acta* **1665**(1-2): 6-19.

Klymchenko, A. S., S. Oncul, P. Didier, E. Schaub, L. Bagatolli, G. Duportail and Y. Mely (2009). "Visualization of lipid domains in giant unilamellar vesicles using an environment-sensitive membrane probe based on 3-hydroxyflavone." *Biochim. Biophys. Acta* **1788**: 495-499.

Klymchenko, A. S., V. G. Pivovarenko and A. P. Demchenko (2003). "Elimination of the Hydrogen Bonding Effect on the Solvatochromism of 3-Hydroxyflavones." *J. Phys. Chem. A* **107**: 4211-4216.

Klymchenko, A. S., V. G. Pivovarenko and A. P. Demchenko (2003). "Perturbation of planarity as the possible mechanism of solvent-dependent variations of fluorescence quantum yield in 2-aryl-3-hydroxychromones." *Spectrochim Acta A* **59**(4): 787-792.

Klymchenko, A. S., V. V. Shvadchak, D. A. Yushchenko, N. Jain and Y. Mely (2008). "Excited-state intramolecular proton transfer distinguishes microenvironments in single-and double-stranded DNA." *J. Phys. Chem. B* **112**(38): 12050-12055.

Koch, S. J., A. Shundrovsky, B. C. Jantzen and M. D. Wang (2002). "Probing protein-DNA interactions by unzipping a single DNA double helix." *Biophys. J.* **83**(2): 1098-1105.

Kochaniak, A. B., S. Habuchi, J. J. Loparo, D. J. Chang, K. A. Cimprich, J. C. Walter and A. M. van Oijen (2009). "Proliferating Cell Nuclear Antigen Uses Two Distinct Modes to Move along DNA." *J. Biol. Chem.* **284**(26): 17700-17710.

Kochetkov, N. K., V. N. Shibaev and A. A. Kost (1971). "New reaction of adenine and cytosine derivatives, potentially useful for nucleic acids modification." *Tetrahedron Lett.* **12**(22): 1993-1996.

Kohn, M. and R. Breinbauer (2004). "The Staudinger ligation-a gift to chemical biology." *Angew. Chem. Int. Ed. Engl.* **43**(24): 3106-3116.

Kondo, E., F. Mammano, E. A. Cohen and H. G. Gottlinger (1995). "The P6(Gag) domain of human-immunodeficiency-virus type-1 is sufficient for the incorporation of Vpr into heterologous viral particles." *J. Virol.* **69**(5): 2759-2764.

Korshun, V. A., I. A. Prokhorenko, S. V. Gontarev, M. V. Skorobogatyi, K. V. Balakin, E. V. Manasova, A. D. Malakhov and Y. A. Berlin (1997). "New pyrene derivatives for fluorescent labeling of oligonucleotides." *Nucleosides Nucleotides* **16**(7-9): 1461-1464.

Koudelka, G. B. and P. Carlson (1992). "DNA twisting and the Effects of noncontacted bases on affinity of 434 operator for 434 repressor." *Nature* **355**(6355): 89-91.

Kourie, J. I. and A. A. Shorthouse (2000). "Properties of cytotoxic peptide-formed ion channels." *Am. J. Physiol.-Cell Physiol.* **278**(6): C1063-C1087.

Kreger, A. S., K. S. Kim, Zaboretz.F and Bernheim.Aw (1971). "Purification and properties of staphylococcal delta hemolysin." *Infect. Immun.* **3**(3): 449-465.

Krishnamoorthy, G., B. Roques, J. L. Darlix and Y. Mely (2003). "DNA condensation by the nucleocapsid protein of HIV-1: a mechanism ensuring DNA protection." *Nucleic Acids Res.* **31**(18): 5425-5432.

Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling and T. R. Cech (1982). "Self-splicing RNA - auto-excision and auto-cyclization of the ribosomal-RNA intervening sequence of tetrahymena." *Cell* **31**(1): 147-157.

Kucherak, O. A., P. Didier, Y. Mely and A. S. Klymchenko (2010). "Fluorene analogues of Prodan with superior fluorescence brightness and solvatochromism." *J. Phys. Chem. Lett.* **1**(3): 616-620.

Kuciak, M., C. Gabus, R. Ivanyi-Nagy, K. Semrad, R. Storchak, O. Chaloin, S. Muller, Y. Mely and J. L. Darlix (2008). "The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities in vitro." *Nucleic Acids Res.* **36**(10): 3389-3400.

Kuppuswamy, M., T. Subramanian, A. Srinivasan and G. Chinnadurai (1989). "Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis." *Nucleic Acids Res.* **17**(9): 3551-3561.

Ladokhin, A. S. (1999). "Analysis of protein and peptide penetration into membranes by depth- dependent fluorescence quenching: theoretical considerations." *Biophys. J.* **76**(2): 946-955.

Ladokhin, A. S. and S. H. White (2001). "'Detergent-like' permeabilization of anionic lipid vesicles by melittin." *Biochim. Biophys. Acta-Biomembr.* **1514**(2): 253-260.

Lakowicz, J. R. (1999). "Principles of Fluorescence Spectroscopy." New York, Kluwer Academic.

Lakowicz, J. R. (2006). Principles of Fluorescence Spectroscopy. New York, Springer-Verlag.

Lam, W. C., A. H. Maki, J. R. Casasfinet, J. W. Erickson, B. P. Kane, R. C. Sowder and L. E. Henderson (1994). "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* **33**(35): 10693-10700.

Lamaziere, A., C. Wolf, O. Lambert, G. Chassaing, G. Trugnan and J. Ayala-Sanmartin (2008). "The homeodomain derived peptide penetratin induces curvature of fluid membrane domains." *PLoS One* **3**(4).

Lamoureux, J. S., D. Stuart, R. Tsang, C. Wu and J. N. M. Glover (2002). "Structure of the sporulation-specific transcription factor Ndt80 bound to DNA." *EMBO J.* **21**(21): 5721-5732.

Langel, U. (2007). Handbook of cell-penetrating peptides, Taylor & Francis.

Lapadattapolsky, M., H. de Rocquigny, D. Vangent, B. Roques, R. Plasterk and J. L. Darlix (1993). "Interactions between HIV-1 nucleocapsid protein and viral-DNA may have important functions in the viral life cycle." *Nucleic Acids Res.* **21**(4): 831-839.

Lapadattapolsky, M., C. Pernelle, C. Borie and J. L. Darlix (1995). "Analysis of the nucleic-acid annealing activities of nucleocapsid protein from HIV-1." *Nucleic Acids Res.* **23**(13): 2434-2441.

Lawson, C. L., D. Swigon, K. S. Murakami, S. A. Darst, H. M. Berman and R. H. Ebright (2004). "Catabolite activator protein: DNA binding and transcription activation." *Curr. Opin. Struct. Biol.* **14**(1): 10-20.

Le Rouzic, E. and S. Benichou (2005). "The Vpr protein from HIV-1: distinct roles along the viral life cycle." *Retrovirology* **2**.

Lee, B. M., R. N. de Guzman, B. G. Turner, N. Tjandra and M. F. Summers (1998). "Dynamical behavior of the HIV-1 nucleocapsid protein." *J. Mol. Biol.* **279**(3): 633-649.

Lee, M. T., F. Y. Chen and H. W. Huang (2004). "Energetics of pore formation induced by membrane active peptides." *Biochemistry* **43**(12): 3590-3599.

Lee, M. T., W. C. Hung, F. Y. Chen and H. W. Huang (2005). "Many-body effect of antimicrobial peptides: on the correlation between lipid's spontaneous curvature and pore formation." *Biophys. J.* **89**(6): 4006-4016.

Lehrer, S. S. and Y. Ishii (1988). "Fluorescence properties of acrylodan-labeled tropomyosin and tropomyosin-actin: evidence for myosin subfragment 1 induced changes in geometry between tropomyosin and actin." *Biochemistry* **27**(16): 5899-5906.

Leitgeb, B., A. Szekeres, L. Manczinger, C. Vagvolgyi and L. Kredics (2007). "The history of alamethicin: a review of the most extensively studied peptaibol." *Chem. Biodivers.* **4**(6): 1027-1051.

Levin, J. G., J. Guo, I. Rouzina and K. Musier-Forsyth (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.

Lewin, M., N. Carlesso, C. H. Tung, X. W. Tang, D. Cory, D. T. Scadden and R. Weissleder (2000). "Tat peptide-derivatized magnetic nanoparticles allow in vivo tracking and recovery of progenitor cells." *Nat. Biotechnol.* **18**(4): 410-414.

Lewis, B. A. and D. M. Engelman (1983). "Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles." *J. Mol. Biol.* **166**(2): 211-217.

Lewis, H. A., K. Musunuru, K. B. Jensen, C. Edo, H. Chen, R. B. Darnell and S. K. Burley (2000). "Sequence-specific RNA binding by a Nova KH domain: Implications for paraneoplastic disease and the fragile X syndrome." *Cell* **100**(3): 323-332.

Li, C. Z. and L. M. Martin (1998). "A robust method for determining DNA binding constants using capillary zone electrophoresis." *Anal. Biochem.* **263**(1): 72-78.

Limon, A., E. Devroe, R. Lu, H. Z. Ghory, P. A. Silver and A. Engelman (2002). "Nuclear localization of human immunodeficiency virus type 1 preintegration complexes (PICs): V165A and R166A are pleiotropic integrase mutants primarily defective for integration, not PIC nuclear import." *J. Virol.* **76**(21): 10598-10607.

Limon, A., N. Nakajima, R. Lu, H. Z. Ghory and A. Engelman (2002). "Wild-type levels of nuclear localization and human immunodeficiency virus type 1 replication in the absence of the central DNA flap." *J. Virol.* **76**(23): 12078-12086.

Lin, M. Z. and L. Wang (2008). "Selective labeling of proteins with chemical probes in living cells." *Physiology (Bethesda)* **23**: 131-141.

Lippert, E. L. (1975). Laser-spectroscopic studies of reorientation and other relaxation processes in solution. Organic Molecular Photophysics. J. B. Birks. New York, Wiley. **2:** 1-31.

Loakes, D., J. Gallego, V. B. Pinheiro, E. T. Kool and P. Holliger (2009). "Evolving a polymerase for hydrophobic base analogues." *J. Am. Chem. Soc.* **131**(41): 14827-14837.

Lohner, K., E. Sevcsik and G. Pasb (2008). Liposome-based biomembrane mimetic systems: implications for peptide-lipid interactions. Amsterdam, Elsevier.

Loret, E. P., P. Georgel, W. C. Johnson, Jr. and P. S. Ho (1992). "Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 trans-activation response RNA and the binding region of Tat, the trans-acting transcriptional activator." *Proc. Natl. Acad. Sci. U. S. A.* **89**(20): 9734-9738.

Lorsch, J. R. (2002). "RNA chaperones exist and DEAD box proteins get a life." *Cell* **109**(7): 797-800.

Loudet, A. and K. Burgess (2007). "BODIPY dyes and their derivatives: syntheses and spectroscopic properties." *Chem. Rev.* **107**(11): 4891-4932.

Loving, G. and B. Imperiali (2008). "A versatile amino acid analogue of the solvatochromic fluorophore 4-N,N-dimethylamino-1,8-naphthalimide: A powerful tool for the study of dynamic protein interactions." *J. Am. Chem. Soc.* **130**(41): 13630-13638.

Loving, G. S., M. Sainlos and B. Imperiali (2010). "Monitoring protein interactions and dynamics with solvatochromic fluorophores." *Trends Biotechnol.* **28**(2): 73-83.

Lu, Z., S. J. Lord, H. Wang, W. E. Moerner and R. J. Twieg (2006). "Long-wavelength analogue of PRODAN: synthesis and properties of Anthradan, a fluorophore with a 2,6-donor-acceptor anthracene structure." *J. Org. Chem.* **71**(26): 9651-9657.

Ludtke, S., K. He and H. Huang (1995). "Membrane thinning caused by magainin 2." *Biochemistry* **34**(51): 16764-16769.

Ludtke, S. J., K. He, W. T. Heller, T. A. Harroun, L. Yang and H. W. Huang (1996). "Membrane pores induced by magainin." *Biochemistry* **35**(43): 13723-13728.

Lunde, B. M., C. Moore and G. Varani (2007). "RNA-binding proteins: modular design for efficient function." *Nat. Rev. Mol. Cell Biol.* **8**(6): 479-490.

Luscombe, N. M., R. A. Laskowski and J. M. Thornton (2001). "Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level." *Nucleic Acids Res.* **29**(13): 2860-2874.

Malim, M. H., S. Bohnlein, R. Fenrick, S. Y. Le, J. V. Maizel and B. R. Cullen (1989). "Functional comparison of the Rev trans-activators encoded by different primate immunodeficiency virus species." *Proc. Natl. Acad. Sci. U. S. A.* **86**(21): 8222-8226.

Malim, M. H., L. S. Tiley, D. F. McCarn, J. R. Rusche, J. Hauber and B. R. Cullen (1990). "HIV-1 structural gene-expression requires binding of the Rev trans-activator to its RNA target sequence." *Cell* **60**(4): 675-683.

Marassi, F. M. and S. J. Opella (2000). "A solid-state NMR index of helical membrane protein structure and topology." *J. Magn. Reson.* **144**(1): 150-155.

Marsden, M. D. and J. A. Zack (2007). "Human immunodeficiency virus bearing a disrupted central DNA flap is pathogenic in vivo." *J. Virol.* **81**(11): 6146-6150.

Mashanov, G. I. and J. E. Molloy (2007). "Automatic detection of single fluorophores in live cells." *Biophys. J.* **92**(6): 2199-2211.

Mataga, N. (1970). Molecular Interactions and Electronic Spectra. New York, Marcel Dekker.

Mataga, N. and T. Kubota (1970). Molecular Interactions and Electronic Spectra. New York, Marcel Dekker.

Matos, P. M., H. G. Franquelim, M. Castanho and N. C. Santos (2010). "Quantitative assessment of peptide-lipid interactions. Ubiquitous fluorescence methodologies." *Biochim. Biophys. Acta-Biomembr.* **1798**(11): 1999-2012.

Matsuzaki, K. (1999). "Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes." *Biochim. Biophys. Acta-Biomembr.* **1462**(1-2): 1-10.

