

École doctorale sciences de la vie et de la santé

THÈSE

Présentée pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR
Discipline : Science du vivant
Domaine : Aspects moléculaires et cellulaires de la biologie

Par

ÉLODIE TOUZÉ

***CRISTALLOGENÈSE ET ÉTUDES STRUCTURALES
APPLIQUÉES AUX AMINOACYL-ARNt SYNTHÉTASES***

Soutenue le 16 novembre 2007 devant la commission d'examen :

Pr Y. MECHULAM	Rapporteur externe
Pr A. DUCRUIX	Rapporteur externe
Pr H.D. BECKER	Membre invité
Dr J.P. RENAUD	Rapporteur interne
Pr E. WESTHOF	Examineur
Dr R. GIEGÉ	Directeur de thèse

ABREVIATIONS.....	5
- INTRODUCTION -	7
INTRODUCTION GENERALE	8
I. Vue d'ensemble des fonctions et structures des aaRS.....	9
A. La réaction d'aminacylation des ARNt.....	9
B. Similarités et divergences des aaRSs.....	10
1. Le classement des aaRS en 2 familles	10
a. Les aaRS de classe I.....	10
b. Les aaRS de classe II	11
2. Quelques exceptions à la règle	13
C. Vision structurale de la fonction et de l'évolution des aaRS	14
1. Site catalytique et spécificité pour l'acide aminé.....	14
2. Compréhension structurale et correction d'erreur	17
3. Reconnaissance de l'ARNt et rôle des éléments d'identité.....	18
4. La nécessité d'erreurs.....	20
5. Divergences et similarités des synthétases : notion de modularité	22
6. Les complexes multisynthétasiques.....	24
D Apparition de fonctions nouvelles et évolution.....	26
1. Les multiples fonctions des synthétases	26
2. Les paralogues	28
II. Cristallogénèse et importance des aaRS pour la biologie structurale.....	29
A. Aspects généraux : croissance des cristaux de macromolécules	29
1. La cristallisation	29
2. Caractéristiques spécifiques des macromolécules	29
3. Méthodes de cristallisation et cristallogénèse.....	31
4. Qualité des produits	31
5. Recherche de conditions	32
6. Les modèles utilisés en cristallogénèse	33
B. Importance des aaRS en cristallogénèse et en biologie structurale	34
1. Les aaRS dans l'histoire de la cristallisation des macromolécules biologiques.....	34
2. Modèles en ingénierie.....	35
C. Vision globale de la cristallographie des aaRS et contribution à leur connaissance générale.....	36
1. Compilation des résultats.....	41
2. Quelques leçons tirées des données structurales des aaRS	46
III. Ma contribution à l'étude des aminoacyl-ARNt synthétases.....	49
- RÉSULTATS ET DISCUSSION -	51
I. Préambule.....	52
A. Originalités de la GlnRS de <i>Deinococcus radiodurans</i>	52
B. La GlnRS : une aaRS chimère issue de deux voies d'aminacylation de l'ARNt	52
1. Structure primaire	52
2. Cœur de la GlnRS-Dr.....	53
C. Rôle de l'extrémité C-terminale et implications du domaine YqeY.....	55
1. Extrémité C-terminale, indispensable à la réaction d'aminacylation.....	55
2. YqeY : un 'module' libre ou lié	55
D. Particularité de l'empilement cristallin de la GlnRS ou la cristallisabilité d'une enzyme désordonnée.....	56
DISORDER CAN EXIST INSIDE WELL-DIFFRACTING CRYSTALS – ARTICLE 1.....	59

II.	Etude cristallographique de la GlnRS de <i>D. radiodurans</i>	63
A.	Les objectifs de l'étude	63
B.	Stratégies de cristallisation	64
1.	Recherche de nouvelles conditions de cristallisation	64
2.	Cristallisation en gel	66
3.	Stabilisation par modification des résidus lysine	66
4.	Données cristallographiques	68
C.	Analyse des nouvelles données structurales	70
1.	GlnRS- <i>Dr</i> cristallisée en présence d'analogue d'adénylate, QSI	70
	Description générale	70
	Site actif de l'enzyme et contacts spécifiques avec le substrat	72
	Position de l'adénylate dans le site actif	72
2.	Comparaison avec la position du glutaminol dans le site actif	75
3.	GlnRS- <i>Dr</i> en présence de glutamine : la surprenante conformation du site actif	76
	Description générale	76
	Contacts spécifiques de la boucle et variation des positions des résidus impliqués dans la reconnaissance des substrats	79
	Position de la boucle du KMSKS	82
	Conclusions	82
4.	Comparaison avec la GlnRS- Δ C	85
III.	Plasticité de l'empilement cristallin	88
A.	Les contacts cristallins	88
B.	Les canaux de solvant	91
IV.	Etudes en solution de la GlnRS	94
A.	Résonance magnétique nucléaire	94
B.	Analyse par dichroïsme circulaire de l'extrémité YqeY	98
- CONCLUSIONS ET PERSPECTIVES -		101
- MATÉRIEL ET MÉTHODES -		107
I.	Clonage des protéines étudiées	108
II.	Cultures de micro-organismes	109
A.	Milieux de culture	109
1.	Milieu LB (Luria Broth)	109
2.	Milieu M9 (minimum)	109
3.	Conduite des cultures	110
III.	Purification des protéines	110
A.	Purification d'affinité sur matrice de chitine	110
B.	Purification sur matrice substituée à l'héparine	112
C.	Chromatographie d'exclusion de taille	112
D.	Séparation des protéines sur gel de polyacrylamide en conditions dénaturantes	113
E.	Dialyses	114
F.	Concentration	114
G.	Détermination de la concentration	114
IV.	Préparation d'ARN de transfert	115
A.	Transcription <i>in vitro</i>	115
B.	Purification des transcrits sur gel préparatif	115
C.	Test d'activité des aminoacyl-ARNt synthétases : aminoacylation <i>in vitro</i>	116
V.	Techniques relatives à l'analyse de l'échantillon contenant la protéine	117
A.	Spectrométrie de masse	117
B.	Diffusion de lumière	117
C.	Méthylation des résidus lysines	118

VI.	Etudes structurales des protéines.....	119
A.	Cristallographie des protéines	119
1.	Méthodes de cristallisation	119
a.	Diffusion de vapeur	119
b.	Cristallisation en “batch”	119
c.	Cristallisation en gel	120
d.	Cristallisation en champ magnétique	120
2.	Criblage de conditions de cristallisation	121
3.	Optimisation des “hits”	121
4.	Observation et analyse du contenu des cristaux	122
5.	Congélation des cristaux.....	122
B.	Analyse cristallographique	122
C.	Résonance Magnétique Nucléaire (RMN) en biologie structurale	124
D.	Dichroïsme circulaire	126
VII.	Outils bioinformatique.....	127
VIII.	Outils bureautiques	128
- ANNEXES -		129
ANNEXE I – LA BACTERIE DEINOCOCCUS RADIODURANS	131	
Une bactérie ultrarésistante.....	131	
Domaines insertions, une routine pour Deinococcus.....	132	
Système traductionnel et aminoacylation espèce spécifique	132	
ANNEXES II – ETUDE STRUCTURALE DE LA TYRRS MITOCHONDRIALE HUMAINE		
ARTICLES 2 ET 3	133	
REFERENCES BIBLIOGRAPHIQUES.....	185	

Abréviations

A	adénine
aa	acide aminé
aaRS	aminoacyl-ARNt synthétase
ADN	acide désoxyribonucléique
Amp	ampicilline
AMP	adénosine monophosphate
ARN	acide ribonucléique
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ATP	adénosine triphosphate
A _x	absorbance à x nanomètres
BSA	serum albumine bovine
BET	bromure d'éthidium
β-ME	beta-mercaptoéthanol
BSA	albumine de sérum bovin
C	cytidine
CCP4	Collaborative computational projet number 4 in protein crystallography
CNS	Crystalligraphy & NMR system
<i>D. radiodurans</i>	<i>Deinococcus radiodurans</i>
Da	dalton
DTE	dithioérrithritol
DTT	dithiothrétol
DTNB	dithionitrobenzoate
DMSO	diméthylsulfoxyde
DO _x	Densité optique à x nanomètres
EDTA	éthylène diamine tétraacétate de sodium
ESRF	European synchrotron radiation facility
<i>E. coli</i>	<i>Escherishia coli</i>
G	guanine
HEPES	acide N-2 hydroxyl éthyl pipérazine N'2 éthane sulfoxyde
IPTG	isopropyl-beta-D-thiogalactopyranoside
K	Kelvin
K _{cat}	constante catalytique
K _M	constante de Michaelis
L	litre

M	molaire
m	mili
min	minute
mm	millimètre
m/v	masse pour volume
μ	micro
n	nano
NTP	nucléoside triphosphate
PDB	Protein data bank
PEG	polyéthylèneglycol
PPI	pyrophosphate
RMSD	root mean square deviation
RMN	resonance magnétique nucléaire
RX	rayons X
rpm	rotation par minute
SDS	sodium dodécyl sulfate
T	thymine
TBE	Tris-borate-EDTA
TCA	acide trichloroacétique
Tris	tris (hydroxyméthyl) amino éthane
U	uridine
<i>U</i>	unité
UV	ultra-violets
v/v	volume pour volume
°C	degré Celsius

Chaque acide aminé est abrégé selon son code à trois lettres

Chaque aaRS est identifiée pour sa spécificité pour l'aa correspondant

- INTRODUCTION -

Introduction générale

Les aminoacyl-ARNt synthétases (aaRS) sont une famille d'enzymes chargées de catalyser le branchement covalent d'un acide aminé donné sur son ARNt homologue. Les complexes d'aminacyl-ARNt, produits de cette réaction appelée aminacylation, et acheminés vers le ribosome, se fixent spécifiquement sur l'ARN messager. Ils agissent donc comme des adaptateurs qui lient l'information génétique à la séquence protéique. La délivrance de l'acide aminé au peptide naissant dépend essentiellement de l'interaction des paires de bases entre les triplets de codon et d'anticodon. Ainsi, la spécificité de la réaction d'aminacylation des ARNt par leur aaRS constitue une étape cruciale à la fidélité de la traduction.

L'ensemble des études investies dans ce domaine combinent des études tant biochimiques que structurales ou bioinformatiques et ont permis d'élucider la base de cette faculté de reconnaissance hautement spécifique. La distribution universelle des aminoacyl-ARNt synthétases à travers les 3 règnes du vivant (bactéries, eucaryotes, archées) fait qu'elles appartiennent à une des plus vieilles familles de polypeptides. C'est pourquoi elles sont d'un intérêt primordial pour mieux connaître l'évolution des espèces et ont fait l'objet d'études variées qui ont permis de clarifier de nombreux phénomènes. Cependant, de multiples questions restent sans réponse car leur diversité s'étend à l'intérieur même d'un système.

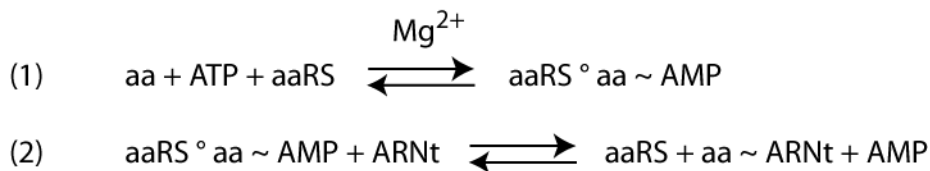
Une première partie de l'introduction de cette thèse sera consacrée aux fonctions et structures des aaRS ainsi qu'aux connaissances sur leur évolution dans les règnes du vivant. Les stratégies utilisées par les aaRS et les ARNt, acteurs de la réaction d'aminacylation, pour opérer une réaction spécifique, seront exposées.

La cristallographie des rayons X est un puissant outil pour comprendre les caractéristiques fonctionnelles d'une protéine et la majeure partie de nos connaissances structurales des aaRS en dépend. Les problèmes rencontrés lors de la cristallogénèse des macromolécules biologiques ont conduit à utiliser des modèles moléculaires pour faciliter et comprendre ce mécanisme. Les aaRS font partie des macromolécules complexes qui ont joué un rôle dans la compréhension de la cristallogénèse. Ceux-ci seront développés dans la seconde partie de cette introduction. J'y dresserai ensuite l'état des lieux de l'ensemble des travaux réalisés ayant conduit à la résolution d'une structure tridimensionnelle d'aminacyl-ARNt synthétase.

I. Vue d'ensemble des fonctions et structures des aaRS

A. La réaction d'aminacylation des ARNt

Mécanisme réactionnel : De manière générale, les aaRS sont au nombre de 22, une par acide aminé. La réaction d'aminacylation nécessite, en plus de l'acide aminé et de l'ARNt, la présence de petits substrats qui ont pour fonction à la fois d'activer les réactifs et de donner l'énergie nécessaire à la catalyse de la réaction. Ainsi, dans une première étape, l'ATP (Adénosine tri-phosphate) active l'acide aminé en s'y branchant via le carboxyle (C=O) α de l'acide aminé, en présence de magnésium (1). L'activation de l'acide aminé permet, lors d'une seconde étape, son transfert sur le 2'OH ou 3'OH de l'adénosine terminale de l'ARNt (2).



Cette réaction en deux étapes permet une discrimination plus efficace de l'acide aminé, puisqu'il subit une double sélection avant d'être chargé sur l'ARNt. Certaines aaRS ont besoin de leur ARNt pour permettre la réaction d'activation. C'est le cas des ArgRS, GluRS, GlnRS et LysRS1 qui, en présence de leur ARNt homologue subissent les réarrangements conformationnels nécessaires pour assurer cette réaction.

Spécificité et erreur : Parmi les 20 acides aminés présents dans la cellule, une aaRS doit, pour opérer une réaction spécifique, différencier (ou discriminer) des acides aminés parfois structuralement et/ou chimiquement très proches. Pour cela, certaines d'entre elles ont évolué en développant une troisième étape dite d'édition qui éliminent les acides aminés non spécifiques accidentellement activés ou liés à l'ARNt (voir § I.B.1 et I.C.3). Cependant, la majorité de ces enzymes ont la capacité de discriminer entre l'un et l'autre acide aminé sans l'aide d'une fonction annexe en développant des capacités de reconnaissance structurales précises et grâce à un éventail de stratégies variées (§ I.C.1).

B. Similarités et divergences des aaRSs

1. Le classement des aaRS en 2 familles

Malgré leur rôle fonctionnel commun dans la biosynthèse protéique, les aaRS présentent des divergences significatives en taille, oligomérisation, séquence et en structure tridimensionnelle. Le panel de connaissances acquises grâce, entre autres, aux recherches d'ordre génétiques, biochimiques et structurales, sur les aminoacyl-ARNt synthétases a permis de les grouper en deux classes et notamment grâce à la découverte de motifs structuraux reflétant la topologie du domaine catalytique (Cusack *et al.*, 1991; Eriani *et al.*, 1990; Nagel et Doolittle, 1991).

a. Les aaRS de classe I

Les aaRS de classe I sont pour la plupart monomériques ou homodimériques. Les 11 aaRS de classe I ont été subdivisées en 3 sous-classes sur la base d'homologies de séquences (Cusack *et al.*, 1991; Eriani *et al.*, 1990). La classe Ia (Met-, Ile-, Leu-, Val-, Cys-, Arg-RS) regroupe les aaRS spécifiques des acides aminés hydrophobes, dont certaines ont développé un mécanisme de correction d'erreur ou édition; la classe Ib est spécifique d'acides aminés à longues chaînes latérales (Glu- et Gln-RS); et la classe Ic d'acides aminés aromatiques (Tyr- et Trp-RS), dont l'organisation dimérique est essentielle à leur activité enzymatique (Brick *et al.*, 1989; Doublé *et al.*, 1995).