Matsuzaki, K., O. Murase, N. Fujii and K. Miyajima (1996). "An antimicrobial peptide, magainin 2, induced rapid flip-flop of phospholipids coupled with pore formation and peptide translocation." *Biochemistry* **35**(35): 11361-11368.

McDonald, D., M. A. Vodicka, G. Lucero, T. M. Svitkina, G. G. Borisy, M. Emerman and T. J. Hope (2002). "Visualization of the intracellular behavior of HIV in living cells." *J. Cell Biol.* **159**(3): 441-452.

McKay, D. B. and T. A. Steitz (1981). "Structure of catabolite gene activator protein at 2.9 a resolution suggests binding to left-handed B-DNA." *Nature* **290**(5809): 744-749.

McMorrow, D. and M. Kasha (1984). "Intramolecular excited-state proton transfer in 3-hydroxyflavone. Hydrogen-bonding solvent perturbations." *J. Phys. Chem.* **88**: 2235-2243.

Melo, M. N., R. Ferre and M. Castanho (2009). " Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations." *Nat. Rev. Microbiol.* **7**(3): 245-250.

Mely, Y., H. de Rocquigny, N. Morellet, B. P. Roques and D. Gerard (1996). "Zinc binding to the HIV-1 nucleocapsid protein: a thermodynamic investigation by fluorescence spectroscopy." *Biochemistry* **35**(16): 5175-5182.

Mely, Y., H. de Rocquigny, M. Sorinasjimeno, G. Keith, B. P. Roques, R. Marquet and D. Gerard (1995). "Binding of the HIV-1 nucleocapsid protein to the primer tRNA(3)(Lys), in-vitro, is essentially not specific." *J. Biol. Chem.* **270**(4): 1650-1656.

Mély, Y., N. Jullian, N. Morellet, H. de Rocquigny, C. Z. Dong, E. Piémont, B. P. Roques and D. Gérard (1994). "Spatial proximity of the HIV-1 nucleocapsid protein zinc fingers investigated by time-resolved fluorescence and fluorescence resonance energy transfer." *Biochemistry* **33**(40): 12085-12091.

Mely, Y., E. Piemont, M. Sorinasjimeno, H. de Rocquigny, N. Jullian, N. Morellet, B. P. Roques and D. Gerard (1993). "Structural and dynamic characterization of the aromatic-amino-acids of the human immunodeficiency virus type-1 nucleocapsid protein zinc fingers and their involvement in heterologous transfer RNA(Phe) binding - a steady-state and time-resolved fluorescence study." *Biophys. J.* **65**(4): 1513-1522.

Merrifield, R. B. (1963). "Solid phase peptide synthesis. The synthesis of a tetrapeptide." *J. Am. Chem. Soc.* **85**: 149-154.

Mertens, H. D. T. and D. I. Svergun (2010). "Structural characterization of proteins and complexes using small-angle X-ray solution scattering." *J. Struct. Biol.* **172**(1): 128-141.

Mes, G. F., B. De Jong, H. J. Van Ramesdonk, J. W. Verhoeven, J. M. Warman, M. P. De Haas and L. E. W. Horsman-van Den Dool (1984). "Excited-state dipole moment and solvatochromism of highly fluorescent rod-shaped bichromophoric molecules." *J. Am. Chem. Soc.* **106**(22): 6524-6528.

Meyer, M. and D. Schomburg (1999). "Protein interactions." *in Biotechnology, Recombinant Proteins* **5a**: 88-107.

Michalet, X., A. N. Kapanidis, T. Laurence, F. Pinaud, S. Doose, M. Pflughoefft and S. Weiss (2003). "The power and prospects of fluorescence microscopies and spectroscopies." *Annu. Rev. Biophys. Biomolec. Struct.* **32**: 161-182.

Mirambeau, G., S. Lyonnais, D. Coulaud, L. Hameau, S. Lafosse, J. Jeusset, I. Borde, M. Reboud-Ravaux, T. Restle, R. J. Gorelick and E. Le Cam (2007). "HIV-1 protease and reverse transcriptase control the architecture of their nucleocapsid partner." *Plos One* **2**(8) : e669.

Mirambeau, G., S. Lyonnais, D. Coulaud, L. Hameau, S. Lafosse, J. Jeusset, A. Justome, E. Delain, R. J. Gorelick and E. Le Cam (2006). "Transmission electron microscopy reveals an optimal HIV-1 nucleocapsid aggregation with single-stranded nucleic acids and the mature HIV-1 nucleocapsid protein." *J. Mol. Biol.* **364**(3): 496-511.

Mishra, A., R. K. Behera, P. K. Behera, B. K. Mishra and G. B. Behera (2000). "Cyanines during the 1990s: a review." *Chem. Rev.* **100**(6): 1973-2012.

Miura, Y., N. Misawa, Y. Kawano, H. Okada, Y. Inagaki, N. Yamamoto, M. Ito, H. Yagita, K. Okumura, H. Mizusawa and Y. Koyanagi (2003). "Tumor necrosis factor-related apoptosis-inducing ligand induces neuronal death in a murine model of HIV central nervous system infection." *Proc. Natl. Acad. Sci. U. S. A.* **100**(5): 2777-2782.

Morellet, N., H. de Rocquigny, Y. Mély, N. Jullian, H. Déméné, M. Ottmann, D. Gérard, J. L. Darlix, M. C. Fournie-Zaluski and B. P. Roques (1994). "Conformational behaviour of the active and inactive forms of the nucleocapsid NCp7 of HIV-1 studied by 1H NMR." *J. Mol. Biol.* **235**(1): 287-301.

Morellet, N., H. Demene, V. Teilleux, T. Huynh-Dinh, H. de Rocquigny, M. C. Fournie-Zaluski and B. P. Roques (1998). "Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC)." *J. Mol. Biol.* **283**(2): 419-434.

Morellet, N., N. Jullian, H. de Rocquigny, B. Maigret, J. L. Darlix and B. P. Roques (1992). "Determination of the structure of the nucleocapsid protein NCp7 from the human immunodeficiency virus type 1 by 1H NMR." *EMBO J.* **11**(8): 3059-3065.

Moreno, F., M. Cortijo and J. Gonalez-Jimanez (1999). "The fluorescent probe Prodan characterizes the warfarin binding site on human serum albumin." *Photochem. Photobiol.* **69**(1): 8-15.

Mori, M., U. Dietrich, F. Manetti and M. Botta (2010). "Molecular dynamics and DFT study on HIV-1 nucleocapsid protein-7 in complex with viral genome." *J. Chem. Inf. Model*. **50**(4):638-650.

Morris, M. C., P. Vidal, L. Chaloin, F. Heitz and G. Divita (1997). "A new peptide vector for efficient delivery of oligonucleotides into mammalian cells." *Nucleic Acids Res.* **25**(14): 2730-2736.

Muir, T. W. (2003). "Semisynthesis of proteins by expressed protein ligation." *Annu. Rev. Biochem.* **72**: 249-289.

Muir, T. W., D. Sondhi and P. A. Cole (1998). "Expressed protein ligation: a general method for protein engineering." *Proc. Natl. Acad. Sci. U. S. A.* **95**(12): 6705-6710.

Muller, B., U. Tessmer, U. Schubert and H. G. Krausslich (2000). "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.* **74**(20): 9727-9731.

Muriaux, D., J. L. Darlix and A. Cimarelli (2004). "Targeting the assembly of the human immunodeficiency virus type I." *Curr. Pharm. Design* **10**(30): 3725-3739.

Muthumani, K., A. Y. Choo, A. Premkumar, D. S. Hwang, K. P. Thieu, B. M. Desai and D. B. Weiner (2005). "Human immunodeficiency virus type 1 (HIV-1) Vpr-regulated cell death: insights into mechanism." *Cell Death Differ.* **12**: 962-970.

Myhra, S. (2004). "A review of enabling technologies based on scanning probe microscopy relevant to bioanalysis." *Biosens. Bioelectron.* **19**(11): 1345-1354.

Naito, A., T. Nagao, K. Norisada, T. Mizuno, S. Tuzi and H. Saito (2000). "Conformation and dynamics of melittin bound to magnetically oriented lipid bilayers by solid-state P-31 and C-13 NMR spectroscopy." *Biophys. J.* **78**(5): 2405-2417.

Nakanishi, J., M. Maeda and Y. Umezawa (2004). "A new protein conformation indicator based on biarsenical fluorescein with an extended benzoic acid moiety." *Anal. Sci.* **20**(2): 273-278.

Neef, A. B. and C. Schultz (2009). "Selective fluorescence labeling of lipids in living cells." *Angew. Chem. Int. Ed. Engl.* **48**(8): 1498-1500.

Nemkovich, N. A., V. G. Pivovarenko, W. Baumann, A. N. Rubinov and A. N. Sobchuk (2005). "Dipole moments of 4'-aminoflavonol fluorescent probes in different solvents." *J. Fluoresc.* **15**(1): 29-36.

Netzel, T. L. (2007). "The spectroscopy, dynamics, and electronic structure of pyrenyl-dU nucleosides: P center dot+/dU(center dot-) charge transfer state photophysics." *Tetrahedron* **63**(17): 3491-3514.

Nguyen, D. H. and J. E. K. Hildreth (2000). "Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts." *J. Virol.* **74**(7): 3264-3272.

Niemz, A. and D. A. Tirrell (2001). "Self-association and membrane-binding behavior of melittins containing trifluoroleucine." *J. Am. Chem. Soc.* **123**(30): 7407-7413.

Nitz, M., A. R. Mezo, M. H. Ali and B. Imperiali (2002). "Enantioselective synthesis and application of the highly fluorescent and environment-sensitive amino acid 6-(2-dimethylaminonaphthoyl) alanine (DANA)." *Chem. Commun. (Camb.)*(17): 1912-1913.

Nomanbhoy, T., A. J. Morales, A. T. Abraham, C. S. Vortler, R. Giege and P. Schimmel (2001). "Simultaneous binding of two proteins to opposite sides of a single transfer RNA." *Nat. Struct. Biol.* **8**(4): 344-348.

Nydegger, S., M. Foti, A. Derdowski, P. Spearman and M. Thali (2003). "HIV-1 egress is gated through late endosomal membranes." *Traffic* **4**(12): 902-910.

Nydegger, S., S. Khurana, D. N. Krementsov, M. Foti and M. Thali (2006). "Mapping of tetraspanin-enriched microdomains that can function as gateways for HIV-1." *J. Cell Biol.* **173**(5): 795-807.

Oberleithner, H., E. Brinckmann, A. Schwab and G. Krohne (1994). "Imaging Nuclear Pores Of Aldosterone-Sensitive Kidney Cells By Atomic Force Microscopy." *Proc. Natl. Acad. Sci. U. S. A.* **91**(21): 9784-9788.

Oehler, S., R. Alex and A. Barker (1999). "Is nitrocellulose filter binding really a universal assay for protein-DNA interactions?" *Anal. Biochem.* **268**(2): 330-336.

Okamoto, A., K. Tainaka, K. Nishiza and I. Saito (2005). "Monitoring DNA structures by dual fluorescence of pyrene derivatives." *J. Am. Chem. Soc.* **127**(38): 13128-13129.

Okamoto, A., K. Tainaka and I. Saito (2003). "Clear distinction of purine bases on the complementary strand by a fluorescence change of a novel fluorescent nucleoside." *J. Am. Chem. Soc.* **125**(17): 4972-4973.

Okamoto, A., K. Tainaka and I. Saito (2003). "Synthesis and properties of a novel fluorescent nucleobase, naphthopyridopyrimidine." *Tetrahedron Lett.* **44**(36): 6871-6874.

Olasagasti, F., K. R. Lieberman, S. Benner, G. M. Cherf, J. M. Dahl, D. W. Deamer and M. Akeson (2010). "Replication of individual DNA molecules under electronic control using a protein nanopore." *Nat. Nanotechnol.* **5**(11): 798-806.

Ono, A. and E. O. Freed (2004). "Cell-type-dependent targeting of human immunodeficiency virus type 1 assembly to the plasma membrane and the multivesicular body." *J. Virol.* **78**(3): 1552-1563.

Op den Kamp, J. A. F. (1979). "Lipid Asymmetry In Membranes." *Annu. Rev. Biochem.* **48**: 47-71.

Oreilly, M. M., M. T. McNally and K. L. Beemon (1995). "2 Strong 5'-splice sites and competing, suboptimal 3'-splice sites involved in alternative splicing of human immunodeficiency virus type-1 RNA." *Virology* **213**(2): 373-385.

Ormson, S. and R. Brown (1994). "Excited-State Intramolecular Proton-Transfer. ESIPT to nitrogen." *Progress In Reaction Kinetics* **19**(1): 45-91.

Ormson, S. and R. Brown (1994). "Excited-State Intramolecular Proton-Transfer. ESIPT to oxygen." *Progress In Reaction Kinetics* **19**(3): 211-275.

Ormson, S. M., R. G. Brown, F. Vollmer and W. Rettig (1994). "Switching between charge- and proton-transfer emission in the excited state of a substituted 3-hydroxyflavone." *J. Photochem. Photobiol. A: Chem.* **81**: 65-72.

Ott, D. E. (2002). "Potential roles of cellular proteins in HIV-1." *Rev. Med. Virol.* **12**(6): 359-374.

Oubridge, C., N. Ito, P. R. Evans, C. H. Teo and K. Nagai (1994). "Crystal structure at 1.92 angstrom resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin." *Nature* **372**(6505): 432-438.

Pabo, C. O. and L. Nekludova (2000). "Geometric analysis and comparison of protein-DNA interfaces: Why is there no simple code for recognition?" *J. Mol. Biol.* **301**(3): 597-624.

Pabo, C. O. and R. T. Sauer (1992). "Transcription factors - structural families and principles of DNA recognition." *Ann. Rev. Biochem.* **61**: 1053-1095.

Paillart, J. C., M. Shehu-Xhilaga, R. Marquet and J. Mak (2004). "Dimerization of retroviral RNA genomes: An inseparable pair." *Nat. Rev. Microbiol.* **2**(6): 461-472.

Pantano, S. and P. Carloni (2005). "Comparative analysis of HIV-1 Tat variants." *Proteins* **58**(3): 638-643.

Parasassi, T., M. Di Stefano, M. Loiero, G. Ravagnan and E. Gratton (1994). "Cholesterol modifies water concentration and dynamics in phospholipid bilayers: a fluorescence study using Laurdan probe." *Biophys J* **66**(3): 763-768.

Parsons, J. and T. Hermann (2007). "Conformational flexibility of ribosomal decoding-site RNA monitored by fluorescent pteridine base analogues." *Tetrahedron* **63**(17): 3548-3552.

Pata, J. D., W. G. Stirtan, S. W. Goldstein and T. A. Steitz (2004). "Structure of HIV-1 reverse transcriptase bound to an inhibitor active against mutant reverse transcriptases resistant to other nonnucleoside inhibitors." *Proc. Natl. Acad. Sci. U. S. A.* **101**(29): 10548-10553.

Patel, L. N., J. L. Zaro and W. C. Shen (2007). "Cell penetrating peptides: Intracellular pathways and pharmaceutical perspectives." *Pharm. Res.* **24**(11): 1977-1992.

Paxton, W., R. I. Connor and N. R. Landau (1993). "Incorporation of Vpr into human immunodeficiency virus type-1 virions - requirement for the P6 region of Gag and mutational analysis." *J. Virol.* **67**(12): 7229-7237.

Pelchen-Matthews, A., B. Kramer and M. Marsh (2003). "Infectious HIV-1 assembles in late endosomes in primary macrophages." *J. Cell Biol.* **162**(3): 443-455.

Peliska, J. A. and S. J. Benkovic (1992). "Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase." *Science* **258**(5085): 1112-1118.

Peloponese, J. M., C. Gregoire, S. Opi, D. Esquieu, J. Sturgis, E. Lebrun, E. Meurs, Y. Collette, D. Olive, A. M. Aubertin, M. Witvrow, C. Pannecouque, E. De Clercq, C. Bailly, J. Lebreton and E. P. Loret (2000). "^{1}H-^{13}C nuclear magnetic resonance assignment and structural characterization of HIV-1 Tat protein." *Comptes Rendus Acad. Sci. Ser. III-Sci. Vie-Life Sci.* **323**(10): 883-894.

Peng, J. M., Y. R. Zhu, J. T. Milton and D. H. Price (1998). "Identification of multiple cyclin subunits of human P-TEFb." *Genes Dev.* **12**(5): 755-762.

Petit, C., O. Schwartz and F. Mammano (2000). "The karyophilic properties of human immunodeficiency virus type 1 integrase are not required for nuclear import of proviral DNA." *J. Virol.* **74**(15): 7119-7126.

Pettit, S. C., J. Simsic, D. D. Loeb, L. Everitt, C. A. Hutchison and R. Swanstrom (1991). "Analysis of retroviral protease cleavage sites reveals 2 types of cleavage sites and the structural requirements of the P1 amino-acid." *J. Biol. Chem.* **266**(22): 14539-14547.

Piller, S. C., L. Caly and D. A. Jans (2003). "Nuclear import of the pre-integration complex (PIC): the Achilles heel of HIV?" *Curr. Drug Targets* **4**(5): 409-429.

Piller, S. C., G. D. Ewart, A. Premkumar, G. B. Cox and P. W. Gage (1996). "Vpr protein of human immunodeficiency virus type 1 forms cation-selective channels in planar lipid bilayers." *Proc. Natl. Acad. Sci. U. S. A.* **93**(1): 111-115.

Planelles, V., J. B. M. Jowett, Q. X. Li, Y. M. Xie, B. Hahn and I. S. Y. Chen (1996). "Vpr-induced cell cycle arrest is conserved among primate lentiviruses." *J. Virol.* **70**(4): 2516-2524.

Poljak, L., S. M. Batson, D. Ficheux, B. P. Roques, J. L. Darlix and E. Kas (2003). "Analysis of NCp7-dependent activation of HIV-1 cDNA integration and its conservation among retroviral nucleocapsid proteins." *J. Mol. Biol.* **329**(3): 411-421.

Polyakov, V., V. Sharma, J. L. Dahlheimer, C. M. Pica, G. D. Luker and D. Piwnica-Worms (2000). "Novel Tat-peptide chelates for direct transduction of technetium-99m and rhenium into human cells for imaging and radiotherapy." *Bioconjugate Chem.* **11**(6): 762-771.

Popot, J. L. and D. M. Engelman (2000). "Helical membrane protein folding, stability, and evolution." *Annu. Rev. Biochem.* **69**: 881-922.

Porcelli, F., B. A. Buck-Koehntop, S. Thennarasu, A. Ramamoorthy and G. Veglia (2006). "Structures of the dimeric and monomeric variants of magainin antimicrobial peptides (MSI-78 and MSI-594) in micelles and bilayers, determined by NMR spectroscopy." *Biochemistry* **45**(18): 5793-5799.

Pornillos, O., B. K. Ganser-Pornillos, B. N. Kelly, Y. Z. Hua, F. G. Whitby, C. D. Stout, W. I. Sundquist, C. P. Hill and M. Yeager (2009). "X-ray structures of the hexameric building block of the HIV capsid." *Cell* **137**(7): 1282-1292.

Postupalenko, V. Y., V. V. Shvadchak, G. Duportail, V. G. Pivovarenko, A. S. Klymchenko and Y. Mely (2011). "Monitoring membrane binding and insertion of peptides by two-color fluorescent label." *Biochim. Biophys. Acta-Biomembr.* **1808**(1): 424-432.

Prabakaran, P., A. S. Dimitrov, T. R. Fouts and D. S. Dimitrov (2007). "Structure and function of the HIV envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors." *Adv. Pharmacol.* **55**: 33-97.

Prats, A. C., L. Sarih, C. Gabus, S. Litvak, G. Keith and J. L. Darlix (1988). "Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and positions the replication primer tRNA onto genomic RNA." *Embo J.* **7**(6): 1777-1783.

Prochiantz, A. (1996). "Getting hydrophilic compounds into cells: lessons from homeopeptides - Commentary." *Curr. Opin. Neurobiol.* **6**(5): 629-634.

Ptacin, J. L., M. Nollmann, E. C. Becker, N. R. Cozzarelli, K. Pogliano and C. Bustamante (2008). "Sequence-directed DNA export guides chromosome translocation during sporulation in Bacillus subtilis." *Nat. Struct. Mol. Biol.* **15**(5): 485-493.

Purcell, D. F. J. and M. A. Martin (1993). "Alternative splicing of human immunodeficiency virus type-1 messenger-RNA modulates viral protein expression, replication, and infectivity." *J. Virol.* **67**(11): 6365-6378.

Qian, S., W. C. Wang, L. Yang and H. W. Huang (2008). "Structure of the alamethicin pore reconstructed by x-ray diffraction analysis." *Biophys. J.* **94**(9): 3512-3522.

Qiang, Z. and J. H. N. Yik (2006). "The Yin and Yang of P-TEFb regulation: Implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation." *Microbiol. Mol. Biol. Rev.* **70**(3): 646-659.

Qin, P. Z. and A. M. Pyle (1998). "The architectural organization and mechanistic function of group II intron structural elements." *Curr. Opin. Struct. Biol.* **8**(3): 301-308.

Raghuraman, H. and A. Chattopadhyay (2007). "Melittin: a membrane-active peptide with diverse functions." *Bioscience Rep.* **27**(4-5): 189-223.

Rajbhandary, U. L., S. H. Chang, A. Stuart, R. D. Faulkner, R. M. Hoskinson and H. G. Khorana (1967). "Studies on polynucleotides, lxviii the primary structure of yeast phenylalanine transfer RNA." *Proc. Natl. Acad. Sci. U.S.A.* **57**(3): 751-8.

Ramalanjaona, N., H. de Rocquigny, A. Millet, D. Ficheux, J. L. Darlix and Y. Mely (2007). "Investigating the mechanism of the nucleocapsid protein chaperoning of the second strand transfer during HIV-1 DNA synthesis." *J. Mol. Biol.* **374**(4): 1041-1053.

Reddy, C. K., A. Das and B. Jayaram (2001). "Do water molecules mediate protein-DNA recognition?" *J. Mol. Biol.* **314**(3): 619-632.

Reicin, A. S., G. Kalpana, S. Paik, S. Marmon and S. Goff (1995). "Sequences in the human immunodeficiency virus type-1 U3 region required for in-vivo and in-vitro integration." *J. Virol.* **69**(9): 5904-5907.

Rein, A., L. E. Henderson and J. G. Levin (1998). "Nucleic-acid chaperone activity of retroviral nucleocapsid proteins: significance for viral replication." *Trends Biochem. Sci.* **23**(8): 297-301.

Ren, R. X. F., N. C. Chaudhuri, P. L. Paris, S. Rumney and E. T. Kool (1996). "Naphthalene, phenanthrene, and pyrene as DNA base analogues: Synthesis, structure, and fluorescence in DNA." *J. Am. Chem. Soc.* **118**(33): 7671-7678.

Rice, P. A., Correll, C. C. (2008). Protein-Nucleic Acid Interactions: Structural Biology, Royal Society of Chemistry.

Richard, J. P., K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. J. Gait, L. V. Chernomordik and B. Lebleu (2003). "Cell-penetrating peptides - a reevaluation of the mechanism of cellular uptake." *J. Biol. Chem.* **278**(1): 585-590.

Richter, R. P., R. Berat and A. R. Brisson (2006). "Formation of solid-supported lipid bilayers: an integrated view." *Langmuir* **22**(8): 3497-3505.

Richter, R. P. and A. R. Brisson (2005). "Following the formation of supported lipid bilayers on mica: A study combining AFM, QCM-D, and ellipsometry." *Biophys. J.* **88**(5): 3422-3433.

Rist, M. J. and J. P. Marino (2002). "Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions." *Curr. Org. Chem.* **6**(9): 775-793.

Riviere, L., J. L. Darlix and A. Cimarelli (2010). "Analysis of the viral elements required in the nuclear import of HIV-1 DNA." *J. Virol.* **84**(2): 729-739.

Rohs, R., X. S. Jin, S. M. West, R. Joshi, B. Honig and R. S. Mann (2010). Origins of specificity in protein-DNA recognition. Annual Review Of Biochemistry. Palo Alto, Annual Reviews. **79:** 233-269.

Rohs, R., S. M. West, A. Sosinsky, P. Liu, R. S. Mann and B. Honig (2009). "The role of DNA shape in protein-DNA recognition." *Nature* **461**(7268): 1248-1281.

Romani, B., S. Engelbrecht and R. H. Glashoff (2010). "Functions of Tat: the versatile protein of human immunodeficiency virus type 1." *J. Gen. Virol.* **91**: 1-12.

Roux, K. H. and K. A. Taylor (2007). "AIDS virus envelope spike structure." *Curr. Opin. Struct. Biol.* **17**(2): 244-252.

Russell, R. S., C. Liang and M. A. Wainberg (2004). "Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably?" *Retrovirology* **1**: 23.

Ruwanpura, S. M., R. I. McLachlan, P. G. Stanton and S. J. Meachem (2008). "Follicle-stimulating hormone affects spermatogonial survival by regulating the intrinsic apoptotic pathway in adult rats." *Biol. Reprod.* **78**(4): 705-713.

Ryu, J. H., Y. J. Seo, G. T. Hwang, J. Y. Lee and B. H. Kim (2007). "Triad base pairs containing fluorene unit for quencher-free SNP typing." *Tetrahedron* **63**(17): 3538-3547.

Sadhu, K. K., S. Mizukami, Y. Hori and K. Kikuchi (2011). "Switching modulation for protein labeling with activatable fluorescent probes." *ChemBioChem* **12**(9): 1299-1308.

Saito, Y., K. Hanawa, K. Motegi, K. Omoto, A. Okamoto and I. Saito (2005). "Synthesis and properties of purine-type base-discriminating fluorescent (BDF) nucleosides: distinction of thymine by fluorescence-labeled deoxyadenosine derivatives." *Tetrahedron Lett.* **46**(44): 7605-7608.

Saito, Y., Y. Miyauchi, A. Okamoto and I. Saito (2004). "Base-discriminating fluorescent (BDF) nucleoside: distinction of thymine by fluorescence quenching." *Chem. Commun.* (15): 1704-1705.

Sakakibara, D., A. Sasaki, T. Ikeya, J. Hamatsu, T. Hanashima, M. Mishima, M. Yoshimasu, N. Hayashi, T. Mikawa, M. Walchli, B. O. Smith, M. Shirakawa, P. Guntert and Y. Ito (2009). "Protein structure determination in living cells by in-cell NMR spectroscopy." *Nature* **458**(7234): 102-110.

Salditt, T. (2003). "Lipid-peptide interaction in oriented bilayers probed by interface-sensitive scattering methods." *Curr. Opin. Struct. Biol.* **13**(4): 467-478.

Salnikov, E. S., M. De Zotti, F. Formaggio, X. Li, C. Toniolo, J. D. J. O'Neil, J. Raap, S. A. Dzuba and B. Bechinger (2009). "Alamethicin topology in phospholipid membranes by oriented solid-state NMR and EPR spectroscopies: a comparison." *J. Phys. Chem. B* **113**(10): 3034-3042.

Samanta, A. and R. W. Fessenden (2000). "Excited state dipole moment of PRODAN as determined from transient dielectric loss measurements." *J. Phys. Chem. A* **104**: 8972-8975.

Sanderson, J. M. (2005). "Peptide lipid interactions: insights and perspectives." *Org. Biomol. Chem.* **3**(2): 201-212.

Sandin, P., G. Stengel, T. Ljungdahl, K. Borjesson, B. Macao and L. M. Wilhelmsson (2009). "Highly efficient incorporation of the fluorescent nucleotide analogs tC and tC(O) by Klenow fragment." *Nucleic Acids Res.* **37**(12): 3924-3933.

Sarai, A. and H. Kono (2005). "Protein-DNA recognition patterns and predictions." *Annu. Rev. Biophys. Biomolec. Struct.* **34**: 379-398.

Schevitz, R. W., Z. Otwinowski, A. Joachimiak, C. L. Lawson and P. B. Sigler (1985). "The 3-dimensional structure of Trp repressor." *Nature* **317**(6040): 782-786.

Schnolzer, M. and S. B. Kent (1992). "Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease." *Science* **256**(5054): 221-225.