Toutes les aaRS de classe I présentent 2 motifs consensus HIGH (Histidine – Isoleucine – Glycine- Histidine) et KMSKS (Lysine – Méthionine – Sérine – Lysine – Sérine) ainsi que des caractéristiques structurales communes. La résolution de la première structure d'aaRS, la TyrRS de *Bacillus stearothermophilus* a permis de mettre en évidence un domaine retrouvé dans toutes les aaRS de classe I : le domaine de Rossmann (Brick et Blow, 1987). Celui-ci est uniformément constitué d'une cage $\alpha/\beta/\alpha$ et d'un cœur interne de 5 brins β parallèles. Toujours trouvé dans la famille des deshydrogénases, sa présence témoigne d'une fonction de liaison à des nucléotides adényliques (ATP, NAD, NADP, FAD). Il a été caractérisé pour la première fois dans la glyceraldéhyde tri-phosphate deshydrogénase (Rossmann *et al.*, 1974) et son rôle, chez les aaRS est d'accommoder l'ATP dans le site actif lors de l'activation de l'acide aminé. Le domaine de Rossmann des aaRS de classe I se distingue également de celui d'autres protéines par la présence de domaines d'insertions CP1 intervenant entre autres dans les fonctions d'édition. Enfin, les régions des séquences consensus HIGH et KMSKS sont impliquées dans les interactions spécifiques avec l'ARNt et l'ATP, et sont mobiles, leur orientation dépendant des substrats en présence.

Le classement des aaRS concerne également le mode de fixation à leurs différents

substrats. Pour les enzymes de classe I, l'acide aminé activé (aa-AMP) est fixé en position 2' OH de l'adénosine A76 terminale de l'ARNt. L'aaRS reconnaît préférentiellement le petit sillon de l'ARNt et la conformation étendue de l'ATP.

Historiquement, les premiers travaux ont été effectués sur des aaRS qui, *a posteriori*, ont été assignées à la classe I. On pensait alors que les aaRS étaient universellement conservées, présentant un profil structural unique.

Classe	Oligométrie	Classe	Oligométrie
Classe I	Motif HiGH, kmsKs	Classe II	Motif (1) g(Φxx)2pxxpΦΦ] ; (2) [fRxEx/RxxxExx(d/e)] ; (3) (gΦ)3(d/e)R(Φ)5
Sous classe Ia		Sous classe IIa	
ArgRS*	α	GlyRS #	(αβ) ₂
CysRS	α	HisRS	α ₂
IleRS	α	ProRS	α ₂
LeuRS	α	ThrRS	α ₂
MetRS #	α ₂	AlaRS #	α ₄
ValRS	α	SerRS	α ₂
LysRS1*°	α		
Sous classe Ib		Sous classe IIb	
GlnRS*	α	AsnRS	α ₂
GluRS*	α	AspRS	α ₂
		LysRS2°	α ₂
Sous classe Ic		Sous classe IIc	
TrpRS	α ₂	PheRS #	(αβ) ₂
TyrRS	α ₂	PylRS	
		SepRS	

* aaRS ARNt-dépendants pour l'activation de l'acide aminé
° La LysRS est trouvée dans les 2 classes
Synthétases dont la structure oligomérique n'est pas conservée dans l'évolution
Φ: acide aminé hydrophile

Tableau I. 1 : Classification des aminoacyl-ARNt synthétases. Les séquences consensus HIGH et KMSKS des aaRS de classe I ne sont pas strictement conservées. La PylRS et la SepRS sont codées dans certains génomes d'archées méthanogènes.

b. Les aaRS de classe II

La séparation des aaRSs en deux classes a été établie grâce à la résolution de la structure de la SerRS d'*Escherichia coli* (Cusack *et al.*, 1990) suivie de celle de l'AspRS de *Saccharomyces cerevisiae* (Ruff *et al.*, 1991). Les aaRS de classe II présentent un échafaudage structural rare seulement retrouvé chez les holoenzyme biotine synthétase¹ (Murzin *et al.*, 1995). Elles sont caractérisées par un repliement α/β comprenant un cœur central de feuillets β flanqué d'hélices α. Ces enzymes sont la plupart du temps retrouvées sous forme d'homodimère, ou de tétramère sauf la PheRS mitochondriale humaine et de levure, monomérique ou la GlyRS trouvée sous forme tétramérique ou dimérique (Bullard *et al.*, 1999; Diaz-Lazcoz *et al.*, 1998, Tang *et al.*, 2005). Elles

¹ Répresseur de transcription de la synthèse de biotine, situé dans la région fonctionnelle de ligase et responsable de l'activation de la biotine et de son transfert à ses protéines acceptrices

ne sont pas caractérisées par un haut degré de similarité mais comportent au moins 3 séquences signatures dégénérées : le motif 1 $[G(\Phi xx)_2PxxP\Phi\Phi]$ intervient dans l'interface dimérique, les motifs 2 $[FRxE/RxxxFxxx(d/e)]$ et 3 $[(G\Phi)_3(d/e)R(\Phi)_5]$ sont situés au niveau du site actif et y jouent un rôle structural dans la reconnaissance de l'ATP et de l'ARNt. Le motif 2 marque un changement de conformation pour le positionnement correct des résidus impliqués lors de l'étape d'activation (Arnez et Moras, 1997; Cavarelli *et al.*, 1994). L'intermédiaire aminoacyl-adényle formé, les résidus polaires du motif libèrent le site actif pour autoriser de nouvelles interactions avec l'extrémité 3' de l'ARNt (Arnez *et al.*, 1995). Les aaRS de classe II fixent l'ATP sous sa forme repliée et lient l'ARNt via le grand sillon en position 3'OH. Précisons que le facteur EF-Tu qui prend en charge les ARNt aminoacylés pour les acheminer jusqu'au ribosome agit comme une isomérase et déplace systématiquement l'acide aminé sur le 3'OH du ribose (Förster *et al.*, 1994).

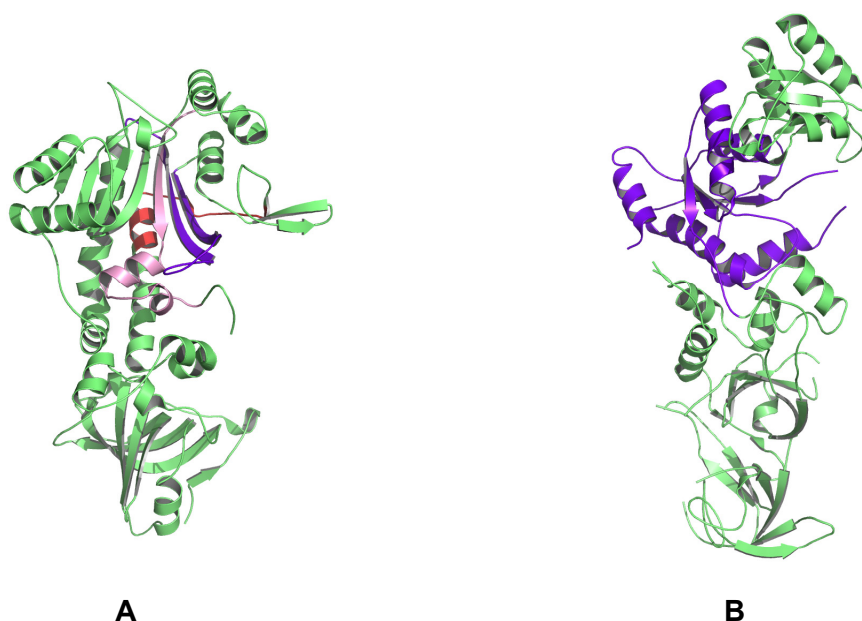


Figure 1. 2 : Mise en relief des différences d'organisation du site actif des aaRS : exemple d'une aaRS de classe II, l'AspRS de levure (A) , et de classe I, la GlnRS d'*Escherichia coli* (B).

A - Les motifs conservés chez les aaRS de classes II sont : le motif 1 (rouge, résidus 258–275), motif 2 est composé de 2 brins antiparallèles (violet, résidus 316–349) connectés par une grande boucle, le motif 3 est constitué d'un brin adjacent aux 2 précédents suivi d'une hélice α (rose pâle, résidus 517–548). Le motif 2 est impliqué dans la liaison avec l'ATP et l'extrémité 3' de l'ARNt et contribue, en adoptant une conformation fermée, à l'ancrage des substrats dans la cavité du site actif. Les 2 étapes de la réaction d'aminoacylation ont d'ailleurs été déduites à partir des structures cristallographiques de l'AspRS de levure (Cavarelli *et al.*, 1994).

B - Le site actif des classes I est constitué du domaine de Rossmann riche en feuillets β et flanqué d'hélices α (en bleu, 8-101, 212-264). Les adaptations induites qui assurent le positionnement correct du substrat de la GlnRS sont tributaires de la fixation de l'ARNt homologue. En dépit de l'architecture conservée du domaine de Rossmann, les modifications conformationnelles causées par la reconnaissance de l'ARNt ne sont pas conservées chez les classes I.

2. Quelques exceptions à la règle

Les aaRS des 2 classes sont généralement monophylétiques². Par exemple, toutes les TyrRS bactériennes et archées forment un groupe distinct excluant les autres aaRS. La règle selon laquelle chaque acide aminé est exclusivement représenté par une synthétase de classe I ou II a été discutée suite à la découverte de la LysRS de *Methanococcus maripaludis* (archée) (Ibba *et al.*, 1997). La LysRS est la seule aaRS représentée dans les 2 classes : une LysRS de classe II est présente chez la majorité des bactéries et des eucaryotes et chez quelques archées, tandis que la LysRS de classe I (LysRS1) est présente chez la majorité des archées et quelques bactéries. Les LysRS de classe II ont un ancêtre commun avec l'AspRS et l'AsnRS tandis que la classe I est proche de la GluRS (Ambrogelly *et al.*, 2002; Woese *et al.*, 2000, Levengood *et al.*, 2007).

Une seconde exception concerne la GlyRS qui présentent 2 types de structures quaternaires. Pour les bactéries et organelles, la GlyRS est un tétramère α_2/β_2 , alors que la GlyRS dimérique (α_2) est retrouvée dans tous les groupes d'organismes. La divergence élevée de séquence entre les GlyRS d'oligomérisation différentes lui donne une propriété unique (Wolf *et al.*, 1999). Cependant, les données structurales ont montré que le cœur catalytique de la sous-unité α des GlyRS dimérique et tétramérique sont homologues contrairement au domaine de liaison à l'anticodon de l'ARNt qui diffère bien d'un type à l'autre (Tang et Huang, 2005). Néanmoins leur séquence et leur taille partagent une faible similarité et laissent à penser qu'elles ont des origines différentes (O'Donoghue et Luthey-Schulten, 2003; Woese *et al.*, 2000).

2 En systématique phylogénétique, groupe qui comprend une espèce ancestrale et tous ses descendants.

C. Vision structurale de la fonction et de l'évolution des aaRS

Lors de la réaction d'aminacylation, les différents substrats en scène subissent des changements conformationnels assurant à la fois la spécificité et l'efficacité de la réaction. Il existe un lien primordial entre la formation du complexe spécifique aaRS-ARNt et les capacités catalytiques de l'aaRS pour la réaction d'aminacylation. L'accumulation des données cristallographiques concernant ces complexes a permis d'observer les modifications conformationnelles des deux partenaires en interaction. Les modifications majeures s'opèrent dans la région du site actif de l'enzyme ainsi que pour l'ensemble du squelette de l'ARN de transfert.

1. Site catalytique et spécificité pour l'acide aminé

La discrimination des acides aminés par les aaRS est potentiellement problématique car elle ne peut simplement reposer sur leur propriété d'encombrement stérique, leur structure ou des éléments spécifiques comme c'est le cas de l'ARNt. Pour les différencier, l'enzyme emploie d'autres stratégies qui reposent d'abord sur le rejet des acides aminés plus encombrés puis sur la conformation du site actif et les modifications structurales induites par la fixation de l'acide aminé. L'échafaudage précis du site actif lui permet d'établir des liaisons préférentielles pour positionner correctement l'acide aminé.

La discrimination d'un acide aminé peut s'opérer à différentes échelles. Par exemple, la GlnRS d'*E. coli* doit discriminer la glutamine du glutamate, qui est très proche et susceptible d'être accueilli dans son site actif. Les auteurs proposent que la discrimination du glutamate par la GlnRS d'*E. coli* s'effectue par l'intermédiaire d'un seul résidu, l'Arg30, qui, en déterminant négatif, positionnerait le glutamate dans une autre orientation (Bullock *et al.*, 2003), (Figure I. 3). Ce mécanisme n'est néanmoins pas encore caractérisé de façon claire.

De manière générale, la liaison à l'acide aminé provoque des modifications structurales ou des adaptations induites du site actif. Dans de nombreuses aaRS de classe I, telles les LeuRS, TrpRS, TyrRS et MetRS, (Ilyin *et al.*, 2000; Kobayashi *et al.*, 2005; Schmitt *et al.*, 1995), la boucle mobile "KMSKS", en conjonction avec le motif HIGH, est essentielle à la stabilisation de l'état de transition et à la discrimination des acides aminés (Cusack *et al.*, 2000; Ilyin *et al.*, 2000). En particulier, la région du KMSKS de la TyrRS d'*E. coli* adopte 3 conformations lors de la réaction d'activation de l'acide aminé : (i) une conformation ouverte lors de la fixation de la tyrosine (ii) une conformation fermée lors de la formation du couple tyrosyl-AMP (iii) une conformation semi-ouverte qui permettrait de guider et lier l'ARNt^{Tyr} (Kobayashi *et al.*, 2005).

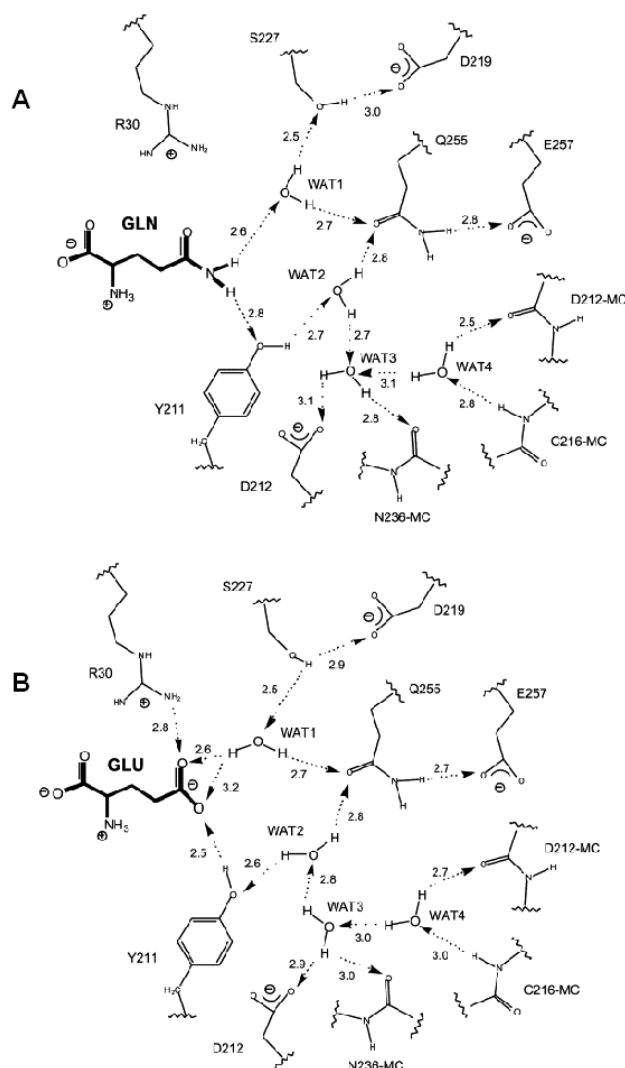


Figure 1. 3 : Représentation schématique de l'organisation des liaisons hydrogène du site actif de la GlnRS d'*E. coli* en présence de deux ligands (glutamine ou GLN, A; glutamate ou GLU, B). Le modèle propose que le site actif de l'enzyme discrimine la glutamine du glutamate grâce à la présence d'une arginine. Le groupe carboxyle qui distingue le glutamate de la glutamine entre en interaction avec un résidu conservé du site actif (Arg30). Schéma d'après (Bullock *et al.*, 2003).