Schuler, W., K. Wecker, H. de Rocquigny, Y. Baudat, J. Sire and B. P. Roques (1999). "NMR structure of the (52-96) C-terminal domain of the HIV-1 regulatory protein Vpr: molecular insights into its biological functions." *J. Mol. Biol.* **285**(5): 2105-2117.

Schwartz, O., V. Marechal, S. Le Gall, F. Lemonnier and J. M. Heard (1996). "Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein." *Nat. Med.* **2**(3): 338-342.

Schwartz, S., B. K. Felber, E. M. Fenyo and G. N. Pavlakis (1990). "Env and Vpu proteins of human immunodeficiency virus type-1 are produced from multiple bicistronic messenger-RNAs." *J. Virol.* **64**(11): 5448-5456.

Scopes, D. I. C., J. R. Barrio and N. J. Leonard (1977). "Defined dimensional changes in enzyme cofactors - fluorescent stretched-out analogs of adenine-nucleotides." *Science* **195**(4275): 296-298.

Secrist, J. A., 3rd, J. R. Barrio and N. J. Leonard (1972). "A fluorescent modification of adenosine triphosphate with activity in enzyme systems: 1,N 6 -ethenoadenosine triphosphate." *Science* **175**(22): 646-647.

Secrist, J. A., 3rd, J. R. Barrio, N. J. Leonard and G. Weber (1972). "Fluorescent modification of adenosine-containing coenzymes. Biological activities and spectroscopic properties." *Biochemistry* **11**(19): 3499-3506.

Seelig, J. (2004). "Thermodynamics of lipid-peptide interactions." *Biochim. Biophys. Acta-Biomembr.* **1666**(1-2): 40-50.

Selig, L., J. C. Pages, V. Tanchou, S. Preveral, C. Berlioz-Torrent, L. X. Liu, L. Erdtmann, J. Darlix, R. Benarous and S. Benichou (1999). "Interaction with the p6 domain of the gag precursor mediates incorporation into virions of Vpr and Vpx proteins from primate lentiviruses." *J Virol* **73**(1): 592-600.

Sengupta, P. K. and M. Kasha (1979). "Excited state proton-transfer spectroscopy of 3-hytdroxyflavone and quercetin." *Chem. Phys. Lett.* **68**(2-3): 382-385.

Serganov, A. and D. J. Patel (2007). "Ribozymes, riboswitches and beyond: regulation of gene expression without proteins." *Nat. Rev. Genet.* **8**(10): 776-790.

Serrano-Andres, L. and M. Merchan (2009). "Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective." *J. Photochem. Photobiol. C-Photochem. Rev.* **10**(1): 21-32.

Sewald, N., S. D. Wilking, R. Eckel, S. Albu, K. Wollschlager, K. Gaus, A. Becker, F. W. Bartels, R. Ros and D. Anselmetti (2006). "Probing DNA-peptide interaction forces at the single-molecule level." *J. Pep. Sci.* **12**(12): 836-842.

Shahin, V., W. Hafezi, H. Oberleithner, Y. Ludwig, B. Windoffer, H. Schillers and J. E. Kuhn (2006). "The genome of HSV-1 translocates through the nuclear pore as a condensed rod-like structure." *J. Cell Sci.* **119**(1): 23-30.

Shai, Y. (1999). "Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides." *Biochim. Biophys. Acta-Biomembr.* **1462**(1-2): 55-70.

Shai, Y. (2002). "Mode of action of membrane active antimicrobial peptides." *Biopolymers* **66**(4): 236-248.

Shcherbakova, I., S. Mitra, R. H. Beer and M. Brenowitz (2006). "Fast Fenton footprinting: a laboratory-based method for the time-resolved analysis of DNA, RNA and proteins." *Nucleic Acids Res.* **34**(6).

Shehu-Xhilaga, M., S. M. Crowe and J. Mak (2001). "Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity." *J. Virol.* **75**(4): 1834-1841.

Shehu-Xhilaga, M., M. Hill, J. A. Marshall, J. Kappes, S. M. Crowe and J. Mak (2002). "The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of Pol protein." *J. Virol.* **76**(9): 4331-4340.

Shehu-Xhilaga, M., H. G. Kraeusslich, S. Pettit, R. Swanstrom, J. Y. Lee, J. A. Marshall, S. M. Crowe and J. Mak (2001). "Proteolytic processing of the P2/nucleocapsid cleavage site is critical for human immunodeficiency virus type 1 RNA dimer maturation." *J. Virol.* **75**(19): 9156-9164.

Sheng, N. J., S. C. Pettit, R. J. Tritch, D. H. Ozturk, M. M. Rayner, R. Swanstrom and S. EricksonViitanen (1997). "Determinants of the human immunodeficiency virus type 1 p15NC-RNA interaction that affect enhanced cleavage by the viral protease." *J. Virol.* **71**(8): 5723-5732.

Sherman, M. P. and W. C. Greene (2002). "Slipping through the door: HIV entry into the nucleus." *Microbes Infect.* **4**(1): 67-73.

Shojania, S. and J. D. O'Neil (2006). "HIV-1 Tat is a natively unfolded protein - The solution conformation and dynamics of reduced HIV-1 Tat-(1-72) by NMR spectroscopy." *J. Biol. Chem.* **281**(13): 8347-8356.

Shubsda, M. F., A. C. Paoletti, B. S. Hudson and P. N. Borer (2002). "Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein." *Biochemistry* **41**(16): 5276-5282.

Shvadchak, V. V., A. S. Klymchenko, H. De Rocquigny and Y. Mely (2009). "Sensing peptide - Oligonucleotide interactions by a two-color fluorescence label: application to the HIV-1 nucleocapsid protein." *Nucleic Acids Res.* **37**(3): e25.

Shynkar, V. V., A. S. Klymchenko, G. Duportail, A. P. Demchenko and Y. Mely (2005). "Two-color fluorescent probes for imaging the dipole potential of cell plasma membranes." *Biochim. Biophys. Acta* **1712**(2): 128-136.

Shynkar, V. V., A. S. Klymchenko, C. Kunzelmann, G. Duportail, C. D. Muller, A. P. Demchenko, J. M. Freyssinet and Y. Mely (2007). "Fluorescent biomembrane probe for ratiometric detection of apoptosis." *J. Am. Chem. Soc.* **129**(7): 2187-2193.

Sierra, S., B. Kupfer and R. Kaiser (2005). "Basics of the virology of HIV-1 and its replication." *J. Clin. Virol.* **34**(4): 233-244.

Singer, S. J. and G. L. Nicolson (1972). "Fluid mosaic model of structure of cell-membranes." *Science* **175**(4023): 720-731.

Sinha, S. and D. P. Grandgenett (2005). "Recombinant human immunodeficiency virus type 1 integrase exhibits a capacity for full-site integration in vitro that is comparable to that of purified preintegration complexes from virus-infecte cells." *J. Virol.* **79**(13): 8208-8216.

Sinkeldam, R. W., N. J. Greco and Y. Tor (2010). "Fluorescent analogs of biomolecular building blocks: design, properties, and applications." *Chem. Rev.* **110**(5): 2579-2619.

Skripkin, E., J. C. Paillart, R. Marquet, B. Ehresmann and C. Ehresmann (1994). "Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro." *Proc. Natl. Acad. Sci. U. S. A.* **91**(11): 4945-4949.

Slavik, J. (1982). "Anilinonaphthalene sulfonate as a probe of membrane composition and function." *Biochim. Biophys. Acta* **694**(1): 1-25.

Sodroski, J., R. Patarca, C. Rosen, F. Wongstaal and W. Haseltine (1985). "Location of the trans-activating region on the genome of human T-cell lymphotropic virus type-III." *Science* **229**(4708): 74-77.

Soomets, U., M. Lindgren, X. Gallet, M. Hallbrink, A. Elmquist, L. Balaspiri, M. Zorko, M. Pooga, R. Brasseur and U. Langel (2000). "Deletion analogues of transportan." *Biochim. Biophys. Acta-Biomembr.* **1467**(1): 165-176.

Soujanya, T., R. W. Fessenden and A. Samanta (1996). "Role of nonfluorescent twisted intramolecular charge transfer state on the photophysical behavior of aminophthalimide dyes." *J. Phys. Chem.* **100**(9): 3507-3512.

South, T. L., P. R. Blake, R. C. Sowder, L. O. Arthur, L. E. Henderson and M. F. Summers (1990). "The nucleocapsid protein isolated from HIV-1 particles binds zinc and forms retroviral-type zinc fingers." *Biochemistry* **29**(34): 7786-7789.

Spadafora, M., V. Y. Postupalenko, V. V. Shvadchak, A. S. Klymchenko, Y. Mely, A. Burger and R. Benhida (2009). "Efficient synthesis of ratiometric fluorescent nucleosides featuring 3-hydroxychromone nucleobases." *Tetrahedron* **65**(37): 7809-7816.

Srivatsan, S. G. and Y. Tor (2007). "Fluorescent pyrimidine ribonucleotide: Synthesis, enzymatic incorporation, and utilization." *J. Am. Chem. Soc.* **129**(7): 2044-2053.

Srivatsan, S. G. and Y. Tor (2009). "Enzymatic incorporation of emissive pyrimidine ribonucleotides." *Chem.-Asian J.* **4**(3): 419-427.

Srivatsan, S. G., H. Weizman and Y. Tor (2008). "A highly fluorescent nucleoside analog based on thieno[3,4-d] pyrimidine senses mismatched pairing." *Org. Biomol. Chem.* **6**(8): 1334-1338.

Strassler, C., N. E. Davis and E. T. Kool (1999). "Novel nucleoside analogues with fluorophores replacing the DNA base." *Helv. Chim. Acta* **82**(12): 2160-2171.

Strebel, K. (2007). "HIV accessory genes Vif and Vpu." *Adv Pharmacol* **55**: 199-232.

Stromstedt, A. A., L. Ringstad, A. Schmidtchen and M. Malmsten (2010). "Interaction between amphiphilic peptides and phospholipid membranes." *Curr. Opin. Colloid Interface Sci.* **15**(6): 467-478.

Summerer, D., S. Chen, N. Wu, A. Deiters, J. W. Chin and P. G. Schultz (2006). "A genetically encoded fluorescent amino acid." *Proc. Natl. Acad. Sci. U. S. A.* **103**(26): 9785-9789.

Summers, M. F., L. E. Henderson, M. R. Chance, J. W. Bess Jr, T. L. South, P. R. Blake, I. Sagi, G. Perez-Alvarado, R. C. Sowder Iii, D. R. Hare and L. O. Arthur (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**(5): 563-574.

Suzuki, M. (1994). "A framework for the DNA-protein recognition code of the probe helix in transcription factors - the chemical and stereochemical rules." *Structure* **2**(4): 317-326.

Suzuki, Y. and R. Craigie (2007). "The road to chromatin - nuclear entry of retroviruses." *Nat. Rev. Microbiol.* **5**(3): 187-196.

Swanson, C. M. and M. H. Malim (2008). "SnapShot: HIV-1 proteins." *Cell* **133**(4): 742.

Swigut, T., A. J. Iafrate, J. Muench, F. Kirchhoff and J. Skowronski (2000). "Simian and human immunodeficiency virus Nef proteins use different surfaces to downregulate class I major histocompatibility complex antigen expression." *J. Virol* .**74**(12): 5691-5701.

Swinney, T. C. and D. F. Kelley (1993). "Proton transfer dynamics in substituted 3-hydroxyflavones: Solvent polarization effects." *J. Chem. Phys.* **99**(1): 211-221.

Sykora, J., P. Jurkiewicz, R. M. Epand, R. Kraayenhof, M. Langner and M. Hof (2005). "Influence of the curvature on the water structure in the headgroup region of phospholipid bilayer studied by the solvent relaxation technique." *Chem. Phys. Lipids* **135**(2): 213-221.

Sykora, J., P. Kapusta, V. Fidler and M. Hof (2002). "On what time scale does solvent relaxation in phospholipid bilayers happen?" *Langmuir* **18**(3): 571-574.

Tafvizi, A., F. Huang, J. S. Leith, A. R. Fersht, L. A. Mirny and A. M. van Oijen (2008). "Tumor suppressor p53 slides on DNA with low friction and high stability." *Biophys. J.* **95**(1): L1-L3.

Tahirov, T. H., N. D. Babayeva, K. Varzavand, J. J. Cooper, S. C. Sedore and D. H. Price (2010). "Crystal structure of HIV-1 Tat complexed with human P-TEFb." *Nature* **465**(7299): 747-752.

Tajmir-Riahi, H. A. (2006). "An overview of protein-DNA and protein-RNA interactions." *J. Iran Chem.Soc.* **3**(4): 297-304.

Takagi, Y., M. Warashina, W. J. Stec, K. Yoshinari and K. Taira (2001). "Recent advances in the elucidation of the mechanisms of action of ribozymes." *Nucleic Acids Res.* **29**(9): 1815-1834.

Takayama, Y., D. Sahu and J. Iwahara (2010). "NMR studies of translocation of the Zif268 protein between its target DNA sites." *Biochemistry* **49**(37): 7998-8005.

Tamm, L. K., A. L. Lai and Y. Li (2007). "Combined NMR and EPR spectroscopy to determine structures of viral fusion domains in membranes." *Biochim. Biophys. Acta-Biomembr.* **1768**(12): 3052-3060.

Tanchou, V., D. Decimo, C. Pechoux, D. Lener, V. Rogemond, L. Berthoux, M. Ottmann and J. L. Darlix (1998). "Role of the N-terminal zinc finger of human immunodeficiency virus type 1 nucleocapsid protein in virus structure and replication." *J. Virol.* **72**(5): 4442-4447.

Tanchou, V., C. Gabus, V. Rogemond and J. L. Darlix (1995). "Formation of stable and functional HIV-1 nucleoprotein complexes in-vitro." *J. Mol. Biol.* **252**(5): 563-571.

Terwilliger, T. C. and D. Eisenberg (1982). "The structure of melittin. Interpretation of the structure." *J. Biol. Chem.* **257**(11): 6016-6022.

Terwilliger, T. C., L. Weissman and D. Eisenberg (1982). "The structure of melittin in the form-I Crystals and its implication for melittins lytic and surface activities." *Biophys. J.* **37**(1): 353-361.