Une autre stratégie consiste en la présence d'un domaine de liaison au zinc, caractérisé par sa propriété de liaison aux acides nucléiques. Chez *E. coli*, la GluRS présente un domaine signature dont les 3 cystéines lient le zinc qui intervient dans le positionnement correct de la tige acceptrice de l'ARNt^{Glu}. Le domaine module la liaison du glutamate au site actif par l'intermédiaire de l'ARNt, indispensable à l'activation de l'acide aminé (Sekine *et al.*, 2006, Banerjee *et al.*, 2004; Liu *et al.*, 1995).

Plus généralement, une adaptation du site actif pour toutes les aaRS dont l'activation de l'acide aminé est ARNt-dépendante (ArgRS, GlnRS, LysRS1) est nécessaire pour le passage d'une forme active à inactive. Ces adaptations induites contribuent à positionner correctement les substrats dans le site actif et sont donc indirectement impliquées dans la discrimination de l'acide aminé (Delagoutte *et al.*, 2000; Sekine *et al.*, 2003; Sherlin et Perona, 2003).

Pendant longtemps, les aaRS de classe II ont été considérés comme plus spécifiques vis-à-vis de l'acide aminé que celles de classe I bien que des activités d'éditions aient été décrites pour les PheRS de levure, de *T. thermophilus* et d'*E. coli* (Lin *et al.*, 1983; Mosyak *et al.*, 1995; Roy *et al.*, 2004). Néanmoins, des études biochimiques sur la LysRS de classe II d'*E. coli* ont montré que l'enzyme est capable d'activer 7 acides aminés hétérologues (Met, Leu, Cys, Ala, Thr, Arg, Ser), tous, excepté l'Arg sont de taille inférieure à la Lysine (Jakubowski, 1999).

2. Compréhension structurale et correction d'erreur

Dans le but d'améliorer la discrimination des acides aminés, une autre stratégie a été, pour certaines aaRS, de développer une fonction d'édition par correction d'erreur. Le premier contrôle consiste à sélectionner les acides aminés par leur encombrement stérique. Un second contrôle est nécessaire pour les plus petits acides aminés comme la Valine capable d'être accueillie dans le site actif de l'IleRS et visualisé dans les structures cristallographiques (Nureki *et al.*, 1998; Silvian *et al.*, 1999). Le mécanisme d'édition consiste soit en l'hydrolyse de l'acide aminé activé non spécifique (correction pré-transfert) soit du couple non homologue aa₍₁₎-ARNt^{aa}₍₂₎ (correction post-transfert), leur interdisant définitivement le chemin du ribosome. Les sites d'éditions des 2 classes sont peu comparables. Trois aaRS de la sous-classe Ia, l'IleRS, la ValRS et la LeuRS, présentent une large insertion CP1 ou site d'édition connecté au domaine de Rossmann et responsable de la correction d'erreurs. Dans le cas de la LeuRS, des résidus conservés du site d'édition reconnaissent la chaîne principale des substrats non spécifiques. Une étroite poche y accueille la chaîne latérale des acides aminés non spécifiques Met et Ile activés où ils sont hydrolysés (Lincecum *et al.*, 2003; Liu *et al.*, 2006). Les résidus des sites d'éditions impliqués dans la reconnaissance des chaînes principales des substrats sont conservés dans les 3 aaRS. Malgré leur faible similarité de séquence globale, le domaine d'édition des IleRS, ValRS et LeuRS est conservé.

Les PheRS, ThrRS, AlaRS et ProRS (classe II) comportent des sites d'éditions plus diversifiés au niveau structural et les mécanismes ne sont pas clairement définis. Il a néanmoins été postulé que les Phe- et Thr-RS partagent un mécanisme commun pour la catalyse de l'édition par l'intermédiaire de l'adenosine terminale de l'ARNt homologue (Ling *et al.*, 2007).

3. Reconnaissance de l'ARNt et rôle des éléments d'identité

La combinaison des données structurales, génétiques et biochimiques a apporté des connaissances conséquentes sur le panel de stratégies employées par les aaRS pour reconnaître spécifiquement un ARNt. Dans de nombreux cas, la liaison des substrats à l'aaRS conduit à des changements de conformations et des mouvements de domaines flexibles. La première structure cristallographique d'un complexe aaRS-ARNt a permis de mettre en évidence les éléments structuraux impliqués dans la spécificité de la réaction. Trois zones du domaine de Rossmann de la GlnRS d'*E. coli*, 2 boucles et une hélice, sont responsables de la spécificité de reconnaissance de la tige acceptrice de l'ARNt. En retour, la présence de l'ARNt aide à structurer le site actif pour déclencher la réaction d'activation de l'acide aminé (Ibba et Söll, 2000; Perona *et al.*, 1993) (Figure I. 4). Nombre d'eucaryotes ont également développé des extensions N- ou C-terminales qui entre autres accroissent l'affinité de l'enzyme pour l'ARNt ou sont impliquées dans les fonctions non canoniques (voir § I.C. 5. et I.D.1).

L'expression correcte du code génétique est gouvernée par des signaux définissant l'identité des ARN de transfert. Celle-ci porte des éléments positifs, appelés déterminants, et négatifs, appelés antidéterminants. Les premiers rendent possible une réaction spécifique, les secondes, empêchent la reconnaissance d'un ARNt par une synthétase non homologue. De manière générale, la structure en "L" de l'ARNt permettrait la présentation correcte de ces éléments à l'enzyme. La détermination de la structure 3D des complexes aaRS-ARNt pour les systèmes glutamine et aspartate (Rould *et al.*, 1989; Ruff *et al.*, 1991) ont d'abord montré que l'interaction avec les synthétases modifie la conformation de l'ARNt au niveau de la boucle anticodon et de l'extrémité acceptrice. La plupart des déterminants d'identité se situent dans ces deux régions de l'ARNt. Les éléments d'identité sont de natures diverses : ils peuvent concerner (i) des groupements chimiques, par exemple des groupements NH₂ des bases G2, G3 et G4 dans le système glutamine, (Hayase *et al.*, 1992), (ii) un nucléotide, dont très souvent N73, (Giegé *et al.*, 1998), (iii) des paires de base (souvent parmi les 5 dernières paires de base de la tige acceptrice), (iv) des bases modifiées comme celle en position 34 de l'anticodon de l'ARNt^{Glu} d'*E. coli*, (Rogers *et al.*, 1995; Gustilo *et al.*, 2007) ou (v) des éléments structuraux tels la boucle D et la région variable de la PheRS (Perret *et al.*, 1992). Un jeu d'identité est complet lorsque sa transplantation dans un ARNt d'une autre spécificité suffit à lui conférer une nouvelle identité. La possibilité de changer l'identité d'un ARNt par de simples mutations ponctuelles apporte également une base pour l'évolution d'ARNt (Frugier *et al.*, 1994).

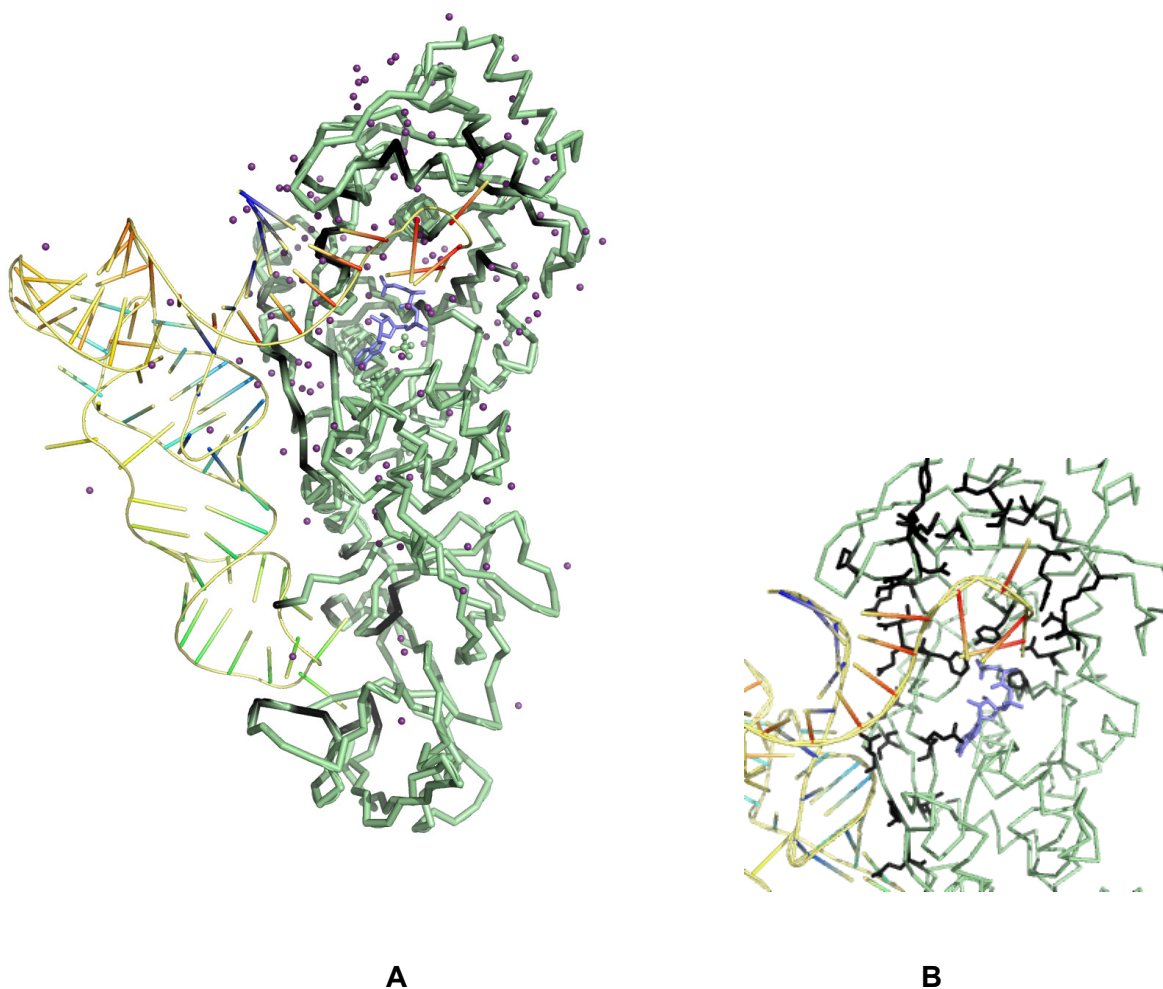


Figure I. 4 : Représentation des régions impliquées dans la discrimination de l'ARNt^{Gln} par la GlnRS d'*E. coli*. Les régions impliquées pour l'ensemble de l'enzyme (A) ou uniquement dans la région du site catalytique (B) sont colorées en noir, l'enzyme en vert pâle et l'ARNt en jaune (pdb-id 1EUQ).

4. La nécessité d'erreurs

On a longtemps pensé qu'il existait une synthétase par acide aminé soit 20 aaRS probablement à cause de l'étroit éventail d'organismes étudiés à savoir *E. coli*, la levure et quelques mammifères. Cette idée a été pour la première fois remise en question à la fin des années 60 (Wilcox, 1969; Wilcox et Nirenberg, 1968) lorsque l'étude du mécanisme de formation du Gln-ARNt^{Gln} a révélé la nécessité de deux enzymes : une GluRS capable de charger de manière non discriminante le glutamate sur l'ARNt^{Gln} (GluRS-ND) pour former la Glu-ARNt^{Gln}, et une amidotransférase responsable du transfert Glu en Gln sur l'ARNt. Le séquençage du génome de l'archée *Methanococcus jannaschi* où manquent trois aaRS (Bult *et al.*, 1996) donna le coup d'envoi à de nombreux travaux. Les études montrèrent qu'il ne s'agit pas de l'exclusivité d'un organisme exotique mais que la plupart des bactéries et des archées manquent la GlnRS, l'AsnRS et parfois la CysRS (Curnow *et al.*, 1997; Curnow *et al.*, 1996; Tumbula *et al.*, 2000). L'absence de certaines aaRS dans la cellule nécessite par conséquent une sorte d' "erreur" de charge d'un ARNt sur une aaRS non discriminante.

Voie indirecte de synthèse de l'Asn-ARNt^{Asn} et la Gln-ARNt^{Gln}

Pour pallier à l'absence de GlnRS et d'AsnRS, la plupart des bactéries et des archées possèdent une GluRS-ND (ND pour non discriminante) et/ou une AspRS-ND. Les amidotransférases sont retrouvées sous 2 formes : la première, GatCAB, est une protéine trimérique retrouvée chez les bactéries et certaines archées et est capable d'amider indistinctement l'Asp-ARNt^{Asn} et la Glu-ARNt^{Gln} (Asp/Glu-AdT); la seconde, GatDE, seulement rencontrée chez les archées, est uniquement responsable de l'amidation du glutamate en glutamine (Glu-AdT) (Tumbula *et al.*, 2000). Dans la plupart des cas, la présence d'une amidotransférase compense l'absence de l'aminacyl-ARNt synthétase correspondante, mais dans certains organismes, les éléments de la voie directe et indirecte d'aminacylation coexistent, comme chez *Thermus thermophilus*, *Deinococcus radiodurans* ou *Bacillus subtilis* (Becker *et al.*, 2000a; Becker *et al.*, 2000b). Chez *D. radiodurans*, par exemple, qui présente une GlnRS, donc utilise la voie directe de glutaminylation, la présence d'une AdT sert uniquement la voie indirecte de synthèse de l'Asn-ARNt^{Asn} (Curnow *et al.*, 1998). Les structures cristallographiques de GatDE et de GatCAB ont permis d'éclaircir le mécanisme et de mettre en évidence l'existence d'un tunnel permettant le passage de l'ion ammonium depuis la poche de l'acide aminé jusqu'à la branche acceptrice de la Glu-ARNt^{Gln} (Feng *et al.*, 2005; Nakamura *et al.*, 2006; Schmitt *et al.*, 2005).

Ces mécanismes impliquant l'aspartate et le glutamate considérés comme les plus anciens acides aminés supportent l'idée que le métabolisme des acides aminés a joué un rôle important

dans l'évolution du système de la synthèse protéique (Wong *et al.*, 1975). Les gènes ancestraux codant la GlnRS et la GluRS ont probablement évolué suite à la duplication d'un gène ancestral conduisant à l'évolution indépendante des GlnRS, GluRS-ND et GluRS eucaryotes et archées (Shulze *et al.*, 2006). Les rares bactéries comme *B. subtilis* ou *D. radiodurans* présentant une GlnRS l'auraient acquis par transfert horizontal de gène (Lamour *et al.*, 1994; Siatecka *et al.*, 1998).

Autre exemple : synthèse de Cys-ARNT^{Cys} chez les archées méthanogènes

La caractérisation de la voie indirecte de cystéinylation de l'ARNT^{Cys} de *Methanocaldococcus jannaschii* a fait l'objet de nombreux tâtonnements mettant en jeu, ici, la SerRS (Kim *et al.*, 1998), là, la ProRS (Stathopoulos *et al.*, 2001), tous réfutés par la suite. La voie de formation du Cys-ARNT^{Cys} a finalement été élucidée récemment : la SepRS, paralogue de la sous-unité α de la PheRS, qui catalyse la charge de O-phosphoserine sur l'ARNT^{Cys}. Dans une seconde étape le Sep-ARNT^{Cys} est reconnu par une SepCysS (Sep-ARNT : Cys-ARNT synthétase) responsable de la conversion en Cys-ARNT^{Cys} (Sauerwald *et al.*, 2005), (Figure I. 5)

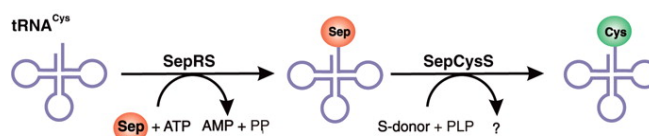


Figure I. 5 : Schéma du mécanisme de formation du Cys-ARNT^{Cys} chez *Methanocaldococcus jannaschii*. La O-phosphoserine (Sep, en orange) chargée sur l'ARNT^{Cys} est prise en charge par la SepCysS qui convertit la Sep en Cys (en vert). Figure d'après (Sauerwald *et al.*, 2005)

5. Divergences et similarités des synthétases : notion de modularité

Les aaRS sont caractérisées par une structure minimale constituée d'un domaine catalytique et d'un domaine de liaison à l'anticodon reliés par un domaine charnière. Comme le montre l'alignement structural des domaines catalytiques, leur architecture est conservée dans chaque classe. Ce domaine pourrait ressembler à l'état ancestral des aaRS au moment de la transition "monde tout ARN" vers un "monde protéines - acides nucléiques". Les sites actifs seraient d'anciens domaines de fixation à l'ATP qui ont évolué pour accepter de manière spécifique l'ARNt en tant que substrat (O'Donoghue et Luthey-Schulten, 2003; Ribas de Pouplana et Schimmel, 2001a) (Figure I. 6).

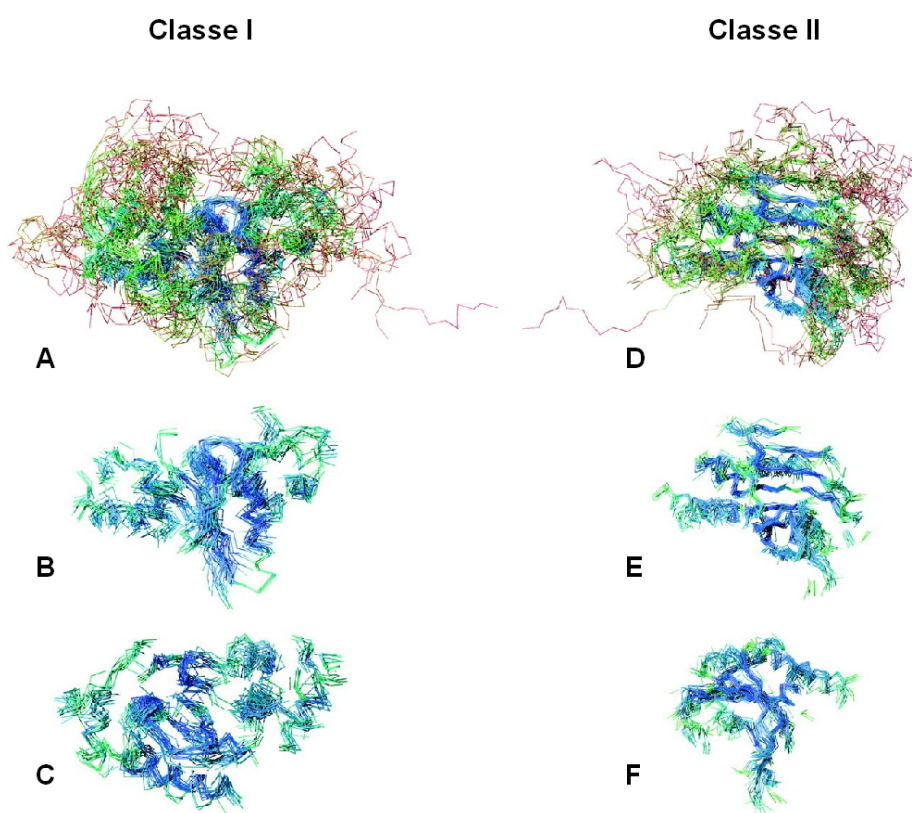


Figure I. 6 : Superposition des domaines catalytiques des classes I et II : vers la définition d'un domaine ancestral? Les structures de tous les domaines catalytiques des aaRS ont été alignées séparément pour les classe I (a à c) et II (dà f), basé sur les alignements structuraux. Les synthétases des deux classes ne sont pas liées structuralement. Le code couleur suit le degré de conservation structurale allant du brun (le moins conservé) au bleu foncé (le plus conservé). Le panneau inférieur de la figure (bc, ef) montre, dans différentes vues, le cœur conservé des domaines catalytiques sans les insertions. Schéma d'après (O'Donoghue et Luthey-Schulten, 2003).

Les domaines de liaisons à l'anticodon des aaRS de classe I sont composés exclusivement d'hélices α ou de feuillets β , ceux des classes II d'une alternance α/β . Le domaine de liaison à l'anticodon DARN, formé d'hélices α , est certainement le plus fréquent chez les aaRS puisqu'il est présent dans 7 aaRS de classe I (Cys-, Lys-, Met-, Val-, Ile-, Leu- et Arg-RS) ainsi que la GlyRS (sous-unité β), seule représentante de la classe II. Parmi les classes II, les domaines structuraux

les plus fréquents sont le domaine "OB-fold" (5β , α), caractéristique de la classe IIb (Asp-, Asn- et Lys-RS2), et un domaine de liaison (5β , 3α) retrouvé dans la classe IIa (Gly-, Thr-, Pro- et His-RS).

L'organisation en 2 domaines des aaRS peut être mise en parallèle à celle des ARN de transfert. Pour l'illustrer, on a par exemple montré que le module N-terminal isolé de la CysRS garde ses capacités d'aminacylation d'une microhélice correspondant à l'extrémité acceptrice de l'ARNt^{Cys} tandis que le module C-terminal garde sa spécificité de liaison à l'ARNt entier. La réaction d'aminacylation n'est néanmoins efficace que lorsque la communication est établie d'un domaine à l'autre c'est-à-dire lorsqu'ils sont liés de manière covalente (Zhang et Hou, 2005). Des témoins de cette modularité se retrouvent au niveau génétique où les domaines fonctionnels des synthétases sont séparés – par ex. : la LeuRS de *A. aeolicus* (Xu *et al.*, 2002) - mais aussi sous formes de domaines d'édition isolés - par ex : Ybak de *H. influenzae* (An et Musier-Forsyth, 2004) – ou encore des domaines d'activation des synthétases qui ont acquis de nouvelles fonctions – par ex. : YadB d'*E. coli* (voir aussi § I.D.2).

Au-delà de ces 2 modules primordiaux, la structure des aaRS est loin d'être consensuelle : on a répertorié 57 modules d'insertions en N- ou C-terminal de fonctions et de structures variées allant de l'édition (Met-, Ile-, Val-, Leu-RS) à la stabilisation de l'ARN de transfert ou des fonctions annexes des synthétases. Les modules additionnels du règne eucaryote ont souvent une fonction de liaison aux acides nucléiques; par exemple, des motifs répétés formant une architecture en "coiled-coil" sont retrouvés en N- ou C-terminal des GluPro-, Gly-, His-, Tyr-, Pro-, Glu- et Met-RS (Cahuzac *et al.*, 2000). Un rôle général d'interaction avec l'ARNt a été attribuée à une séquence riche en lysines de l'extrémité N-terminale de l'AspRS et de la LysRS eucaryotes (Frugier *et al.*, 2000; Francin *et al.*, 2002). L'AspRS de levure est également impliquée dans l'autorégulation de l'expression de l'enzyme par interaction avec son ARN messenger (Frugier *et al.*, 2005; Ryckelynck *et al.*, 2005b). Ces domaines additionnels caractéristiques des eucaryotes ont également un rôle dans l'échafaudage du complexe multisynthétasique (§ I.C. 6).

Le domaine EMAPII stimule lui aussi l'interaction avec l'ARNt. Il est trouvé sous forme d'un domaine chez certaines aaRS bactériennes, par exemple dans l'extension C-terminale de la MetRS bactérienne, alors qu'il est absent dans les formes eucaryotes. Ainsi, chez la levure, ce domaine s'exprime sous forme libre dans la protéine Arc1p et agit en trans pour stimuler l'interaction de la MetRS et de son ARNt homologue (Simader *et al.*, 2006ab; Simos *et al.*, 1998). D'autres domaines additionnels bactériens se sont révélés indispensables à l'efficacité de d'aminacylation de l'ARNt, comme l'extrémité C-terminale de la GlnRS de *D. radiodurans* dont il sera question dans ce travail (voir la partie *Résultats*).

6. Les complexes multisynthétasiques

L'architecture compartimentée et la taille de la cellule eucaryote nécessitent une adaptation fonctionnelle des aaRS reflétée par la présence de domaines additionnels d'interaction avec les protéines et les acides nucléiques (Kisselev et Wolfson, 1994; Mirande, 1991; Yang, 1996). Chez les eucaryotes supérieurs, de la drosophile à l'humain, il existe un complexe multienzymatique constitué de 9 aaRS (Glu-, Pro-, Ile-, Leu-, Met-, Gln-, Arg-, Lys- et Asp-RS) associées à trois protéines non enzymatiques (p43/AIMP1, p38/AIMP2, p18/AIMP3). Pour comprendre le rôle structural de cette organisation complexe, de nombreuses approches ont été mises en oeuvre utilisant des techniques de chromatographie, de microscopie électronique (Han *et al.*, 2003; Wolfe *et al.*, 2003) (Figure I. 7). Les extensions spécifiques des aaRS eucaryotes peuvent être impliquées dans la stabilisation du complexe multisynthétasique, celles de l'ArgRS ou la MetRS par exemple, sans que cela soit une règle générale. Les cofacteurs associés aux aaRS jouent un rôle central dans l'assemblage du complexe, en particulier la présence de p38 y est cruciale tandis que p18 et p43 sont plus spécifiques de l'ArgRS et de la MetRS, respectivement et pourraient être impliqués dans la formation de complexes intermédiaires (Lee *et al.*, 2004; Robinson *et al.*, 2000).

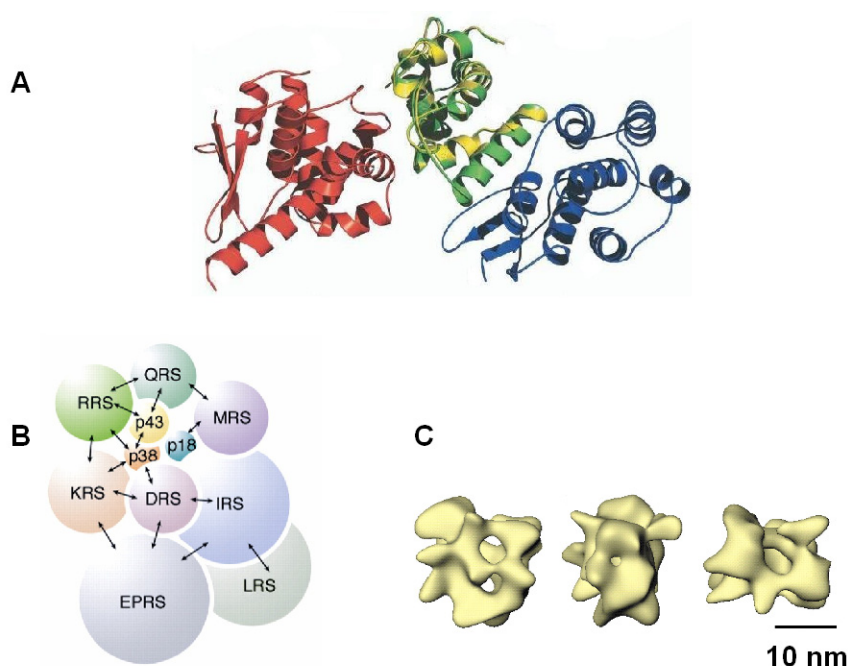


Figure I. 7 : Structures tridimensionnelles et modèles de complexes multisynthétasiques d'eucaryotes inférieurs et supérieurs. **A** - Modèle du complexe ternaire MetRS-N - Arc1P-N - GluRS-N de *Saccharomyces cerevisiae*. Le domaine d'interaction de la MetRS (MetRS-N) est représenté en bleu, de la GluRS (GluRS-N) en rouge et les domaines Arc1p correspondants (MetRS-N - Arc1P-N et Arc1P-N - GluRS-N) en vert et jaune, respectivement. Le domaine d'interaction de la protéine Arc1p exhibe des sites de liaison avec ces aaRS qui lui permettent de les lier simultanément ou indépendamment (Simos *et al.*, 1998). Ce domaine (GST) est souvent retrouvé chez les protéines impliquées dans le complexe multisynthétasique eucaryote. Figure d'après (Simader *et al.*, 2006b). **B** - Schéma d'interaction des éléments du complexe multisynthétasique des eucaryotes supérieurs. Le schéma utilise le code acide aminé à une lettre (Lee *et al.*, 2004). **C** - Structure tridimensionnelle du complexe obtenue par cryomicroscopie électronique (Wolfe *et al.*, 2003) respectivement et de gauche à droite en vue de face, vue de profil (rotation de 90° sur l'axe vertical) ou du dessus (rotation de -90° sur l'axe horizontal).

Le rôle du complexe multisynthétasique est encore flou mais quelques hypothèses ont néanmoins été proposées : (i) en facilitant les multiples interactions protéine-protéine et protéine-ARN, le complexe pourrait former une voie de passage pour le flot d'ARNt et accroître l'efficacité de la synthèse protéique, grâce par exemple, au facteur p43 qui facilite le placement de l'ARNt^{Arg} sur son aaRS correspondante (Park *et al.*; 1999, Park *et al.*, 2005); (ii) l'association des aaRS serait requise pour éviter leur dégradation; (iii) Les aaRS comme les cofacteurs sont souvent impliqués dans des fonctions annexes de la cellule. Pour contrôler l'équilibre dynamique entre leur fonction d'aminacylation et celle de régulation, le complexe servirait de réservoir moléculaire pour la distribution des enzymes et cofacteurs vers leur différentes fonctions (Lee *et al.*, 2004; Sampath *et al.*, 2004).

Si de nombreuses études ont eu pour but de connaître les éléments en interaction avec les aaRS eucaryotes (Kim *et al.*, 2000; Quevillon *et al.*, 1999), peu d'études structurales ont porté leurs fruits. Chez *Saccharomyces cerevisiae* (eucaryote inférieur), un équivalent du complexe multisynthétasique existe sous la forme MetRS-Arc1p-GluRS. Ces 3 éléments protéiques sont nécessaires et suffisants pour produire le complexe trimérique. Le domaine de liaison à l'ARNt de Arc1p est impliqué notamment dans l'efficacité catalytique pour l'ARNt homologue. Les structures des domaines d'interaction en complexe (Arc1p-N – MetRS-N et Arc1p-N – GluRS-N) ont été résolues séparément et un modèle du trimère présenté (Figure I. 7). Ce complexe a été proposé comme un "prototype" du complexe multisynthétasique des eucaryotes supérieurs composé de 5 orthologues du complexe eucaryote inférieur : la MetRS, la GluProRS et les cofacteurs p18, p38, p43 (Simader *et al.*, 2006a; Simos *et al.*, 1998).

Des traces du complexe multisynthétasique ont également été trouvées dans le règne archée : une protéine métabolique de *Methanocaldococcus janaschii*, Mj1338, a été co-purifiée avec la ProRS du même organisme. Mj1338 a une affinité générale pour l'ARNt et présente des capacités d'interaction avec la LysRS et l'AspRS (Lipman *et al.*, 2003).