Thomas, J. A., T. D. Gagliardi, W. G. Alvord, M. Lubomirski, W. J. Bosche and R. J. Gorelick (2006). "Human immunodeficiency virus type 1 nucleocapsid zinc-finger mutations cause defects in reverse transcription and integration." *Virology* **353**(1): 41-51.

Thomas, J. A. and R. J. Gorelick (2008). "Nucleocapsid protein function in early infection processes." *Virus Res.* **134**(1-2): 39-63.

Thordarson, P., B. Le Droumaguet and K. Velonia (2006). "Well-defined protein-polymer conjugates-synthesis and potential applications." *Appl. Microbiol. Biotechnol.* **73**(2): 243-254.

Tieleman, D. P., S. J. Marrink and H. J. C. Berendsen (1997). "A computer perspective of membranes: Molecular dynamics studies of lipid bilayer systems." *Biochim. Biophys. Acta – Rev. Biomembr.* **1331**(3): 235-270.

Tisne, C., B. P. Roques and F. Dardel (2003). "Specific recognition of primer tRNA(3)(Lys) by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension." *Biochimie* **85**(5): 557-561.

Tor, Y. (2007). "Fluorescent nucleoside analogs: synthesis, properties, and applications." *Tetrahedron* **63**(17): 3425-3426.

Tor, Y., S. Del Valle, D. Jaramillo, S. G. Srivatsan, A. Rios and H. Weizman (2007). "Designing new isomorphic fluorescent nucleobase analogues: the thieno[3,2-d]pyrimidine core." *Tetrahedron* **63**(17): 3608-3614.

Torchilin, V. P., R. Rammohan, V. Weissig and T. S. Levchenko (2001). "TAT peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors." *Proc. Natl. Acad. Sci. U. S. A.* **98**(15): 8786-8791.

Tse, W. C. and D. L. Boger (2004). "A fluorescent intercalator displacement assay for establishing DNA binding selectivity and affinity." *Accounts Chem. Res.* **37**(1): 61-69.

Tunnemann, G., R. M. Martin, S. Haupt, C. Patsch, F. Edenhofer and M. C. Cardoso (2006). "Cargo-dependent mode of uptake and bioavailability of TAT-containing proteins and peptides in living cells." *Faseb J.* **20**(11): 1775-1784.

Turner, B. G. and M. F. Summers (1999). "Structural biology of HIV." *J Mol Biol* **285**(1): 1-32.

Twine, S. M. and A. G. Szabo (2003). "Fluorescent amino acid analogs." *Methods Enzymol.* **360**: 104-127.

Tyagi, M., M. Rusnati, M. Presta and M. Giacca (2001). "Internalization of HIV-1 Tat requires cell surface heparan sulfate proteoglycans." *J. Biol. Chem.* **276**(5): 3254-3261.

Ulrich, G., R. Ziessel and A. Harriman (2008). "The chemistry of fluorescent bodipy dyes: Versatility unsurpassed." *Angew. Chem. Int. Ed. Engl.* **47**(7): 1184-1201.

Urbaneja, M. A., B. P. Kane, D. G. Johnson, R. J. Gorelick, L. E. Henderson and J. R. Casas-Finet (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**(1): 59-75.

Vaghefi, M. (2005). Nucleoside Triphosphates and Their Analogs: Chemistry, Biotechnology, and Biological Applications, CRC Press, Taylor & Francis Group: Boca Raton, FL.

Vamvouka, M., J. Cieslak, N. Van Eps, W. Hubbell and A. Gross (2008). "The structure of the lipid-embedded potassium channel voltage sensor determined by double-electron-electron resonance spectroscopy." *Protein Sci.* **17**(3): 506-517.

van den Bogaart, G., J. T. Mika, V. Krasnikov and B. Poolman (2007). "The lipid dependence of melittin action investigated by dual-color fluorescence burst analysis." *Biophys. J.* **93**(1): 154-163.

van Duyne, R., R. Easley, W. L. Wu, R. Berro, C. Pedati, Z. Klase, K. Kehn-Hall, E. K. Flynn, D. E. Symer and F. Kashanchi (2008). "Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR." *Retrovirology* **5**.

van Meer, G., D. R. Voelker and G. W. Feigenson (2008). "Membrane lipids: where they are and how they behave." *Nat. Rev. Mol. Cell Biol.* **9**(2): 112-124.

Vardabasso, C., L. Manganaro, M. Lusic, A. Marcello and M. Giacca (2008). "The histone chaperone protein nucleosome assembly protein-1 (hNAP-1) binds HIV-1 Tat and promotes viral transcription." *Retrovirology* **5**.

Vazquez, E. M., D. M. Rothman and B. Imperiali (2004). "A new environment-sensitive fluorescent amino acid for Fmoc-based solid phase peptide synthesis." *Org. Biomol. Chem.* **2**(14): 1965-1966.

Vazquez, M. E., J. B. Blanco and B. Imperiali (2005). "Photophysics and biological applications of the environment-sensitive fluorophore 6-N,N-dimethylamino-2,3-naphthalimide." *J. Am. Chem. Soc.* **127**(4): 1300-1306.

Venkatraman, P., T. T. Nguyen, M. Sainlos, O. Bilsel, S. Chitta, B. Imperiali and L. J. Stern (2007). "Fluorogenic probes for monitoring peptide binding to class II MHC proteins in living cells." *Nat. Chem. Biol.* **3**(4): 222-228.

Verma, S. and F. Eckstein (1998). "Modified oligonucleotides: synthesis and strategy for users." *Annu. Rev. Biochem.* **67**: 99-134.

Vives, E., P. Brodin and B. Lebleu (1997). "A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus." *J. Biol. Chem.* **272**(25): 16010-16017.

Vives, E., J. P. Richard, C. Rispal and B. Lebleu (2003). "TAT peptide internalization: Seeking the mechanism of entry." *Curr. Protein Pept. Sci.* **4**(2): 125-132.

Vogel, B. E., S. J. Lee, A. Hildebrand, W. Craig, M. D. Pierschbacher, F. Wongstaal and E. Ruoslahti (1993). "A novel integrin specificity exemplified by binding of the alpha(V)beta(5) integrin to the basic domain of the HIV Tat protein and vitronectin." *J. Cell Biol.* **121**(2): 461-468.

von Heijne, G. (2006). "Membrane-protein topology." *Nat. Rev. Mol. Cell Biol.* **7**(12): 909-918.

von Hippel, P. H. (2007). "From "simple" DNA-protein interactions to the macromolecular machines of gene expression." *Annu. Rev. Biophys. Biomolec. Struct.* **36**: 79-105.

von Schwedler, U., R. S. Kornbluth and D. Trono (1994). "The nuclear-localization signal of the matrix protein of human immunodeficiency virus type-1 allows the establishment of infection in macrophages and quiescent T-lymphocytes." *Proc. Natl. Acad. Sci. U. S. A.* **91**(15): 6992-6996.

Vuilleumier, C., E. Bombarda, N. Morellet, D. Gerard, B. P. Roques and Y. Mely (1999). "Nucleic acid sequence discrimination by the HIV-1 nucleocapsid protein NCp7: a fluorescence study." *Biochemistry* **38**(51): 16816-16825.

Wagenknecht, H. A. (2008). Fluorescent DNA base modifications and substitutes: Multiple fluorophore labeling and the DETEQ concept. Fluorescence Methods And Applications: Spectroscopy, Imaging, And Probes. Oxford, Blackwell Publishing. **1130:** 122-130.

Wapling, J., S. Srivastava, M. Shehu-Xhilaga et G. Tachedjian (2007). "Targeting human immunodeficiency virus type 1 assembly, maturation and budding." *Drug Target Insights* **2**: 159-182.

Ward, D. C., E. Reich and L. Stryer (1969). "Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives." *J. Biol. Chem.* **244**(5): 1228-37.

Warrilow, D. and D. Harrich (2007). "HIV-1 replication from after cell entry to the nuclear periphery." *Curr. HIV Res.* **5**(3): 293-299.

Watkins, J. D., G. R. Campbell, H. Halimi and E. P. Loret (2008). "Homonuclear (1)H NMR and circular dichroism study of the HIV-1 Tat Eli variant." *Retrovirology* **5**.

Watson, J. D. and F. H. C. Crick (1953). "Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid." *Nature* **171**(4356): 737-738.

Watts, J. M., K. K. Dang, R. J. Gorelick, C. W. Leonard, J. W. Bess, R. Swanstrom, C. L. Burch and K. M. Weeks (2009). "Architecture and secondary structure of an entire HIV-1 RNA genome." *Nature* **460**(7256): 711-787.

Wayment, J. R. and J. M. Harris (2009). "Biotin-avidin binding kinetics measured by single-molecule imaging." *Anal. Chem.* **81**(1): 336-342.

Weaver, A. J., M. D. Kemple, J. W. Brauner, R. Mendelsohn and F. G. Prendergast (1992). "Fluorescence, CD, Attenuated Total Reflectance (ATR) FTIR, and C-13 NMR characterization of the structure and dynamics of synthetic melittin and melittin analogs in lipid environments." *Biochemistry* **31**(5): 1301-1313.

Weber, G. and F. J. Farris (1979). "Synthesis and spectral properties of a hydrophobic fluorescent probe: 6-propionyl-2-(dimethylamino)naphthalene." *Biochemistry* **18**(14): 3075-3078.

Wei, P., M. E. Garber, S. M. Fang, W. H. Fischer and K. A. Jones (1998). "A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA." *Cell* **92**(4): 451-462.

Weller, A. (1956). "Innermolekularer protonenubergang im angeregten zustand." *Z. Elektrochem.* **60**: 1144-1147.

Welsch, S., O. T. Keppler, A. Habermann, I. Allespach, J. Krijnse-Locker and H. G. Krausslich (2007). "HIV-1 buds predominantly at the plasma membrane of primary human macrophages." *Plos Pathogens* **3**(3): e36.

Wender, P. A., D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard (2000). "The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters." *Proc. Natl. Acad. Sci. U. S. A.* **97**(24): 13003-13008.

Westendorp, M. O., R. Frank, C. Ochsenbauer, K. Stricker, J. Dhein, H. Walczak, K. M. Debatin and P. H. Krammer (1995). "Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat and gp120." *Nature* **375**(6531): 497-500.

Westendorp, M. O., M. Li-Weber, R. W. Frank and P. H. Krammer (1994). "Human immunodeficiency virus type 1 Tat upregulates interleukin-2 secretion in activated T cells." *J Virol* **68**(7): 4177-4185.

White, S. H. and W. C. Wimley (1998). "Hydrophobic interactions of peptides with membrane interfaces." *Biochim. Biophys. Acta-Rev. Biomembr.* **1376**(3): 339-352.

Wieprecht, T., M. Beyermann and J. Seelig (1999). "Binding of antibacterial magainin peptides to electrically neutral membranes: thermodynamics and structure." *Biochemistry* **38**(32): 10377-10387.

Wilhelmsson, L. M. (2010). "Fluorescent nucleic acid base analogues." *Q. Rev. Biophys.* **43**(2): 159-183.

Willey, R. L., F. Maldarelli, M. A. Martin and K. Strebel (1992). "Human immunodeficiency virus type-1 Vpu protein induces rapid degradation of Cd4." *J. Virol.* **66**(12): 7193-7200.

Williams, M. C. and L. J. Maher (2011). Biophysics of DNA -Protein Interactions: From Single Molecules to Biological Systems. New York, Springer.

Wilson, J. N. and E. T. Kool (2006). "Fluorescent DNA base replacements: reporters and sensors for biological systems." *Org. Biomol. Chem.* **4**(23): 4265-4274.

Wisniewski, M., M. Balakrishnan, C. Palaniappan, P. J. Fay and R. A. Bambara (2000). "The sequential mechanism of HIV reverse transcriptase RNase H." *J. Biol. Chem.* **275**(48): 37664-37671.

Wisniewski, M., Y. Chen, M. Balakrishnan, C. Palaniappan, B. P. Roques, P. J. Fay and R. A. Bambara (2002). "Substrate requirements for secondary cleavage by HIV-1 reverse transcriptase RNase H." *J. Biol. Chem.* **277**(32): 28400-28410.

Wolf, D. and S. P. Goff (2008). "Host restriction factors blocking retroviral replication." *Annu. Rev. Genet.* **42**: 143-163.

Woodbury, C. P. and P. H. von Hippel (1983). "On the determination of deoxyribonucleic-acid protein-interaction parameters using the nitrocellulose filter-binding assay." *Biochemistry* **22**(20): 4730-4737.

Wu-Baer, F., W. S. Lane and R. B. Gaynor (1995). "The cellular factor TRP-185 regulates RNA polymerase II binding to HIV-1 TAR RNA." *Embo J.* **14**(23): 5995-6009.

Wu, H. H., A. Henras, G. Chanfreau and J. Feigon (2004). "Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III." *Proc. Natl. Acad. Sci. U. S. A.* **101**(22): 8307-8312.

Wu, W. X., L. E. Henderson, T. D. Copeland, R. J. Gorelick, W. J. Bosche, A. Rein and J. G. Levin (1996). "Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract." *J. Virol.* **70**(10): 7132-7142.

Wu, Y., H. W. Huang and G. A. Olah (1990). "Method of oriented circular dichroism." *Biophys. J.* **57**(4): 797-806.

Wu, Z. B., J. Alexandratos, B. Ericksen, J. Lubkowski, R. C. Gallo and W. Y. Lu (2004). "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.* **101**(32): 11587-11592.

Xian, J., M. G. Harrington and E. H. Davidson (1996). "DNA-protein binding assays from a single sea urchin egg: a high-sensitivity capillary electrophoresis method." *Proc. Natl. Acad. Sci. U. S. A.* **93**(1): 86-90.

Xiao, Q., R. T. Ranasinghe, A. M. P. Tang and T. Brown (2007). "Naphthalenyl- and anthracenyl-ethynyl dT analogues as base discriminating fluorescent nucleosides and intramolecular energy transfer donors in oligonucleotide probes." *Tetrahedron* **63**(17): 3483-3490.