D Apparition de fonctions nouvelles et évolution

1. Les multiples fonctions des synthétases

En plus de la réaction d'aminacylation, les aaRS ont de multiples fonctions dans la cellule. Certaines fonctions non canoniques découlent directement de l'utilisation des propriétés structurales des aaRS. Dans certains cas, le mimétisme moléculaire des régions de l'ARNm permet son autorégulation traductionnelle : (i) la GlnRS humaine régule sa propre transcription par des interactions de l'extrémité N-terminale avec son propre ARN messenger. En présentant le même jeu d'identité que l'ARNt^{Gln} humain, l'ARNm le mimerait en sa région 5'UTR (Schray *et al.*, 1991; Wang *et al.*, 2000) ;(iii) De même l'AspRS de *S. cerevisiae* agit sur l'expression de son propre gène en fixant la région 5' UTR de son ARN messenger. Celle-ci adopterait une structure secondaire en tige boucle similaire à celle de la boucle anticodon de l'ARNt^{Asp} (Frugier *et al.*, 2005; Ryckelynck *et al.*, 2005a; Ryckelynck *et al.*, 2005b). On trouve également ce type de mimétisme chez la bactérie. La région 5'UTR en amont de la séquence codant pour la MetRS d'*E. coli* prend pour modèle le domaine anticodon de l'ARNt^{Met} (Dardel *et al.*, 1990). La ThrRS du même organisme reconnaît spécifiquement la région de l'opérateur de son ARN messenger qui mime le bras anticodon et la tige acceptrice de l'ARNt^{Thr} et entre en compétition avec la sous-unité 30S du ribosome (Romby *et al.*, 1996; Marzi *et al.*, 2007), ([Tableau I. 8](#)).

Les aaRS sont également connues pour participer à une large variété de fonctions qui peuvent rester dans le domaine de la traduction, où elles jouent un rôle régulateur (ThrRS, GluProRS, MetRS), ou encore intervenir dans la fonction d'épissage (TyrRS, LeuRS), d'inflammation (EPRS), d'angiogenèse (TyrRS, TrpRS) ou d'apoptose (GlnRS) (Park *et al.*, 2005). L'expansion de ces adaptation est en particulier retrouvée pour la LysRS humaine qui exhibe le plus large panel de fonctions allant de l'encapsidation du VIH, à la régulation de la transcription, l'inflammation et la signalisation cellulaire (Park *et al.*, 2005, Barat *et al.*, 1989; Isel *et al.*, 1995). Ces fonctions multiples représentent une stratégie économique de l'évolution mise en place pour maximiser les diversités de fonctions des protéines et utiliser une ressource génétique limitée.

Espèce	aaRS	Localisation	Cible	Activité	Référence
<i>E. coli</i>	ThrRS		5'UTR	Contrôle traductionnel	(Romby et al., 1996)
<i>N. crassa</i>	TyrRS	Mitochondrie	Groupe I intron	Epissage	(Mohr et al., 2001)
<i>S. cerevisiae</i>	LeuRS	Mitochondrie	Groupe I intron	Epissage	(Li et al., 1996)
<i>Levure</i>	AspRS	Cytoplasme	5'UTR	Autorégulation	(Frugier et al., 2000)
<i>H. sapiens</i>	Mini TyrRS	Extracellulaire	Cell.endothéliales	angiogénique	(Wakasugi and Schimmel, 1999)
<i>H. sapiens</i>	Mini TrpRS	Extracellulaire	Cell.endothéliales	Anti-angiogénique	(Otani et al., 2002)
<i>H. sapiens</i>	LysRS	Noyau	MITF ^a	Control transcriptionnel	(Lee et al., 2004b)
<i>H. sapiens</i>	LysRS	Extracellulaire	Macrophage	Inflammation	(Park et al., 2005b)
<i>H. sapiens</i>	LysRS	Membranaire	Gag ^b (VIH)	Assemblage viral	(Halwani et al., 2004)
<i>H. sapiens</i>	GlnRS	Cytoplasme	ASK1 ^c	Anti-apoptose	(Ko et al., 2001)
<i>H. sapiens</i>	MetRS	Noyau	Nucléole	Transcription ARNr	(Ko et al., 2000)

Tableau I. 8 : Quelques exemples des fonctions non canoniques jouées par les aaRS.

^a : MITF : Protéine de liaison à l'ADN

^b : Gag : protéine d'enveloppe du VIH

^c : ASK1 : kinase impliquée dans la régulation du signal apoptotique

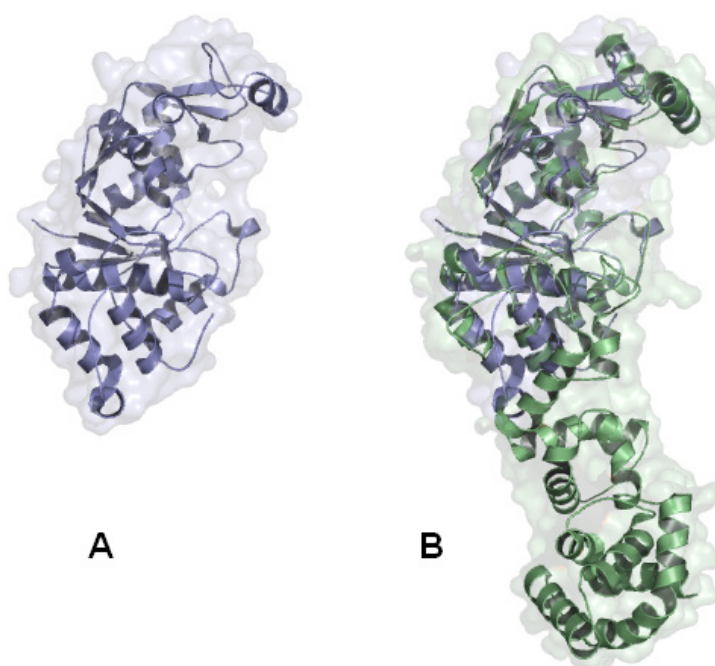


Figure I. 9 : Structures tridimensionnelles du paralogue YadB d'*E. coli* (pdb-ID :1NZJ) seul (A) ou superposé à la GluRS (pdb-ID :1N78) de *Thermus thermophilus* (B). YadB, bien que privé de domaine de liaison à l'anticodon, et, à l'inverse de toutes les GluRS connues active le glutamate en l'absence d'ARNt.

2. Les paralogues

Grâce aux progrès réalisés lors de cette dernière décennie, nombre de génomes de différentes espèces sont disponibles dans les banques de données. L'analyse des homologies de séquences a permis d'isoler des protéines qui dériveraient des modules des aaRS. Un exemple, YadB, présente 34% d'homologie avec le coeur catalytique de la GluRS mais est dépourvue de site de liaison à l'anticodon. Ce « module » solitaire est responsable de la glutaminylation de la queuosine 34 (nucléotide modifiée) de l'anticodon de l'ARNt^{Asp}, nucléotide qui occuperait la place de l'extrémité 3' de l'ARNt. YadB est le premier cas d'aminacylation déviant à la réaction localisée sur la tige acceptrice de l'ARNt (Blaise *et al.*, 2005; Dubois *et al.*, 2004, Salazar *et al.*, 2004), (Figure I. 9)

D'autres paralogues sont par exemple retrouvés (i) chez *E. coli* sous la forme de la protéine HisZ, impliquée dans la biosynthèse de l'histidine et paralogue du domaine d'activation de l'HisRS ou (ii) chez la *Mycobactérie*, par un paralogue de la CysRS nommé MshC (Sareen *et al.*, 2002; Sissler *et al.*, 1999). Enfin, (iii) les protéines AlaX présentent une homologie avec le domaine d'édition des AlaRS. Parmi elles, l'AlaX-M de *P. horikoshii* dont la structure a été résolue est par sa capacité d'hydrolyser les couples Ser-ARNt^{Ala} et Gly-ARNt^{Ala} dotée de propriétés d'édition en trans (Fukunaga *et al.*, 2007).

II. Cristallogénèse et importance des aaRS pour la biologie structurale

A. Aspects généraux : croissance des cristaux de macromolécules

1. La cristallisation

Le principe de la cristallisation d'une macromolécule ne diffère pas particulièrement de celui des petites molécules organiques. Elle nécessite le passage d'un état dispersé en solution à un état ordonné solide. Le cristal est le résultat d'un arrangement régulier périodique d'atomes, de molécules ou d'assemblage moléculaire (Giegé et Ducruix, 1999). La production d'un cristal de macromolécule biologique est un processus complexe et multiparamétrique. Le processus de cristallisation dépend à la fois de facteurs biologiques (stabilité conformationnelle) et de variables physico-chimiques (additifs, force ionique, pH), (voir [Tableau I. 10](#)).

2. Caractéristiques spécifiques des macromolécules

Le premier cristal d'hémoglobine a été reporté par Hünfeldt 100 ans avant que Perutz n'en résolve la structure en 1960 (Perutz *et al.*, 1960). La stabilité intrinsèque des macromolécules biologiques dépend d'une large palette de critères comme le pH et la température, définis pour maintenir leurs propriétés structurales et fonctionnelles (voir [Tableau I. 10](#)). Les macromolécules (i) cristallisent dans 65 groupes d'espaces sur les 230 existants, causé par leurs propriétés de chiralité; les petites molécules préfèrent cristalliser dans des groupes qui remplissent l'espace alors que les macromolécules favorisent leur connectivité et seule une petite partie de leur surface est impliquée dans les contacts intermoléculaires (Carugo et Argos, 1997) (ii) ont de faibles propriétés biréfringentes, (iii) sont fragiles de part leur organisation maintenue par de faibles liaisons intermoléculaires (Van der Waals, liaisons ioniques et hydrogène, interaction dipôle-dipôle...) et leur richesse en solvant. La différence majeure est liée à la flexibilité conformationnelle des macromolécules et leur plus forte sensibilité structurale aux conditions extérieures.

Paramètres généraux	Paramètres spécifiques
- pH - Température, Pression - Agent cristallisant, force ionique, pureté des réactants - Diffusion et convection (microgravité, gels...) - Champs électrique et magnétique, vibration, ultrasons	- Additifs (substrats, cofacteurs, ions métalliques) - Hétérogénéité conformationnelle (domaines flexibles, conformères...) - Pureté (contaminants, hétérogénéité de séquence et de modifications post-transcriptionnelles) - Stabilité en solution (dégradation, vieillissement, dénaturation) - Source biologique (organismes thermophiles³ ou mésophiles⁴)

Tableau I. 10 : Résumé des paramètres généraux et spécifiques agissant sur la cristallisation des macromolécules biologiques (d'après (Ducruix et Giegé, 1999a).

3 Thermophile : organisme extrémophile vit à de hautes températures (supérieure à 45°C) et représenté dans les 3 règnes du vivant (ex: *Bacillus stearothermophilus* (procaryote), *Algues Thermophiles* (eucaryote), *Pyrococcus abyssi* (archée)).

4 Mésophile : organisme qui croit à une température modérée, entre 25 e t 40°C

3. Méthodes de cristallisation et cristallogénèse

La cristallisation des protéines nécessite des techniques en solution adaptées à leur sensibilité et compatibles avec la quantité de molécules accessibles. Les méthodes de cristallisation les plus utilisées sont la diffusion de vapeur, la microdialyse et la cristallisation en batch, la plus populaire restant sans doute la méthode de diffusion de vapeur (voir aussi le chapitre *Matériel et Méthodes*).

Les difficultés rencontrées pour obtenir des cristaux de qualité ont encouragé la communauté à l'étude de la cristallogénèse et élargir ces méthodes à d'autres environnements. La cristallogénèse, qui suit la naissance des premiers noyaux, est la phase durant laquelle se déroule la croissance cristalline dont dépend la taille et la qualité du cristal. Par exemple, on a montré l'influence positive de la microgravité lors du processus de cristallisation. En revanche, les expérimentations au sol sont caractérisées par des mouvements de convection dans la goutte, qui favorisent l'incorporation d'impuretés lors de la croissance cristalline, et par la sédimentation du cristal, qui induit une croissance non isotrope. Ainsi, des missions de cristallisation dans l'espace (sous microgravité) ont permis d'en réduire les effets et d'obtenir de cristaux tridimensionnels de meilleure qualité (Lorber *et al.*, 1999; Ramachandran et Leslie, 2005; Vergara *et al.*, 2003). Les résultats encourageants obtenus lors de ces missions ont conduit à des applications « sur Terre » qui miment l'effet de la microgravité comme la cristallisation en capillaire, en champ magnétique, en microfluidique ou en gel. Cette dernière, de faible coût, a fait ses preuves pour augmenter la qualité du cristal entre autres en ce qui concerne l'amélioration de la mosaïcité et des limites de résolution (Ng *et al.*, 1997; Lorber *et al.*, 1999; Biertumpfel *et al.*, 2002).

4. Qualité des produits

La qualité d'un cristal est corrélée à la régularité de l'empilement des molécules, sa morphologie externe étant liée à sa structure interne. La présence d'impuretés (contaminants, conformères), interfère avec la nucléation, perturbe l'ordre du cristal, diminue sa stabilité, provoque des interactions non spécifiques, d'où l'importance de la pureté et de l'homogénéité du matériel biologique comme des agents cristallins qui sont utilisés à forte concentration. La technique de diffusion de lumière (DLS ou Dynamic Light Scattering) est un bon outil pour contrôler l'homogénéité d'une solution (Lorber *et al.*, 1993). Les techniques d'AFM permettent de visualiser la cartographie de la surface cristalline et donc d'observer le processus de cristallogénèse (Ng *et al.*, 1997a).

5. Recherche de conditions

Les difficultés rencontrées lors de la recherche de conditions de cristallisation encouragent la mise au point de nouvelles stratégies et techniques pour augmenter les chances de succès. La première fut la technique de crible factorielle dont le principe est de couvrir efficacement l'ensemble de l'espace expérimental (Carter et Carter Jr, 1979). L'exploitation de l'information liée au nombre exponentiel de structures cristallographiques disponibles dans la PDB et à la collection des conditions de cristallisation sur des banques de données comme la BMCD ou plus récemment la MPCD (Biological Macromolecule Crystallization Database <http://xpdb.nist.gov:8060/BMCD/>, Gillilan *et al.*, 1988; [Marseille Protein Crystallization Database http://www.crmcn.univ-mrs.fr/mpcd](http://www.crmcn.univ-mrs.fr/mpcd)) a conduit à la commercialisation des cribles. Combinée au développement de robots de cristallisation de plus en plus perfectionnés, cette approche a considérablement facilité la tâche du cristallographe (Figure I. 11). Néanmoins, l'obtention de cristaux d'une qualité suffisante pour être exploités par diffraction aux rayons X nécessite souvent l'optimisation des conditions (revue dans (Kundrot, 2004). En cela, si ces techniques apportent un gain d'échelle et de productivité incontestables, elles restent d'un coût élevé et peuvent requérir un matériel biologique conséquent avant d'aboutir à une structure tridimensionnelle.

Afin de limiter la quantité de matériel nécessaire au criblage de conditions de cristallisation, d'autres techniques de cribles de cristallisation ont été développées. Elles reposent sur une variation limitée des paramètres de cristallisation. Par exemple, A. McPherson a proposé de combiner une base d'agent précipitant (30% PEG 3350 ou 50% TacsimateTM) à un large crible de petites molécules d'additifs (composés polyvalents chargés, ligands, composés phosphatés...etc...) ((McPherson et Cudney, 2006). Une autre approche utilise la contre diffusion en capillaire (Capillary Counterdiffusion Kits[®]) avec une combinaison de PEGs de tailles différentes dans le même réservoir (PEG 400-8000 MW) et avec le pH pour variable unique (Garcia-Ruiz, 2003, www.trianatech.com). Une autre stratégie est en développement au laboratoire, entre autres pour réduire le nombre d'étapes conduisant à la résolution d'une structure 3D : l'élaboration de puces microfluidiques est dédiée à la recherche et à l'optimisation de conditions de cristallisation et doit permettre leur observation et leur analyse par diffraction des rayons X *in situ*.

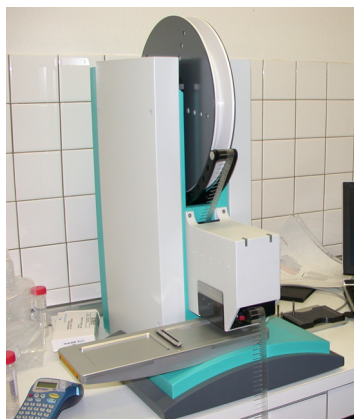


Figure I. 11 : Exemple de matériel utilisé pour le criblage des conditions de cristallisation. Robot de cristallisation type Mosquito®.

6. Les modèles utilisés en cristallogénèse

Les modèles couramment utilisés pour mettre en application ce type d'étude sont le plus souvent le lysozyme, la thaumatine ou les particules virales sphériques dont les conditions de cristallisation sont bien établies (Lorber *et al.*, 1999; Ng *et al.*, 1997b). Ces petites protéines ont accompagné les débuts de la cristallogénèse et sont encore largement utilisées, mais sont représentatives des protéines plus complexes qui font l'attrait de la communauté (protéines modulaires, complexes multiprotéiques). De ce fait, les aaRS ont, pour de nombreux points, contribué à l'évolution de la science de la cristallogénèse. Les multiples tentatives de cristallisation de ces molécules sont à l'origine de techniques de cristallisation répandues.

B. Importance des aaRS en cristallogénèse et en biologie structurale

1. Les aaRS dans l'histoire de la cristallisation des macromolécules biologiques

Les chercheurs du domaine ont utilisé l'« outil » aaRS pour développer des techniques plus appropriées à la réalité quotidienne du cristallographe. Ainsi, c'est à la fin des années 60 qu'émerge la technique de diffusion de vapeur en goutte suspendue, suivie un peu plus tard de la microdialyse (Hampel *et al.*, 1968). En partant du pré requis que la nature a développé des structures stables pour permettre à certaines espèces de survivre en conditions extrêmes, les chercheurs se sont intéressés à des organismes thermostables ce qui a certainement facilité la résolution de la structure 3D de la TyrRS de *Bacillus stearothermophilus* (Brick et Blow, 1987). L'étude de la cristallogénèse de la TrpRS du même organisme et de l'AspRS de *S. cerevisiae* ont permis respectivement d'élaborer des techniques pour la recherche de conditions de cristallisation (Carter et Carter Jr, 1979), ou de mettre en évidence l'importance de la pureté et de l'homogénéité structurale d'une protéine (Lorber *et al.*, 1987). L'AspRS de levure a également fait l'objet d'investigations pour montrer l'influence du champ magnétique terrestre sur la qualité du processus de cristallisation depuis les expériences de cristallisation dans l'espace, en capillaire, et l'utilisation du gel d'agarose. Des cristaux de quelques aaRS obtenus sur puces microfluidiques permettent également au laboratoire de tester, en plus des protéines classiques (type, lysozyme ou thaumatine), la cristallisabilité de protéines plus complexes dans les microcanaux de ces puces. Ces tests apportent un degré supérieur de difficulté et s'adaptent plus justement à la réalité quotidienne du cristallographe.

Si les aaRS sont des acteurs importants de ce domaine, les ARN de transfert y ont également une place de choix puisqu'ils ont été les premiers acides nucléiques cristallisés dont la structure a été résolue à l'état libre, comme l'ARNt^{Phe} de levure (Kim *et al.* 1974, Robertus *et al.*, 1974, Stout *et al.*, 1976), l'ARNt^{Asp} de levure (Giegé *et al.*, 1977; Moras *et al.*, 1980, Westhof *et al.*, 1983), l'ARNt^{fMet} initiateur d'*E. coli* et de levure (Basavappa et Sigler, 1991; Woo *et al.*, 1980). De plus, les premières observations des mécanismes de croissance d'un cristal d'acide nucléique par la technique d'AFM ont été faites à partir d'un cristal d'ARNt^{Phe} de levure (Ng. *et al.*, 1997).

2. Modèles en ingénierie

Les aaRS sont également les précurseurs de l'ingénierie génétique : la digestion par protéolyse ménagée de la MetRS de *E. coli* a conduit au premier état de diffraction d'une aaRS (Waller *et al.*, 1971). Ce principe est encore très usité notamment face aux difficultés rencontrées pour la cristallisation des protéines eucaryotes à nombreux domaines flexibles.

La TyrRS de *B. stearothermophilus* dont la structure cristallographique avait été résolue à haute résolution a servi d'outil pour le développement de ces techniques. L'identification des résidus impliqués dans l'interaction avec ses substrats ont aidé le développement des premières techniques de mutagenèse ponctuelle *in vitro* (Wilkinson *et al.*, 1983; Bedouelle *et al.*, 1985). Pour déterminer la contribution d'une région C-terminale désordonnée de la TyrRS, une technique de délétion *in vitro* par l'intermédiaire du bactériophage M13 a été développée. On a ainsi montré que l'extrémité C-terminale était impliquée dans la reconnaissance de l'ARNt^{Tyr} (Waye *et al.*, 1983; Winter *et al.*, 1982).

C. Vision globale de la cristallographie des aaRS et contribution à leur connaissance générale

Ce paragraphe donne une vision globale des études de cristallogénèse des aaRS qui ont conduit à une résolution de structure par cristallographie, complété par les structures de domaines résolues par RMN. Nous nous intéresserons aux structures des aaRS résolues et déposées dans la PDB. Il faut néanmoins prendre en compte le fait que les structures résolues avant les années 80 n'étaient pas automatiquement déposées dans la Protein Data Bank. Ce bilan donne un point de vue sur les similitudes entre les règnes du vivant, entre les classes, et plus généralement sur l'élargissement des études à des organismes "exotiques" ([Tableau I. 12](#)).

L'étendue des connaissances structurales accumulées justifie une étude généralisée sur ces aspects des aminoacyl-ARNt synthétases. À ce jour, les structures des 20 types d'aaRS sont résolues, ainsi que récemment la structure de la PylRS d'archée (Kavran *et al.*, 2007). Elles couvrent quelquefois les 3 règnes du vivant et sont très largement étudiées (ex : MetRS, AspRS), d'autres fois, le travail n'est qu'esquissé et de nombreuses facettes restent inexplorées (ex : AlaRS, AsnRS).

Il existe souvent une longue transition entre le cristal initial et la structure finale d'une macromolécule et la littérature regorge de cas de protéines cristallisées mais qui n'ont pas conduit à des structures tridimensionnelles. Pour les aaRS par exemple, la LysRS de levure, cristallisée en 1970 (Rymo et Lagerkvist, 1970), n'a pas conduit à la résolution d'une structure tridimensionnelle. Mais ces échecs ont souvent servi de leçon pour la compréhension des paramètres importants dans la cristallisation des aaRS. On pense par exemple à la découverte des propriétés de maintien des interactions ARNt-aaRS à forte concentration en sulfate d'ammonium qui a permis par la suite de résoudre de nombreuses structures de ce type (Leberman *et al.*, 1982; Giegé *et al.*, 1980; Lorber *et al.*, 1983).

Tableau récapitulatif des structures résolues et déposées dans la Protein Data Bank

CLASSE I

CLASSE I

Classe Ia												
aaRS	n	Espèce	Règne	PDB-ID	Date de dépôt	ARNt	ligand-id	Résolution (Å)	Groupe d'espace	Méthode de cristallisation	Références	
LeuRS	12	<i>Pyrococcus horikoshii</i>	A	1WKB	05/04		SO4	2,0	R 3	DV	Fukunaga et al., 2005	
			1WZ2	02/05	X		3,2	P 2, 2, 2 ₁	DV			
		<i>Escherichia coli</i>	B	2AJG	08/05			2,0		DV	Liu et al., 2005	
				2AJH	08/05		Met	2,4	P 6, 2 2	DV		
		2AJI		08/05		Ile	3,2		DV			
		<i>Thermus thermophilus</i>		1H3N	09/02		Zn	2,0	C 2 2 2 ₁		Cusack et al., 2000	
				1OB C	09/02		Methionyl-adenylate ^{an} Zn	2,0	C 2 2 2 ₁		Linccum et al., 2003	
				1OBH	01/03		Norvaline, Zn	2,2	C 2 2 2 ₁			
				2BTE	05/05		Norvaline, Zn	2,9	C 2		Tukalo et al., 2005	
				2BYT	08/05		Zn	3,2	C 2			
				2VOC	05/07		Antifungique	1,8	C 2 2 2 ₁		Rock et al., 2007	
				2V0G	05/07		Zn	3,5	C 2			
IleRS	11			<i>Thermus thermophilus</i>	B	1ILE	02/98		Zn	2,5	C 2	DV
				1QU3		07/99		Mupirocine Zn	2,9		E	Silvian et al., 1999 et 2000
		<i>Staphylococcus aureus</i>		1QU2		07/99	X	Mupirocine Mg Zn K	2,2	P 2, 2, 2 ₁	E	
				1FFY		07/00		Mupirocine Mg Zn K	2,2		DV	Nakama et al., 2001
				1JZS		09/01		Mupirocine Zn	2,5		DV	
				1JZQ		09/01		Isoleucyl-adenylateTyrosine ^{an} Zn	3,0	C 2	DV	
		<i>Thermus thermophilus</i>		1UE0		05/03		Val	2,0	P 4, 2, 2	DV	Fukunaga et al., 2004 et 2006
				1UDZ		05/03			1,8	P 4, 2, 2	DV	
				1WK8		05/04		Valyl-adenylate ^{an}	1,7	P 4, 2, 2	DV	
				1WNY		05/04			1,6	P 4, 2, 2	DV	
				1WNZ		05/04		Valyl-adenylate ^{an}	1,7	P 4, 2, 2	DV	
ValRS	4	<i>Thermus thermophilus</i>		B	1GAX	06/00	X	Valyl-adenylate ^{an} Zn	2,9	P 4, 2, 2	DV	Fukai et al., 2000 et 2003
		<i>Thermus aquaticus</i>			1IVS	03/02		Valyl-adenylate ^{an}	2,9		DV	
					1IYW	03/02			4,0	P 2, 2, 2	DV	
		<i>Thermus thermophilus</i>		B	1WK9	05/04		Threonyl-adenylate ^{an}	1,8	P 6, 2 2	DV	Fukunaga et al., 2005
ArgRS	4	<i>Thermus thermophilus</i>	B	1WKA	05/04			1,7				
		<i>Thermus thermophilus</i>	B	1IQO	11/01			2,3	P 6		Shimada et al., 2001	
			1BS2	08/98			2,8	P 4, 2, 2		Cavarelli et al., 1998		
		<i>Saccharomyces cerevisiae</i>	E	1F7U	06/00	X	SO4	2,2	P 2, 2, 2 ₁	DV	Delagoutte et al., 2000	
	1F7V	06/00			2,9	I 2 2 2	DV					
CysRS	3	<i>Escherichia coli</i>	B	1LIS	04/02		ZN	2,3	P 4, 2, 2	DV	Newberry et al., 2002	
			1L17	04/02		ZN	2,6		DV			
			1U0B	07/04	X	ZN	2,3		DV	Hauenstein et al., 2004		
MetRS	22	<i>Escherichia coli</i>	B	1MED	11/92		ZN				Fourmy et al., 1993	
		1MEA		11/92		ZN						
		<i>Thermus aquaticus</i>		1A8H	03/98		ZN	2,0	P 2, 2, 2 ₁		Sugiura et al., 1998	
				1QQT	06/99		ZN	2,0		DV	Mechulam et al., 1999	
				1F4L	06/00		ZN	1,9		DV	Serre et al., 2001	
				1P7P	05/03		Met ^{an} Zn	1,8		DV		
				1PFY	05/03		Methionyl-adenylate ^{an} Zn	1,9		DV	Crepin et al., 2003	
				1PFU	05/03		Met ^{an} Zn	1,9	P 2 ₁	DV		
				1PG0	05/03		Methionyl-adenylate ^{an}	1,9		DV		
				1PG2	05/03		Adenosine	1,8		DV		
				1PFW	05/03		Met ^{an} Zn	1,8		DV		
				1PFV	05/03		Met ^{an} Zn	1,7		DV		
		<i>Thermus thermophilus</i>		1WOY	08/04		ZN	2,0		DV		Konno et al., 2005
				2D54	10/05		ZN	2,0		DV		
				2D5B	10/05		ZN	1,8	P 2, 2, 2 ₁	DV		
				<i>Aquifex aeolicus</i>	2CSX	05/05			2,7		DV	Nakanishi et al., 2005
		2CT8			05/05	X	Methionyl-adenylate ^{an}	2,7		DV		
		<i>Pyrococcus abyssi</i>			A	1MKH	08/02			2,0	P 3, 2 1	DV
					1RQG	12/03		ZN	2,9	I 4, 2 2	DV	Crepin et al., 2004
		<i>Pyrococcus horikoshii</i>			2CWP	06/05			2,1	P 2, 2, 2 ₁	DV	Murayama et al., 2005
		<i>Homo sapiens</i>		E	2DJV	04/06						Sato et al., 2006
		<i>Saccharomyces cerevisiae</i>			2HSN	07/06			2,2	P 4, 2, 2		Simader et al., 2006

Tableau I. 12 : Structures des aaRS déposées dans le Protein Data Bank⁵ : Les aaRS sont groupées par classes et sous classes. La colonne "n" donne le nombre des aaRS résolues par type, la colonne "Règne" classe les aaRS en fonction de leur appartenance au règne bactérien (B), archée (A) ou eucaryote (E). Les dates correspondent à la date de dépôt de la structure. Les structures résolues en complexe avec l'ARNt sont annotées d'un "X" et, s'il y a lieu, la nature des autres ligands est précisée dans la colonne "ligands". La limite de résolution le groupe d'espace et la technique de cristallisation sont précisées. Les structures dont les codes PDB sont surlignés en cyan correspondent à des structures résolues par Résonance magnétique nucléaire.

Annotations :

- l'indice 'an' pour analogue (de l'acide amine, de l'ATP, de l'aa-adenylate)
- Xa1 : ARNt^{Asp} d' *E. coli*
- Xa2 : ARNt^{Asp} *S. cerevisiae*
- NP : non publié / ND : non déposée sur la PDB

classe Ib

aaRS	n	Espèce	Règne	PDB-ID	Date de dépôt	ARNt	ligand-id	Résolution (Å)	Groupe d'espace	Méthode de cristallisation	Références
GluRS	18	<i>Thermus thermophilus</i>	B	1GLN	07/94			2,5			Nureki et al., 1995
				1G59	10/00	X		2,4	C 2 2 2 ₁	DV	Sekine et al., 2001
				1J09	11/02		Glu ATP Mg	1,8	P 2 ₁ 2 ₁ 2 ₁		Sekine et al., 2003
				1N75	11/02		ATP Mg	1,9			
				1N77	11/02		ATP Mg	2,4			
		1N78		11/02	X	Glutamyl-adenylate ^{Asp} Mg	2,1	C 2 2 2 ₁	DV		
		2CUZ		05/05		Glu	2,0	P 2 ₁ 2 ₁ 2 ₁	DV	Sekine et al., 2006	
		2CV0		05/05		Glu MG	2,4		DV		
		2CV1		05/05	X	Glutamate ^{Asp} ATP MG	2,4	C 2 2 2 ₁	DV		
		2CV2		05/05		Glutamyl-adenylate ^{Asp} CL MG	2,7		DV		
		2DXI		08/06		Glu ^{Asp} ATP MG	2,2		DV		
		<i>Synechococcus elongatus</i>		2CFO	02/06		Glu	2,5	P 2 ₁		Schulze et al., 2006
		<i>Thermotoga maritima</i>		2O5R	12/06		GOL CL	2,3	P 3 ₁ 2 ₁ 1	DV	JCSG, Wilson, I.A, 2006
		<i>Saccharomyces cerevisiae</i>		2HRA			IOD	1,9	C 2 2 2 ₁	DV	Simader et al., 2006 et 2007
				2HQT				1,9	C 2	DV	
2HRK				CL	2,0	P 21	DV				
2HSM					3,0						
2HSN					2,2	P 4 ₁ 2 ₁ 2 ₁					
GlnRS	18	<i>Escherichia coli</i>	B	1GS G	04/90			2,8			Rould et al., 1989
				1GTS	09/93		AMP	2,8			Perona et al., 1995
				1QRS	06/96		ATP	2,6			Arnez et al., 1996
				1QRU	06/96		ATP	3,0			
				1QRT	06/96		ATP	2,7			
				1QTQ	01/98	X	Glutamyl-adenylate ^{Asp} SO4	2,3	C 2 2 2 ₁	DV	Rath et al., 1998
				1EUQ	04/00		Glutamyl-adenylate ^{Asp}	3,1		DV	Sherlin et al., 2000
				1EUY	04/00		Glutamyl-adenylate ^{Asp}	2,6		DV	
				1EXD	05/00		SO4	2,7		DV	Bullock et al., 2000
				1NYL	02/03			2,6	P 2 ₁ 2 ₁ 2 ₁		Sherlin et al., 2003
		1O0B		02/03		Gln AMP SO4	2,7		DV	Bullock et al., 2003	
		1O0C		02/03	X	Glu AMP SO4	2,7	C 2 2 2 ₁	DV		
		1ZJW		05/05		Gln AMP SO4	2,5		DV	Gruic-Sovulj et al., 2005	
		<i>Deinococcus radiodurans</i>		2H27	08/06			2,3	P 2 ₁ 2 ₁ 2 ₁	Microbatch	Deniziak et al., 2006
				ND	01/07		QSI	2,2	P 2 ₁ 2 ₁ 2 ₁	Batch	Ce travail
				ND	01/07		GOL	2,4	P 2 ₁ 2 ₁ 2 ₁	Batch	Ce travail
				ND	01/07		Gln	2,5	P 2 ₁ 2 ₁ 2 ₁	Batch	Ce travail
				ND	01/07			2,0	P 2 ₁ 2 ₁ 2 ₁	Batch	Ce travail
LysRS1	1		<i>Pyrococcus horikoshii</i>	A	1IRX	10/01	ZN	2,6	P 2 ₁	DV	Terada et al., 2002

Annotations :

- l'indice 'an' pour analogue (de l'acide amine, de l'ATP, de l'aa-adenylate)
- Xa1 : ARNt^{Asp} d' E. coli
- Xa2 : ARNt^{Asp} S. cerevisiae
- NP : non publié
- DV : diffusion de vapeur, E : évaporation, MD : microdialyse, MB : microbatch.