Yamaguchi, S., D. Huster, A. Waring, R. I. Lehrer, W. Kearney, B. F. Tack and M. Hong (2001). "Orientation and dynamics of an antimicrobial peptide in the lipid bilayer by solid-state NMR spectroscopy." *Biophys. J.* **81**(4): 2203-2214.

Yamashita, M. and M. Emerman (2005). "The cell cycle independence of HIV infections is not determined by known karyophilic viral elements." *PLoS Pathog* **1**(3): e18.

Yamashita, M. and M. Emerman (2006). "Retroviral infection of non-dividing cells: old and new perspectives." *Virology* **344**(1): 88-93.

Yang, L., T. A. Harroun, T. M. Weiss, L. Ding and H. W. Huang (2001). "Barrel-stave model or toroidal model? A case study on melittin pores." *Biophys. J.* **81**(3): 1475-1485.

Yang, O. O., P. T. Nguyen, S. A. Kalams, T. Dorfman, H. G. Gottlinger, S. Stewart, I. S. Y. Chen, S. Threlkeld and B. D. Walker (2002). "Nef-mediated resistance of human immunodeficiency virus type 1 to antiviral cytotoxic T lymphocytes." *J. Virol.* **76**(4): 1626-1631.

Yeaman, M. R. and N. Y. Yount (2003). "Mechanisms of antimicrobial peptide action and resistance." *Pharmacol. Rev.* **55**(1): 27-55.

Yesylevskyy, S. O., A. S. Klymchenko and A. P. Demchenko (2005). "Semi-empirical study of two-color fluorescent dyes based on 3-hydroxychromone." *J. Mol. Struct.* **755**(1-3): 229-239.

You, J. C. and C. S. McHenry (1993). "HIV nucleocapsid protein - expression in Escherichia-Coli, purification, and characterization." *J. Biol. Chem.* **268**(22): 16519-16527.

Zachowski, A. (1993). "Phospholipids In Animal Eukaryotic Membranes - Transverse Asymmetry And Movement." *Biochem. J.* **294**: 1-14.

Zasloff, M. (1987). "Magainins, a class of antimicrobial peptides from Xenopus skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor." *Proc. Natl. Acad. Sci. U. S. A.* **84**(15): 5449-5453.

Zasloff, M. (2002). "Antimicrobial peptides of multicellular organisms." *Nature* **415**(6870): 389-395.

Zemmel, R. W., A. C. Kelley, J. Karn and P. J. G. Butler (1996). "Flexible regions of RNA structure facilitate co-operative Rev assembly on the Rev-response element." *J. Mol. Biol.* **258**(5): 763-777.

Zennou, V., C. Petit, D. Guetard, U. Nerhbass, L. Montagnier and P. Charneau (2000). "HIV-1 genome nuclear import is mediated by a central DNA flap." *Cell* **101**(2): 173-185.

Zennou, V., C. Serguera, C. Sarkis, F. Colin, E. Perret, J. Mallet and P. Charneau (2001). "The HIV-1 DNA flap stimulates HIV vector-mediated cell transduction in the brain." *Nat. Biotechnol.* **19**(5): 446-450.

Zhang, J. L. and C. S. Crumpacker (2002). "Human immunodeficiency virus type 1 nucleocapsid protein nuclear localization mediates early viral mRNA expression." *J. Virol.* **76**(20): 10444-10454.

Zhang, Y. L., Z. Q. Xi, R. S. Hegde, Z. Shakked and D. M. Crothers (2004). "Predicting indirect readout effects in protein - DNA interactions." *Proc. Natl. Acad. Sci. U. S. A.* **101**(22): 8337-8341.

Zhao, L. J., S. Mukherjee and O. Narayan (1994). "Biochemical mechanism of HIV-I Vpr function - specific interaction with a cellular protein." *J. Biol. Chem.* **269**(22): 15577-15582.

Zhao, L. J., L. Wang, S. Mukherjee and O. Narayan (1994). "Biochemical mechanism of HIV-1 Vpr function. Oligomerization mediated by the N-terminal domain." *J. Biol. Chem.* **269**(51): 32131-32137.

Zhao, M., M. F. Kircher, L. Josephson and R. Weissleder (2002). "Differential conjugation of tat peptide to superparamagnetic nanoparticles and its effect on cellular uptake." *Bioconjugate Chem.* **13**(4): 840-844.

RESUME DE LA THESE EN FRANÇAIS

Fluorophores Ratiométriques pour le Marquage de Peptides et d'Oligonucléotides: Applications à la Protéine de la Nucléocapside de VIH-1

INTRODUCTION. La luminescence est un phénomène photophysique qui correspond à l'émission de la lumière par des atomes et des molécules dans un état électronique excité. Certains types de luminescence moléculaire peuvent être considérés comme dépendants de la nature de l'état excité. Si l'émission a lieu à partir d'un état excité singulet, le processus s'appelle fluorescence. La fluorescence a lieu avec un nombre limité de molécules (généralement des hydrocarbures polyaromatique ou hétérocycliques) dites fluorophores, certaines étant utilisées comme sondes fluorescentes. La spectroscopie de fluorescence est un puissant outil d'investigation des interactions peptide-peptide, peptide-acide nucléique et peptide-membrane et des changements conformationnels qui y sont associés. Les sondes fluorescentes sensibles à l'environnement permettent de suivre ces modifications au voisinage d'un site spécifiquement marqué. La plupart des sondes utilisées comme marqueurs solvatochromiques de peptides sondent la polarité de l'environnement *via* des déplacements de leur maximum d'émission. Des exemples typiques de ces sondes sont le Prodan, son benzo-analogue (Anthradan) et des dérivés phtalimide et naphthymide: 4-DMAP, 6-DMN, 4-DMN (Figure 1).

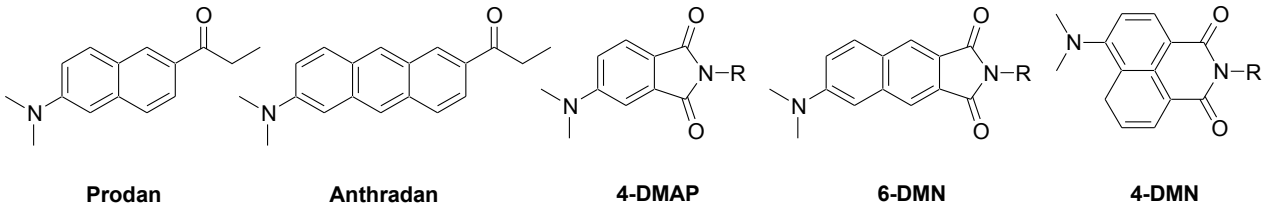


Figure 1. Le Prodan, son benzo-analogue Anthradan, les dérivés phtalimide et naphthymide.

Parmi ces sondes, les sondes ratiométriques à deux bandes d'émission présentent un intérêt particulier car leur réponse aux propriétés de l'environnement dans des systèmes hétérogènes est concentration-indépendante. Plus particulièrement, les sondes dérivées de la 3-hydroxyflavone (3HF) ont été utilisées avec succès tant pour marquer des protéines et des membranes cellulaires que pour suivre des interactions protéines-protéines et des interactions ADN-polycation. Etant donné leur réaction de transfert de proton intramoléculaire à l'état excité (ESIPT), ces fluorophores présentent deux formes à l'état excité, la forme initialement excitée dite normale (N*) et la forme tautomère (T*). Comme la réaction ESIPT de ces sondes est très sensible à la polarité du milieu, aux liaisons hydrogène et aux divers champs électriques, le rapport des deux bandes d'émission peut être utilisé pour suivre avec une grande sensibilité les modifications de leur environnement (Figure 2). Un avantage décisif de ces sondes ratiométriques sur les sondes plus conventionnelles (intensiométriques) réside dans le fait que leur rapport d'intensités ne dépend que de ces modifications d'environnement et non de leur concentration locale ou de l'instrument utilisé.

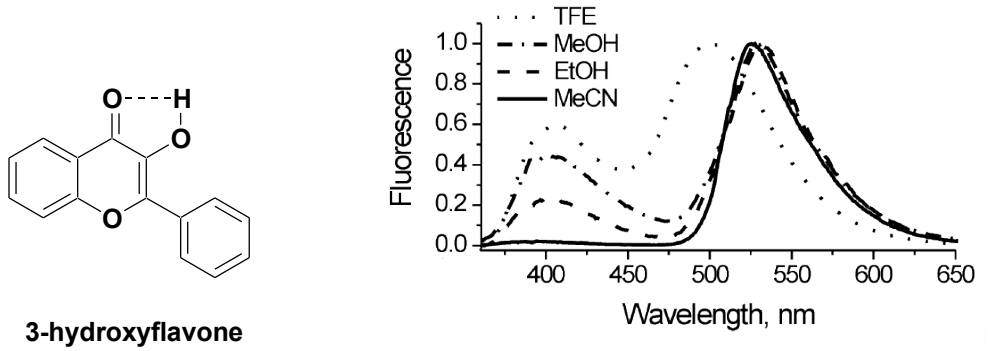


Figure 2. Structure de la 3-hydroxyflavone (3HF) et spectres de fluorescence de 3HF dans les solvants protiques et aprotiques.

La protéine de la nucléocapside de VIH-1 (NCp7) est une petite protéine basique (55 acides aminés) possédant deux motifs à doigt de zinc hautement conservés (Figure 3). Cette protéine joue un rôle important dans le cycle viral de VIH-1 et constitue de ce fait une cible prometteuse dans un but thérapeutique. NCp7 fixe avec une forte affinité deux ions zinc, qui sont nécessaires au repliement de la protéine. Une grande partie des fonctions de NCp7 est liée à ses activités chaperonnes des acides nucléiques qui sont importantes pour la transcription réverse, l'intégration et l'assemblage de la particule virale. En outre, il n'est pas exclu que NCp7 puisse aussi fixer plusieurs protéines du virus et de l'hôte. Les membranes lipidiques pourraient également être une cible additionnelle de fixation pour NCp7, même si cela n'a jamais été clairement démontré.

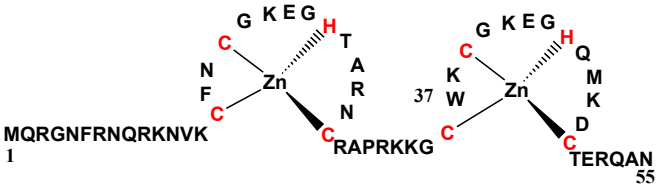


Figure 3. Structure de la NC du VIH-1, les acides aminés chargés positivement sont en bleu et les acides aminés hydrophobes du plateau sont en rouge.

ACTUALITE. Dans ce contexte, les objectifs de cette thèse furent de développer une méthodologie pour la détection des interactions des protéines avec des oligonucléotides et des membranes à base des sondes sensibles à l'environnement de la famille 3-hydroxychromone. À cette fin, la synthèse de nouvelles sondes fluorescents 3HC présentant des propriétés améliorées était nécessaire. Aussi, afin de caractériser le rôle des domaines de la protéine lors de son interaction avec les ODNs, des analogues d'acides aminés fluorescents et de nucléosides ont du être synthétisés et fixés à différentes positions du peptide et de l'ODN, respectivement. En outre, un objectif supplémentaire fût de concevoir de nouvelles sondes ratiométriques pour étudier de manière site-sélective les interactions de NCp7 avec des membranes lipidiques, afin de confirmer ces dernières comme cibles potentielles pour NCp7 et apporter de nouvelles clefs pour mieux comprendre le rôle pléiotropique de cette protéine dans le cycle viral.

RESULTATS. Récemment, une sonde 3-hydroxychromone (3HC) présentant une forte sensibilité aux solvants polaires a été développée au laboratoire et couplée à la partie N-terminale de NCp7 par synthèse peptidique en phase solide (Shvadchak, Klymchenko et al. 2009). Au contraire de la plupart des sondes utilisées à ce jour, dont la fluorescence est fortement inhibée lors de leur liaison à un oligonucléotide, aucune inhibition n'est constatée lors de la liaison de la NCp7 marquée à la 3-HC avec des oligonucléotides (ODNs). En outre, le fait que le rapport d'intensité des deux bandes d'émission de 3HC-NCp7 dépende fortement de la séquence de l'ODN étudié apporte un intérêt supplémentaire. Cette propriété a pu être corrélée sans ambiguïté à la proximité de l'ODN avec la partie N-terminale de NCp7 et la possibilité d'un *stacking* du résidu 3HC avec les bases de l'ODN, permettant ainsi d'envisager une caractérisation site-spécifique de l'interaction.

Dans le présent travail, un analogue d'acide aminé fluorescent dérivé de la sonde 3HC (Figure 4) a été synthétisé et inséré en différentes positions ciblées du peptide NCp7 (à la place des résidus Phe16, Ala30 et Trp37) par synthèse peptidique en phase solide (Figure 4). Cet analogue a une taille inférieure à deux fois celle du tryptophane.

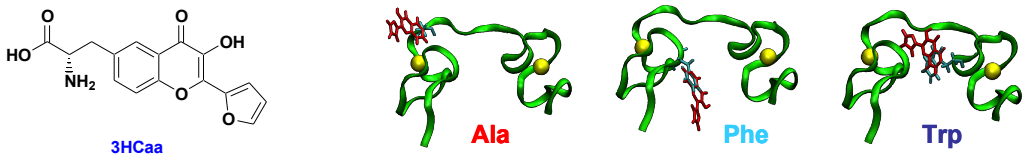


Figure 4. Structures de l'acide aminé fluorescent et 3D de NCp7 avec les résidus Ala30, Phe16 et Trp37 substitués par cet analogue d'acide aminé. Les ions Zn sont représentés par des boules jaunes. Ces structures sont obtenues à partir de données RMN.