Classe Ic										
TyrRS	36	<i>Bacillus Stearothermophilus</i>	-	07/76 ^{SP}			2,7	P 3, 2 1		Irwin et al., 1976
			4TS1	06/89			2,5			Brick et al., 1987
			2TS1	06/89			2,3			
			3TS1	06/89		Tyrosine ^{an}	2,7			Brick et al., 1989
			1TYC	07/92			2,5			
			1TYB	07/92			2,5			Brown et al., NP
			1TYA	07/92			2,8			
			1TYD	07/92			2,5			Brick et al., 1989
			1H3	06/01						Guizarro et al., 2001
			1JLJ	07/01		Tyrosine ^{an}	3,2	C 2 2 2 ₁	DV	Qiu et al., 2001
		<i>Staphylococcus aureus</i>	1JIL	07/01		Tyrosine ^{an}	2,2	I 2, 2, 2 ₁		
			1JIK	07/01		Tyrosine ^{an}	2,8			
			1JII	07/01		Tyrosine ^{an}	3,2	C 2 2 2 ₁		
		<i>Thermus thermophilus</i>	1H3F	08/02		Tyrosine ^{an} SO ₄	2,0	P 2, 2, 2 ₁		Yaremchuk et al., 2002
			1H3E	08/02	X	Tyrosine ^{an} ATP	2,9	P 3, 2 1		
		<i>Escherichia coli</i>	1VBM	02/04		Tyrosyl-adenylate ^{an} SO ₄	2,7	P 3, 2 1	DV	Kobayashi et al., 2005a,b
			1VBN	02/04		Tyrosine-adenylate ^{an}				
			1X8X	08/04						
			1WQ3	09/04		Tyrosine ^{an}		P 3, 2 1		
			1WQ4	09/04			2,0			
		<i>Methanocaldococcus jannaschii</i>	1J1U	08/04	X	Tyrosine ^{an}				
			1U7D	08/04			2,7	P 2, 2, 2 ₁	DL	Zhang et al., 2005
			1U7X	08/04			3,0		DL	
			1ZHO	04/05		Ala ^{an}	1,9			Turner et al., 2005 et 2006
			1ZH6	04/05		Phe ^{an}	2,5	P 4, 2, 2		
			2CYA	07/05			2,2			Kuratani et al., 2005
			2CYB	07/05			1,8	P 21		
			2CYC	07/05			2,2	P 2, 2, 2 ₁		
			2AG6	07/05		Phe ^{an}	1,9			Turner et al., 2006
			2HGZ	06/06		Phe ^{an}	2,5	P 4, 2, 2		Liu et al NP
		<i>Homo sapiens</i>	1U7D	08/04			2,7	P 2, 2, 2 ₁	DL	Zhang et al., 2005
			1N3L	10/02			1,2	P 2, 2, 2		Yang et al., 2002 et 2003
			1NTG	01/03			2,2	P 1		
			1Q11	07/03		Tyrosyl-adenylate ^{an}	1,6	P 2, 2, 2 ₁		
			1Y42	11/04			2,0	C 2		Paukstelis et al., 2005
			2DLC	04/06	X	Tyrosyl-adenylate ^{an} Mg	2,4	P 4, 2, 2		Tsunoda et al NP
			2PID ^{PROC}	07/07			2,2	P 2, 2, 2 ₁		Bonnefond et al., 2007
		<i>Neurospora crassa</i>	1D2R	09/99	X		2,9	C 2		Ilyin et al., 1999
			1I6L	03/01		Tryptophanyl-adenylate ^{an}	1,7			Retaillieu et al., 2001
			1I6M	03/01			1,7			
			1I6K	03/01			1,7	P 4, 2, 2		
			1M83	07/02		ATP	2,2		MD	Retaillieu et al., 2003
			1MAU	08/02		ATP, Tryptophane ^{an}	2,2		MD	
			1MAW	08/02		ATP	2,2		MD	
			1MB2	08/02			2,7		MD	Buddha et al., 2005,
			1Y18	01/05			2,1		DV	
			1Y1A	01/05		Tryptophane ^{an}	3,7		DV	
TrpRS	15	<i>Bacillus Stearothermophilus</i>	1Y1D	01/05		ATP	2,4		DV	JCSG, NP
			2A4M	06/05			2,3		DV	
			2G36	02/06			2,5	C 2 2 2 ₁	DV	
			2OV4	02/07		ATP ^{an}	2,5	P 4, 2, 2	MD	Retaillieu et al., 2007
			1ULH	09/03			2,3	C 2		Kise et al., 2004
			1OST	10/03			2,5	P 6, 2 2		Yu et al., 2004
			1R6U	10/03		Tryptophanyl-adenylate ^{an}	2,0	C 2		Yang et al., 2003
			1R6T	10/03		Tryptophanyl-adenylate ^{an}	2,1	C 2		
			2AZX	09/05		MG SO4	2,8	I 2, 2, 2 ₁		
		<i>Bos taurus</i>	2AKE	08/05		SO4	3,1	P 4, 2 2		Shen et al., 2006
			2DR2	06/06		SO4	3,0			
		<i>Saccharomyces cerevisiae</i>	2IP1	10/06			1,8	P 4, 2, 2	MB	Malkowski et al., 2007
		<i>Dainococcus radiodurans</i>	1Y18	01/05			2,1		DV	Buddha et al., 2005,
			1Y1A	01/05		Tryptophane ^{an}	3,7		DV	
			2A4M	06/05			2,3		DV	
			2G36	02/06			2,5	C 2 2 2 ₁	DV	
			2OV4	02/07		ATP ^{an}	2,5	P 4, 2, 2	MD	
		<i>Thermotoga maritima</i>	1ULH	09/03			2,3	C 2		Yang et al., 2006
			1OST	10/03			2,5	P 6, 2 2		
			1R6U	10/03		Tryptophanyl-adenylate ^{an}	2,0	C 2		
			1R6T	10/03		Tryptophanyl-adenylate ^{an}	2,1	C 2		
			2AZX	09/05		MG SO4	2,8	I 2, 2, 2 ₁		

Annotations :

- l'indice 'an' pour analogue (de l'acide amine, de l'ATP, de l'aa-adenylate)
- Xa1 : ARNt^{Asp} d' E. coli
- Xa2 : ARNt^{Asp} S. cerevisiae
- NP : non publié
- DV : diffusion de vapeur, E : évaporation, MD : microdialyse, MB : microbatch.
- Les cases colorées en cyan correspondent aux structures résolues par RMN

CLASSE II

Classe IIa												
aaRS	n	Espèce	Règne	PDB-ID	Date de dépôt	ARNt	ligand-id	Résolution (Å)	Groupe d'espace	Méthode de cristallisation	Références	
SerRS	17	<i>Escherichia coli</i>	B	nd	01/90 ^{OP}			2,5			Cusack et al., 1990	
		1SRV		08/93			2,5	P ₂		Fujinaga et al., 1993		
		1SES		02/94		Seryl-adenylate ²⁺ AMP	2,5	P ₂		Belrhali et al., 1994		
		1SET				Seryl-adenylate ²⁺	2,5	P ₂				
		1SER			X		2,9			Biou et al., 1994		
		1FYF		09/00		Seryl-adenylate ²⁺ Zn	1,7	P ₂ 2 ₁ 2 ₁	DV	Sankaranarayanan et al., 2000		
		1TKG		06/04		Seryl-adenylate ²⁺	1,5	P ₂	DV	Dock-Bregeon et al., 2004		
		1TKY		06/04		Serine ²⁺	1,5	P ₂ 2 ₁ 2 ₁	DV			
		2DQ3		05/06		Seryl-adenylate ²⁺	3,0	P ₄	DV	Itoh et al.NP		
		2CIM		06/06		Zn, Pt	2,5			Bilokapic et al., 2006		
		2CJ9	06/06		Seryl-adenylate ²⁺ Zn	2,3	P ₃ 2 ₁					
		2CJA	06/06		ATP, Zn	2,2						
		2CJB	06/06		Zn	2,7						
		2HL2	08/06		Seryl-adenylate ²⁺	2,6			Hussain et al., 2006			
2DQ0	05/06		Seryl-adenylate ²⁺ SO ₄	2,6	P ₂ 2 ₁ 2 ₁	DV	Itoh et al.NP					
2DQ1	05/06		AMP SO ₄	2,8								
2DQ2	05/06			3,0								
2DQ2	05/06											
<i>Bostaurus</i>	E	1WLE	06/04		Seryl-adenylate ²⁺	1,6	C 2 2 2			Chimmaronk et al., 2005		
ThrRS	17	<i>Escherichia coli</i>	B	1QF6	04/99	X	AMP ZN	2,9	P ₃ 2 ₁		Sankaranarayanan et al., 1999 et 2000	
				1EVK	04/00		ZN	2,0	P ₂ 2 ₁ 2 ₁			
				1EVL	04/00		Threonyl-adenylate ²⁺	1,5	P ₂		Dock-Bregeon et al., 2000	
				1FYF	09/00		Seryl-adenylate ²⁺ ZN	1,6	P ₂ 2 ₁ 2 ₁			
		1KOG		12/01	X	Threonyl-adenylate ²⁺ ZN	3,5	P ₂		Torres-Larrios et al., 2002 et 2003		
		1NYQ		02/03		Threonyl-adenylate ²⁺ ZN	3,2	P ₂ 2 ₁ 2 ₁				
		1NYR		02/03		ATP ZN	2,8			Dock-Bregeon et al., 2004		
		1TJE					1,5					
		1TKG		06/04		Seryl-adenylate ²⁺	1,5	P ₂				
		1TKE					1,5	P ₂ 2 ₁ 2 ₁				
		1TKY				Seryl-adenylate ²⁺	1,5					
		<i>Homo sapiens</i>	E	1VWT	01/05					Chikayama et al., 2005		
		<i>Pyrococcus abyssi</i>	A	1Y2Q	11/04			2,0	P ₃ 2 ₁		Dwived et al., 2004	
				2HKZ			2,1			Dwived et al., 2006		
				2HL0		Seryl-adenylate ²⁺	1,9	I 2 2 2				
				2HL1		Seryl-adenylate ²⁺	2,3	P ₂ 2 ₁ 2 ₁				
				2HL2		Seryl-adenylate ²⁺	2,6	P ₂ 2 ₁ 2 ₁				
ProRS	19			<i>Thermus thermophilus</i>	B	nd	01/98 ^{OP}	X		3,5		
		1H4Q	05/01			Proline ²⁺ , ATP, Zn	2,9				Yaremchuk et al., 2001	
		1H4S				Prolyl-adenylate ²⁺ Zn	2,9		P ₄ 2 ₁ 2 ₁			
		1H4T				ZN	2,9					
		1H4V					2,4		P ₂ 2 ₁ 2 ₁			
		1HC7					2,4		C 2 2 2		Murayama et al., 2005	
		<i>Escherichia coli</i>		2CI5		06/05			1,9	P 21		
				2DVA				1,6	P ₂ 2 ₁ 2 ₁			
				2I4L				2,8	C 2 2 2			
				2I4M			Prolyl-adenylate ²⁺	2,8			Crepin et al., 2006	
		<i>Rhodospseudomonas palustris</i>		2I4N			Cysteinyl-adenylate ²⁺	2,8	C 2			
				2I4O			ATP, Mg	2,4	C 2 2 2			
				2J3M			Proline ²⁺	2,3	P ₂			
		<i>Enterococcus faecalis</i>		2J3L			Prolyl-adenylate ²⁺	2,3	P ₄ 2 ₁ 2 ₁		Kartekar et al., 2003	
		A	1NU1	12/02		Cysteinyl-adenylate ²⁺ MG ZN	2,5					
			1NU2			MG ZN	3,1					
			1NU5			Prolyl-adenylate ²⁺ MG ZN	2,8	P ₆ 2 ₁ 2 ₁				
			1NU6			alanyl-adenylate ²⁺ MG ZN	2,8					
			1NU8				3,2					
		<i>Methanocaldococcus jannaschii</i>		1W0V	05/04			1,7	P ₂		Murayama et al., 2005	
HisRS	9	<i>Escherichia coli</i>	B	1HTT	03/96		AMP	2,6	P ₁		Arnez et al., 1995	
		<i>Thermus thermophilus</i>		1ADJ	02/97		SO ₄	2,7			Aberg et al., 1997	
				1ADY			Histidyl-adenylate ²⁺ SO ₄	2,8	P ₂ 2 ₁ 2 ₁			
				1KMM			Histidyl-adenylate ²⁺	2,6	P ₁		Arnez et al., 1997	
		<i>Escherichia coli</i>		1KMN				2,8	P ₁			
		<i>Staphylococcus aureus</i>		1QE0	07/99			2,7	P ₆	DV	Qiu et al., 1999	
		<i>Thermus thermophilus</i>		1H4V	05/01		SO ₄	2,4	C 2 2 2		Yaremchuk et al., 2001	
		<i>Thermoplasma acidophilum</i>	A	1WU7	12/04			2,4	P ₂ 2 ₁ 2 ₁	DV	Tanaka et al., 2005	
<i>Homo sapiens</i>	E	1X59	05/05						Nemeki et al., RSGI, 2005			
GlyRS	8	<i>Thermus thermophilus</i>	B	1ATI	04/96			2,8			Logan et al., 1997	
				1B76	01/99		ATP	3,4	C 2 2 2		Arnez et al., 1999	
				1GGM			Glycyl-adenylate ²⁺	3,4				
		<i>Thermotoga maritima</i>		1J5W	07/02			2,0	P ₂ 2 ₁ 2 ₁		JCSG, Wilson, LA, souris	
		E	2PME	04/07			2,9			Xie W., 2007		
			<i>Homo sapiens</i>		2PMF		CL	2,8	P ₄ 2 ₁ 2 ₁		Cader et al., 2007	
					2Q5H			3,0				
2Q5L					SO ₄	2,8						
AlaRS	7	<i>Aquifex aeolicus</i>	B	1RIQ	11/03			2,1	P ₄ 2 ₁ 2 ₁		Swarjo et al., 2004	
				1YFR	01/05		ATP MG	2,2	P ₂ 2 ₁ 2 ₁		Swarjo et al., 2005	
				1YFS				2,1				
				1YFT				2,2	P ₄ 2 ₁ 2 ₁			
				1YGB				2,5				
		<i>Pyrococcus horikoshii</i>	A	1V7O	12/03			2,6	P ₂ 2 ₁ 2 ₁		Sokabe et al., 2005	
				1V4P	11/03		Zn	1,5	P ₂		Ishijima et al., 2006	