Nous avons montré que les peptides portant cet acide aminé fluorescent conservent leur activité et peuvent ainsi être utilisés pour suivre les interactions avec les acides nucléiques. La réponse ratiométrique de fluorescence obtenue avec différentes cibles ADN ou ARN se corrèle bien avec la structure 3D des complexes NC/ODN, indiquant ainsi que les modifications des rapports d'intensité peuvent être utilisés pour caractériser de manière site-sélective les sites de fixation et déterminer les paramètres de liaison du peptide avec les ODNs. A titre d'exemple, la Figure 5 montre que, dans le complexe de la NC avec l'ARN SL3, les trois positions substituées par l'acide aminé fluorescent ne sont pas équivalentes. Le rapport N*/T* élevé, observé pour le peptide NC(11-55)-A30-3HCaa, suggère une localisation très exposée de 3HC dans le complexe, en accord avec l'absence d'interaction directe avec l'ODN. Au contraire, les valeurs beaucoup plus faibles observées avec NC(11-55)-W37-3HCaa et NC(11-55)-F16-3HCaa suggèrent un *stacking* efficace du fluorophore 3HC avec l'ODN. A notre connaissance, c'est la première fois qu'un analogue d'acide aminé présentant une fluorescence duale a été spécifiquement incorporé dans un peptide en des positions bien ciblées, ouvrant ainsi de nouvelles perspectives dans l'étude des fonctions de ces peptides (**Publication 1**).

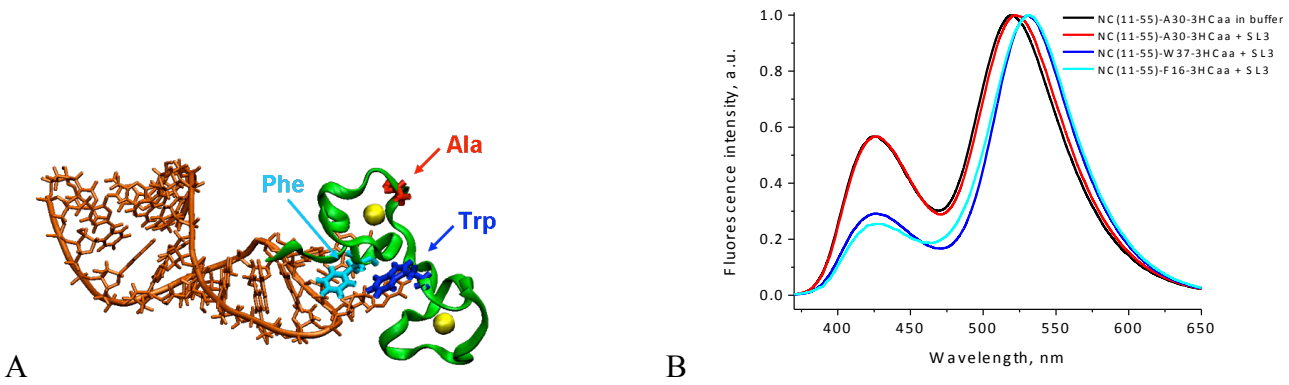


Figure 5. Structure 3D du complexe de la NC-SL3 basée sur des données RMN (A) ; Spectres normalisés de fluorescence des peptides NC(11-55)-F16-3HCaa, NC(11-55)-A30-3HCaa, NC(11-55)-W37-3HCaa libres en solution ou complexés à l'ARN SL3 (B).

En parallèle, afin de pallier aux défauts de la sonde 3HC, trois sondes optimisées (F6A, F6C et F4O) pour le marquage de l'extrémité N-terminale ont été synthétisées (Figure 6) (**Publication 2**).

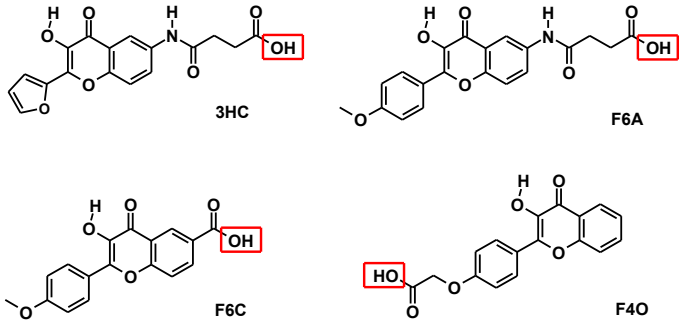


Figure 6. Structures des sondes étudiées.

Comparées à la sonde mère 3HC, les nouvelles sondes montrent une meilleure sensibilité du rapport d'intensité de leurs deux bandes d'émission (N*/T*) à la polarité et la capacité "donneur de liaisons hydrogène" du solvant, de meilleurs rendements quantiques en solution aqueuse et une sensibilité accrue à la liaison avec un ODN. Ces sondes ont été fixées de manière covalente à l'extrémité N-terminale du peptide Tat(44-61), pris comme modèle, et utilisées pour étudier la liaison de ce peptide avec différents ODNs. Par rapport à 3HC, l'ensemble de ces nouvelles sondes montrent une sensibilité au moins trois fois supérieure de leur émission duale aux interactions peptide-ODN, résultant de leur meilleure sensibilité à l'hydratation (Figure 7). La meilleure réponse a été observée avec F4O, qui présente la géométrie la plus compacte permettant le meilleur stacking avec les bases nucléïques et donc l'hydratation la plus faible. Ces trois sondes sont proposées pour fournir de nouveaux outils plus efficaces pour suivre de manière site-sélective les interactions peptide-ODN.

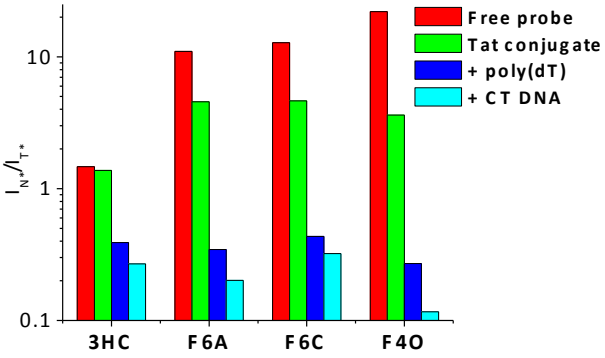


Figure 7. Rapports d'intensité N*/T* pour les sondes 3HC, F6A, F6C et F4O libres (rouge), conjuguées avec le peptide Tat(44-61) seul (vert) ou en présence de poly(dT) simple brin (bleu) ou de CT ADN double brin (cyan) en solution tamponnée.

Par ailleurs, une nouvelle approche pour l'estimation de la concentration locale en donneurs de liaisons hydrogène (HB) a été développée en utilisant des sondes 3HF à fluorescence duale affichant une forte sensibilité aux donneurs HB, mais une faible sensibilité à la présence de molécules dipolaires, au pH et à la force ionique du milieu (**Publication 3**). Pour ces sondes dans des mélanges binaires de solvant, une dépendance linéaire a été observée pour le logarithme du ratio de l'intensité des émissions de deux bandes, log(N*/T*), en fonction de la concentration locale de donneurs HB. La courbe de calibration obtenue nous a permis de déterminer quantitativement l'hydratation de la sonde, exprimée par le volume partiel de l'eau dans l'environnement de la sonde, ceci pour différents peptides marqués en position N-terminale et leurs complexes avec des ADN.

En collaboration avec l'équipe de A. Burger (Nice), de nouveaux analogues de nucléosides (Figure 8) ont été synthétisés dans lesquels la base naturelle a été substituée par un résidu 3-HC (**Publication 4**).

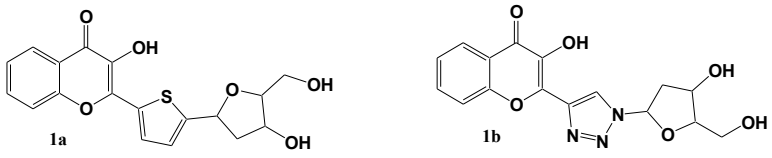


Figure 8. Structures des nucléosides 2-aryl-3HC **1a** et **1b**.

Les études de fluorescence ont montré que les molécules **1a** (*vide infra*) et dans une moindre mesure **1b** conservaient une émission duale fortement sensible à l'environnement (Figure 9). Ainsi, une diminution de la polarité du solvant induit un changement significatif du rapport N*/T* pour **1a**, ainsi qu'un déplacement des bandes T* et N*, respectivement vers le rouge et vers le bleu. Etant donnée la différence de polarité entre l'intérieur et la surface de l'hélice d'ADN, l'introduction d'un 3HC-nucléotide dans l'ADN devrait permettre de suivre les modifications du microenvironnement et de la dynamique de l'ADN au niveau du site de substitution.

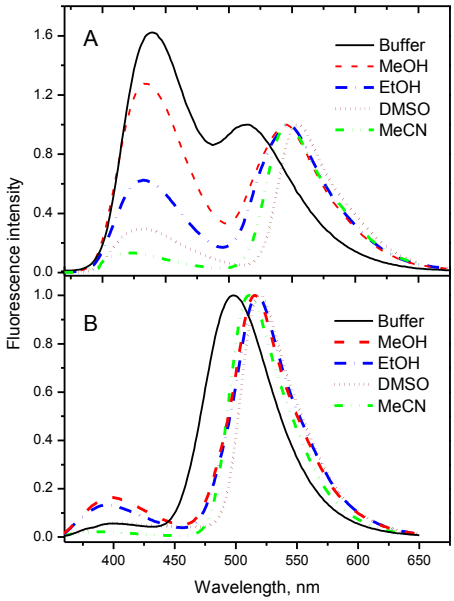


Figure 9. Spectres de fluorescence de **1a** (A) et **1b** (B) dans différents solvants.

L'incorporation de l'analogue 1a dans la séquence du *Primer Binding Site* (PBS) ou dans un duplex de courtes séquences d'ODN (double brin de 15-mers) ont montré que cet analogue était capable de se substituer aux bases naturelles et de s'empiler avec les bases adjacentes (Figure 6). En outre, des courbes de fusion par spectrophotométrie UV ont montré que ces analogues de nucléosides n'affectaient que marginalement la stabilité des ODNs ainsi marqués et que ces dérivés 3-HC pouvaient se substituer à toute base nucléique. Ces données indiquent que la 3-HC-thiényldéoxyribose possède les caractéristiques propres à une base universelle. Les ODNs marqués, double ou simple brin, montrent en outre une importante modification du rapport d'intensité N*/T* et une augmentation du rendement quantique en comparaison avec les valeurs obtenues pour les nucléosides libres, indiquant que la sonde couplée à l'ODN est dans un environnement hydrophobe. L'addition de NCp7 à des ODNs marqués provoque une modification significative du rapport N*/T*, indiquant ainsi que cette sonde est à même de suivre l'interaction (Figure 10). L'augmentation du rapport N*/T* s'accompagne d'une augmentation du rendement quantique, ce qui suggère que la sonde est déplacée dans un environnement moins hydrophobe, conséquence probable de l'interaction avec certains acides aminés de NCp7. L'ensemble de ces résultats suggère que la sonde permet de suivre de manière site-spécifique les interactions NCp7/ODN (**Publication 5**).

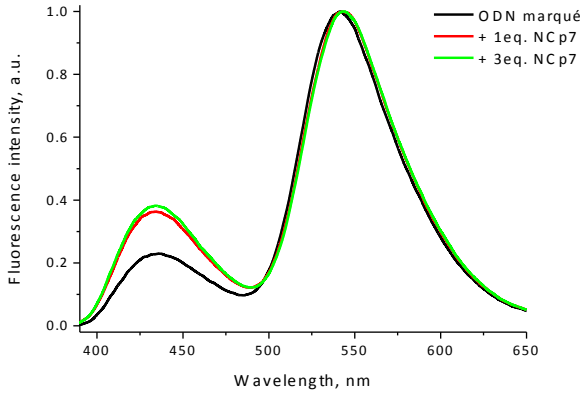


Figure 10. Spectres de fluorescence de l'ODN marqué avec la sonde 3HC en absence et en présence de NCp7.

La seconde partie de mon travail a consisté à développer une approche permettant de suivre la fixation et l'insertion de peptides, notamment la NCp7, à des biomembranes en utilisant des sondes fluorescentes de la famille des 3HC, sensibles à l'environnement, Une sonde 3HC très sensible aux changements de polarité dans des solvants peu polaires a été synthétisée. Afin de tester sa capacité à suivre les interactions peptides-membranes, nous avons marqué la partie N-terminale de trois peptides modèles, la mélittine, la magaïnine 2 et la poly-L-lysine, dont les caractéristiques de liaison aux membranes lipidiques étaient déjà connues (**Publication 6**).

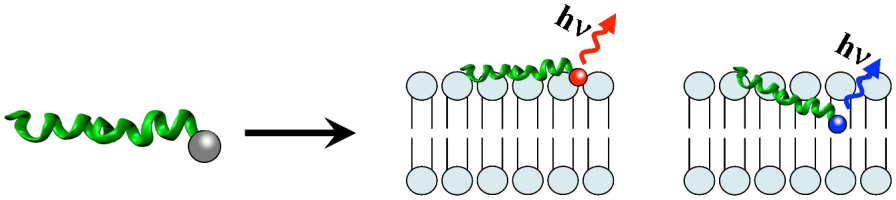


Figure 11. Liaison peptide-membrane. Le peptide (en vert) est marqué avec une sonde fluorescente (boule grise).

La liaison de ces peptides marqués à des vésicules lipidiques induit une forte augmentation de l'intensité de fluorescence, ce qui permet de quantifier l'interaction. En outre, l'émission duale de la sonde liée aux peptides est bien corrélée avec la profondeur dans la bicouche atteinte lors de son insertion, profondeur mesurée par la méthode d'inhibition de fluorescence dite du parallaxe. Le principe du suivi de l'interaction des peptides marqués avec les membranes lipidiques est montré sur la Figure 11. Avec la mélittine et la magaïnine 2, qui présentent l'insertion de leur extrémité N-terminale la plus profonde, la sonde montre une émission duale correspondant à un environnement de faible polarité, alors que l'extrémité N-terminale de la poly-L-lysine reste dans un environnement plutôt polaire, ce qui est cohérent avec une localisation proche de la surface de la bicouche. Cette sonde constitue donc un nouvel outil intéressant, cette fois pour suivre la fixation et l'insertion de peptides dans les membranes.

Nous avons ensuite synthétisé un peptide NCp7 marqué à sa partie N-terminale avec la sonde 3HC mentionnée ci-dessus en vue de détecter une éventuelle interaction de NCp7 avec des

biomembranes. Les propriétés du peptide marqué ont été trouvées voisines de celles du peptide natif. La fluorescence duale du peptide NCp7 marqué s'est révélé être extrêmement sensible lors de l'interaction avec des vésicules lipidiques unilamellaires (LUVs) considérées comme membranes modèles. Le peptide marqué en solution ne montre qu'une bande d'émission unique de faible intensité. En présence de vésicules anioniques, chargées négativement (DOPS, DOPG), l'intensité de fluorescence augmente fortement et les deux bandes d'émission dues à l'ESIPT apparaissent dans le spectre de fluorescence, confirmant ainsi une forte interaction (Figure 12). Au contraire, en présence de LUVs neutres (DOPC), les propriétés de fluorescence de 3HF-NCp7 ne sont pas affectées, montrant ainsi l'absence d'interaction. Dans une étape suivante, nous avons utilisé un mélange de lipides (DOPC: DOPE: DOPS: PI: PIP_2 dans une proportion molaire de 38:30:5:24:3 comparable à celle de la membrane nucléaire (*nuclear membrane like composition* – NMLC). L'addition de NC marqué à ce mélange conduit à un effet similaire à celui observé avec des vésicules chargées négativement.

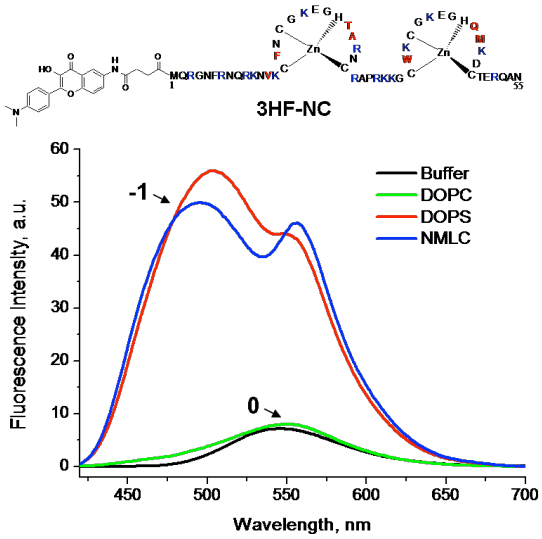


Figure 12. Structure du peptide NCp7 marqué à son extrémité N-terminale par la sonde 3HF et son spectre de fluorescence en absence et en présence de différents types de vésicules.

Afin d'étudier l'orientation du peptide NCp7 lié à une membrane, nous avons comparé sa fluorescence avec la série de peptides modèles mentionnés précédemment. Nous avons trouvé pour NCp7 une fluorescence proche de celle observée avec la poly-L-lysine, indiquant une localisation de NCp7 à l'interface membranaire. Une preuve supplémentaire est apportée par des mesures d'inhibition de fluorescence selon la méthode du parallaxe qui montrent que son extrémité N-terminale est située à ~ 17 Å du centre de la bicouche, c'est-à-dire au niveau des groupements phosphates des phospholipides.

De manière intéressante, l'interaction de NCp7 avec l'ADN résulte en une forte augmentation de l'affinité de l'interaction de ce peptide avec des vésicules neutres (DOPC). La formation d'un complexe ternaire peptide-ADN-membrane a été confirmée par microscopie de force atomique (AFM) sur des bicouches lipidiques supportées (Figure 13). En outre, l'AFM et l'imagerie bi-photonique de vésicules unilamellaires géantes indiquent que la NCp7 aussi bien en forme libre que

liée à des ODNs peut déstabiliser et rompre de manière concentration- et temps-dépendante la bicouche lipidique (Figure 13) [Manuscrit en préparation].

Considérés dans leur ensemble, nos résultats indiquent que le peptide NCp7 libre et son complexe avec des acides nucléiques peuvent tous deux déstabiliser des membranes lipidiques, suggérant que NCp7 puisse participer à l'entrée dans le noyau du complexe de pré-intégration (PIC) en déstabilisant la membrane nucléaire.

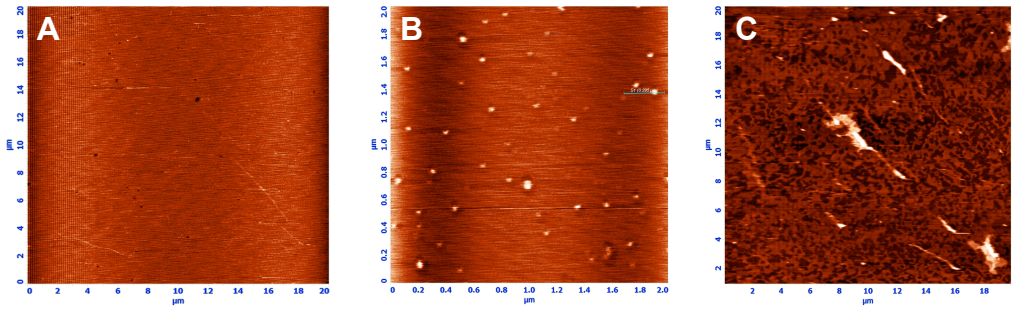


Figure 13. Images d'AFM sur des bicouches lipidiques supportées (A); complexes NCp7-DNA sur des bicouches supportées de DOPC (B); NCp7 sur des bicouches supportées de DOPS (C).

CONCLUSIONS. Dans ce travail, nous avons développé une méthodologie pour la détection des interactions des protéines avec des oligonucléotides et des membranes basée sur la sensibilité à l'environnement de sondes 3-hydroxychromone (3HC). De nouvelles séries de sondes sensibles à l'hydratation pour le marquage de peptides en position N-terminale ont été synthétisées. Ces nouvelles sondes présentent une meilleure sensibilité de leur émission duale à l'environnement. Afin de démontrer le rôle individuel de chaque résidu dans les interactions peptide-ODN, des acides aminés et des analogues nucléosidiques fluorescents ont été synthétisés et incorporés à différentes positions du peptide et de l'ODN, respectivement. Par ailleurs, nous avons développé une approche afin de suivre la liaison avec la membrane et l'insertion de ces peptides en utilisant une sonde fluorescente de la famille des 3-hydroxyflavones sensible à l'environnement. Ultérieurement, cette sonde a été utilisée pour étudier l'interaction de NCp7 avec des membranes lipidiques. Nous avons montré que NCp7, aussi bien à l'état libre que complexée à l'ADN, interagit et déstabilise les membranes lipidiques, ce qui suggère que NCp7 puisse participer à l'import nucléaire du complexe de pré-intégration (PIC). Sur la base de ces résultats, deux hypothèses quant au rôle de NCp7 dans l'internalisation nucléaire du PIC ont été proposées. Considérés dans leur ensemble, nos résultats montrent que des sondes fluorescentes à base de 3HC constituent des outils multi-paramétriques particulièrement intéressants pour le suivi d'interactions biomoléculaires impliquant des peptides ou des protéines.

PERSPECTIVES. Le présent travail peut se poursuivre dans différentes directions. Par exemple, la NC marquée en différentes positions avec l'analogue d'acide aminé fluorescent 3HCaa pourrait être utilisée pour caractériser la dynamique de son interaction avec PBS et TAR), ceci en s'appuyant sur des techniques résolues en temps. D'importantes informations complémentaires sont attendues en utilisant PBS et cTAR marqués avec des dérivés 3HC-nucléosides. Ces séquences marquées permettront de suivre de manière site-sélective les changements structuraux induits par

NC au niveau des ODNs. La comparaison des données au niveau du peptide et de l'ODN devrait permettre de mieux caractériser les mécanismes moléculaires de l'activité chaperonne de NV vis-à-vis des acides nucléiques.

Les sondes développées pourraient aussi être utilisés pour mieux caractériser les interactions NC-membranes. En utilisant les techniques de FRET et de FLIM-FRET, il sera possible de mieux déterminer la nature des aggrégats (particules de lipoprotéines) qui apparaissent lors de l'interaction de NC avec une bicouche lipidique supportée. En outre, la technique AFM / TIRF permettra sans ambiguïté de suivre des changements dans la structure et les propriétés de la bicouche induits par NC ou des complexes NC-ADN. Enfin, une collaboration avec un groupe de virologie sera mise en oeuvre afin de démontrer la pertinence fonctionnelle des interactions NC-membranes dans les cellules infectées. Les données obtenues devraient mettre en évidence pour la première fois une nouvelle propriété de NC, probablement importante quant à ses fonctions dans le cycle rétroviral, et ainsi initier un nouvel axe de recherche sur cette protéine.

Une autre orientation importante de ce travail consistera à poursuivre le développement des sondes sensibles à l'environnement pour l’étude des interactions site-spécifiques des protéines avec différentes cibles, ceci au niveau de la molécule unique. Une amélioration des sondes peut également être obtenue par la synthèse d’acides aminés non-naturels basés sur la 4'-(diméthylamino)-3-hydroxyflavone, afin de poursuivre l’étude des interactions peptide-membrane. Les peptides seront marqués en différents sites afin de déterminer l'implication des différents résidus dans ce type d’interactions. Par ailleurs, des informations site-spécifiques sur la conformation des protéines et leur position dans la membrane pourraient être obtenues.

Résumé: La spectroscopie de fluorescence est un outil très sensible pour étudier les changements de conformation et les interactions de protéines avec leurs cibles. À cet égard, les sondes fluorescentes sensibles à l'environnement, capables de détecter les changements d'environnement sur un site spécifique marqué, ont un intérêt particulier. Le but de ce travail était de développer une méthodologie basée sur des sondes de la famille des 3-hydroxychromones (3HC) pour détecter des interactions de protéines avec des oligonucléotides (ODNs) et des membranes. Du fait de leur réaction de transfert de proton intramoléculaire à l'état excité (ESIPT), ces fluorophores présentent deux bandes d'émission bien séparées et différemment sensibles à l'environnement, permettant ainsi de détecter des interactions via le changement du rapport d'intensité de ces deux bandes. Dans un premier temps, un analogue d'acide aminé fluorescent dérivé de la sonde 3HC, hautement sensible aux solvants protiques, a été synthétisé puis inséré à des positions ciblées de la protéine de la nucléocapside (NC) du VIH-1. Les peptides ainsi marqués ont été utilisés pour mieux caractériser l'interaction peptide-ODNs et fournir des informations site-spécifique sur les changements environnementaux, induits par l'interaction à proximité du site marqué. En utilisant différents oligonucléotides dont les structures en complexe avec NC ont été résolues, nous avons montré que la réponse ratiométrique de l'acide aminé fluorescent est corrélée avec la proximité de la base la plus proche au sein des complexes. Comme alternative, pour caractériser les interactions peptide-ODN, des analogues nucléosidiques fluorescents ont été synthétisés et introduits en différentes positions des ODNs. Ces séquences marquées ont permis d'obtenir des informations sur les changements environnementaux locaux induits par l'interaction. Par ailleurs, de nouvelles sondes sensibles à l'hydratation ont été synthétisées pour suivre de manière site-sélective les interactions peptide-ODN. Pour étudier les interactions peptide-membrane, nous avons développé une sonde 3HC sensible aux changements de polarité dans les solvants apolaires. Pour valider cette sonde, nous avons couplé cette sonde à l'extrémité N-terminale de différents peptides synthétiques comme la mélittine, la magainine 2 et la poly-L-lysine, connus pour interagir avec les membranes lipidiques. Nous avons montré que le rapport d'intensité de la sonde est bien corrélé à la profondeur d'insertion de la région N-terminale des peptides dans les membranes. Enfin, nous avons appliqué cette sonde pour détecter des interactions possibles de la NC avec les membranes. En suivant la fluorescence duale de la NC marquée à sa partie N-terminale, nous avons montré que la NC se lie avec une forte affinité sur les membranes contenant des lipides négativement chargés. Cette interaction, principalement de nature électrostatique, permet de localiser la protéine au niveau des têtes lipidiques. A forte concentration, la NC est capable de déstabiliser et perturber la membrane de manière dose-dépendante. Par ailleurs, nous avons montré que les complexes NC-ADN peuvent également se lier aux membranes et les déstabiliser d'une manière similaire à la NC libre, suggèrant que la NC pourrait participer à l'import nucléaire du complexe de pré-intégration.

Abstract: Fluorescence spectroscopy is a highly sensitive tool for investigating protein interactions and conformational changes. In this respect, environment-sensitive fluorescent probes, which monitor environment changes at a specifically labeled site, are of particular interest. The aim of this work was to develop a methodology for sensing interactions of proteins with oligonucleotides (ODNs) and membranes based on environment-sensitive labels from the 3-hydroxychromone (3HC) family. Due to an excited state intramolecular proton transfer (ESIPT), these dyes exhibit two highly resolved emission bands, differently sensitive to the environment, thus allowing to sense interactions through changes in their intensity ratio. Primarily, a fluorescent amino acid analogue, based on 3HC dye which is highly sensitive to protic solvents, was synthesized and inserted at selected positions of the HIV-1 nucleocapsid protein (NC) to further characterize the peptide-ODNs interaction and provide site-specific information on the environmental changes induced by the interaction close to the labelling site. Using different oligonucleotides whose structures in complex with NC have been solved, we found that the ratiometric response of the fluorescent amino acid correlates well with its proximity to the closest base in the complexes. As an alternative, to reveal the role of individual residues in peptide-ODN interactions, fluorescent nucleoside analogues were synthesized and applied to different positions of ODNs. Afterwards, the interaction of the labeled ODNs with NC was characterized, providing site-specific information on local environment changes induced by the interaction. Moreover, new hydration-sensitive probes were synthesized to site-selectively monitor the binding of peptides to ODNs. For investigation of peptide-membrane interactions, we further developed a 3HC probe sensitive to polarity changes in apolar solvents. To validate this probe, we synthesized melittin, magainin 2 and poly-L-lysine, which are known to interact with lipid membranes, and labeled their N-terminus with this probe. The observed intensity ratio of the probe was found to correlate well with the insertion depth of the N-terminal region of the peptides. Finally, we applied this probe for the detection of possible interactions of NC with membranes. The two-color fluorescence of the N-terminally labeled NC was exquisitely sensitive to the binding to lipid membranes, showing high affinity to membranes containing negatively charged lipids. This interaction was mainly driven by electrostatic forces and locates the protein at the level of the lipid heads. At high concentrations, NC was able to destabilize and disrupt the membrane in a concentration-dependent manner. Moreover, it was shown that NC-DNA complexes can also bind and destabilize the membrane similarly to free NC, suggesting that NC peptide could participate in the nuclear import of the pre-integration complex.