Classe IIb

aaRS	n	Espèce	Règne	PDB-ID	Date de dépôt	ARnt	ligand-id	Résolution (Å)	Groupe d'espace	Méthode de cristallisation	Références
AspRS	13	<i>Escherichia coli</i>	B	1C0A	07/99	X	Aspartyl-adenylate ^{an} AMP SO4	2,4	P 4, 2, 2	DV	Eiler et al., 1999
				1EQR	04/00		MG	2,7	C 2	DV	Rees et al., 2000
				1N9W	11/02			2,3	P 2, 2, 2	DV	Charron et al., 2003
				1L0W	02/02			2,0	P 2, 2, 2	Sous micrographie	Ng et al., 2002
				2CF0	02/02			2,4	P 21		Schulze et al., 2006
		1G51		10/00		AMP Aspartyl-adenylate ^{an} SO4	2,4	P 2, 2, 2	DV	Potetzman et al., 2000	
		1EFW		02/00	Xa 1		3,0	P 6	DV	Briand et al., 2000	
		1IL2		05/01	Xa 2	AMO SO4	2,6	P 21	DV	Moulinier et al., 2001	
		1B8A		01/99		ATP MN	1,9	P 2, 2, 2		Schmitt et al., 1999	
		1WYD	02/05		Piperazine	2,3	P 2, 2, 2	DV	Maeda et al., NP		
		<i>Thermus thermophilus</i>	A	nd	06/91 ^{NP}	X	?	3			Ruff et al., 1991
				1ASY	10/95			2,9	P 2, 2, 2		Ruff et al., 1995
				1ASZ	10/95		ATP	3	P 2, 2, 2		Cavarelli et al., 1995
1EOV	03/00					2,3	P 4, 2, 2	DV	Sauter et al., 2000		
Saccharomyces cerevisiae	E										
AsnRS	4	<i>Thermus thermophilus</i>	B	-	01/98		ATP, ade	2,7	P 6		Berthet-Colominas et al., 1998
		<i>Pyrococcus horikoshii</i>	A	1X54	05/05		methy lan AMPan Mg	1,5	C 2	DV	Iwasaki et al., 2006
				1X55			Asparagyl-adenylate ^{an} MG PO4	1,8			
1X56			2,0								
LysRS	11	<i>Escherichia coli</i>	B	1KRS	06/95						Commans et al., 1995
				1KRT	06/95	X					
				1LYL	05/95			2,8	C 2 2 2		Onesti et al., 1995
				1BBU	04/98			2,7			Onesti et al., 2000
				1BBW	04/98			2,7			
				1E1O	05/00			2,1			
				1E1T	05/00		Mg Lysyl-adenylate ^{an} PPI	2,4	P 6, 2 2		Desogus et al., 2000
				1E22	05/00		ATP ^{an} Mg	2,4			
				1E24	05/00		ATP Mn	2,3			
		<i>Pyrococcus horikoshii</i>	A	1IRX	10/01		Zn	2,6	P 2	DV	Terada et al., 2002

Classe II c

PheRS	10	<i>Thermus thermophilus</i>	B	1PYS	11/96		MG	2,9	P 3, 2 1		Mosyak et al., 1995
		<i>Thermus aquaticus</i>		1B70			MG	2,7		DV	
		<i>Thermus thermophilus</i>		1B7Y	01/99		Phenylalanyl-adenylate ^{an} MG	2,5		DV	Reshetnikova et al., 1999
				1ETV	02/00	X		3,3		DV	Goldgur et al., 1997
				1JJC	07/01		Phenylalanyl-adenylate ^{an} SO4 MN	2,6		DV	Fishman et al., 2001
				2ALY	08/05		Tyrosyl-adenylate ^{an} SO4 MN	2,6		DV	
				2AKW	08/05		Phe ^{an} SO4 MN	2,8		DV	Kotik-Kogan et al., 2005
				2AMC	08/05		SO4 MG	2,7		DV	
				2IV5	07/06	X	Phenylalanyl-adenylate ^{an} Mg	3,1		P 3, 2	Moor et al., 2006
		<i>Pyrococcus horikoshii</i>		A	2CXI	06/05		1,9		C 2	MB
PylRS	3	<i>Methanosarcina mazei</i>	A	2Q7E	06/07		ATP		P6 _i	DV	Kavran et al., 2007
		2Q7G									
		2Q7H				Pyrolysine ^{an}					

Annotations :

- l'indice 'an' pour analogue (de l'acide amine, de l'ATP, de l'aa-adenylate)
- Xa1 : ARnt^{Asp} d' *E. coli*
- Xa2 : ARnt^{Asp} *S. cerevisiae*
- NP : non publié
- Les cases colorées en cyan correspondent aux structures résolues par RMN

1. Compilation des résultats

Bilan général : Cette compilation nous permet tout d'abord de dresser un bilan historique sur l'évolution de la qualité des structures, de leur complexité, du rendement annuel des structures résolues et de la nature de la source étudiée. A ce jour, la PDB compte plus de 250 structures d'aaRS sous forme entière ou tronquée dont 30 structures en complexe avec l'ARnt (le premier étant le complexe GlnRS-ARnt^{Gln} d' *Escherichia coli*), (Rould et al., 1989)). Ces 10 dernières années furent particulièrement fructueuses puisque plus de 77% des structures déposées dans la Protein Data Bank le sont depuis 1999. Un bilan statistique est exposé (Tableau I. 13A).

Qualité des données : Intéressons-nous plus particulièrement à la qualité des données. Le graphe présenté Figure I. 14 représente la valeur moyenne de la résolution des structures pour chaque année ainsi que les minima et maxima. Il nous a semblé important de la compléter du cas particulier des structures résolues en complexe avec l'ARNt homologue qui, par leur degré supérieur de difficulté contribue à baisser la moyenne globale. La limite de résolution moyenne des structures résolues entre la fin des années 80 et 2007 n'a pas significativement augmenté. Mais, il faut noter que les valeurs maximales et minimales s'élargissent en fonction du temps et que le nombre de protéines dont la limite haute de résolution dépasse 2 Å augmente significativement avec les années. Si la résolution de complexes macromoléculaires (aaRS-ARNt) contribue à baisser la limite de résolution, on observe une nette augmentation de la proportion des structures résolues à haute résolution ces dernières années : ainsi, entre 1999 et 2006, la limite des 2 Å est franchie pour près de 40% des structures résolues (Figure I. 14).

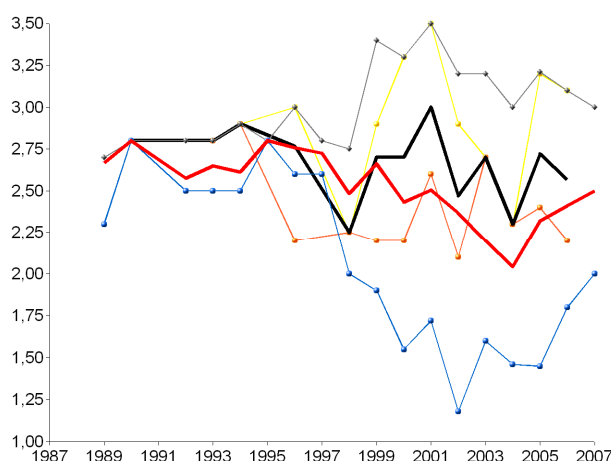


Figure I. 14 : Représentation graphique de la moyenne des résolutions des structures cristallographiques. La moyenne globale est représentée en rouge, la moyenne des structures résolues en complexe en noir. Les minima sont représentés en bleu et orange pour les résolutions globales et en complexe, respectivement, les maxima sont en jaune (global) et gris (complexe). Si la résolution moyenne globale ne diminue pas nettement, l'historique des résolutions maximales augmente sans contestes vers des valeurs de résolution atomique.

Degré d'étude pour un système donné – Prenons en considération l'ensemble des structures résolues par type d'aaRS : celui-ci est donné dans la Tableau I. 12. Si tous les types d'aaRS ont aujourd'hui une structure référencée dans la PDB, elles sont loin d'être étudiées de façon homogène. On remarque que l'éventail est très large puisqu'il s'étend de 3 (PylRS, CysRS, AlaRS) à 36 (TyrRS) structures déterminées pour un système donné. Une étude est complète lorsqu'on peut assembler les formes intermédiaires de l'enzyme lors de la réaction d'aminacylation dans une même espèce. Le système acide aspartique de levure reste un cas isolé : on dispose en effet des structures de l'AspRS seule (Sauter *et al.*, 2000), sous forme complexée (Ruff *et al.*, 1988), et de la structure de l'ARNt isolé (Moras *et al.*, 1980), (Figure I. 15).

Ces études ont permis de visualiser les états des différents substrats et de caractériser les modifications conformationnelles en réponse à la formation du complexe. L'AspRS fait également partie des rares aaRS dont la structure est connue dans les 3 règnes du vivant. Le seul élément manquant de ce système est la structure de l'AspRS non discriminante. Au contraire, les divergences évolutives appuyées sur des études structurales d'enzymes comme la CysRS, dont les 3 structures proviennent d' *E. coli* ou la GlnRS dont on ne connaît que les structures d'origine bactérienne sont peu connues. Ainsi, si les informations disponibles peuvent être très complètes pour certains systèmes, couvrant l'ensemble des règnes du vivant et les formes libre et complexée de l'enzyme, de nombreuses pièces du puzzle doivent être ajoutées pour reconstituer l'histoire complète des aaRS.

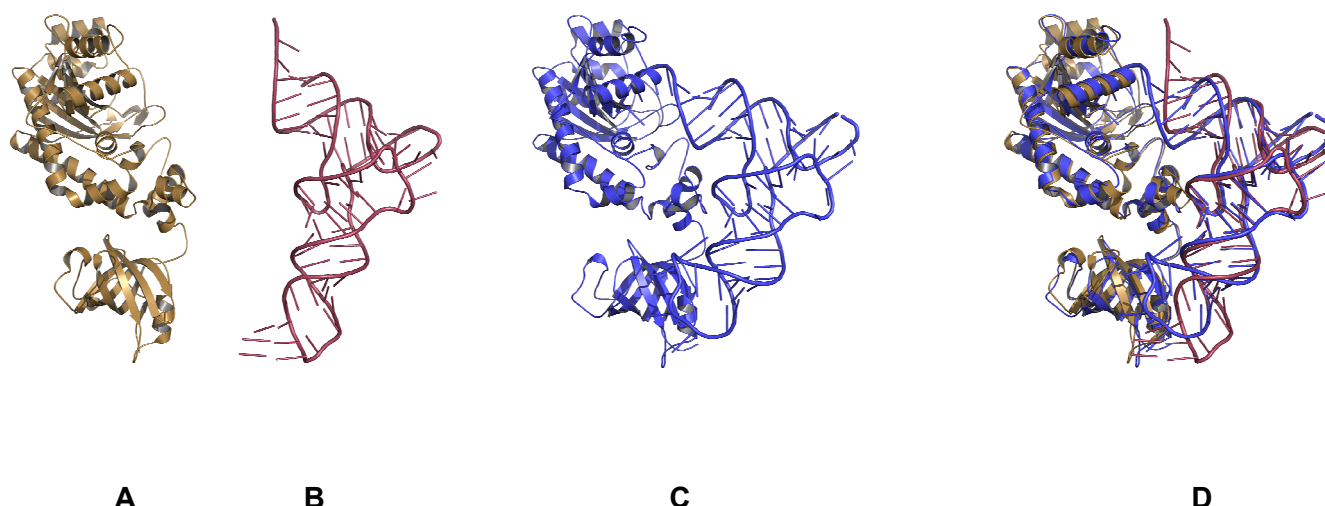


Figure I. 15 : Structures de l'AspRS de levure sous formes libre ou complexée. La structure de l'AspRS sous sa forme apo (**A**) et de l'ARNt^{Asp} libre (**B**) est comparée au complexe AspRS-ARNt^{Asp} (**C**). Les 3 structures sont superposées en (**D**). La conformation de l'ARNt est plus compacte sous forme complexée et présente des modifications de la boucle anticodon qui facilite sa reconnaissance par la synthétase (Ruff *et al.*, 1991). L'adaptation induite par l'ARNt est plus importante que celle de la synthétase avec laquelle il forme le complexe. Les modifications les plus importantes de l'enzyme consistent en une rotation du domaine de liaison à l'anticodon ainsi que 2 boucles conservées au niveau du site catalytique qui interviennent dans la liaison des substrats.

A

Année	Nombre total	Nombre total en complexe	Résolution moyenne (Res)	Résolution maximale (Res_max)	Résolution minimale (Res_min)	Résol en complexe (Res-c)	Res-c maximale (Res-c_max)	Res-c minimale (Res-c_min)	A	B	E
1989	3	0	2,67	2,30	2,70					3	
1990	1	1	2,80	2,80	2,80	2,80	2,80	2,80		1	
1991	1	1	3,00	3,00	3,00	3,00	3,00	3,00		1	
1992	4	0	2,57	2,50	2,80					6	
1993	2	1	2,65	2,50	2,80	2,80	2,80	2,80		2	
1994	4	1	2,61	2,50	2,90	2,90	2,90	2,90		4	
1995	1	1	2,80	2,80	2,80	2,80	2,80	2,80		3	
1996	6	3	2,76	2,60	3,00	2,77	2,2	3		6	
1997	4	0	2,72	2,60	2,80					3	
1998	6	2	2,48	2,00	2,75	2,25	2,25	2,25		5	1
1999	13	5	2,66	1,90	3,40	2,70	2,2	2,9	1	11	
2000	20	10	2,43	1,55	3,30	2,70	2,2	3,3		19	3
2001	22	4	2,50	1,70	3,50	3,00	2,6	3,5	1	19	
2002	27	4	2,37	1,20	3,20	2,47	2,1	2,9	2	19	1
2003	28	2	2,20	1,60	3,20	2,70	2,7	2,7	3	16	6
2004	28	2	2,04	1,45	3,00	2,30	2,30	2,30	8	18	2
2005	44	9	2,32	1,45	3,20	2,72	2,4	3,2	13	26	4
2006	37	4	2,41	1,80	3,10	2,57	2,2	3,1	13	14	10
2007	10		2,81	2,80	3,00				3	3	5
Total	261	50							44	179	32
Pourcentage		19							17	70	13

B

B	<i>Aquifex aeolicus</i>	AlaRS, MetRS*, SerRS
	<i>Bacillus Stearothermophilus</i>	TrpRS*, TyrRS
	<i>Deinococcus radiodurans</i>	GlnRS, TrpRS
	<i>Enterococcus faecalis</i>	ProRS
	<i>Escherichia coli</i>	AspRS*, CysRS*, GlnRS*, HisRS, LeuRS, LysRS*, MetRS, ProRS, SerRS, ThrRS*, TyrRS
	<i>Rhodopseudomonas palustris</i>	ProRS
	<i>Staphylococcus aureus</i>	HisRS, IleRS*, ThrRS, TyrRS
	<i>Synechococcus elongatus</i>	GluRS
	<i>Thermotoga maritima</i>	GluRS, GlyRS, TrpRS
	<i>Thermus aquaticus</i>	PheRS, MetRS, ValRS*
A	<i>Thermus thermophilus</i>	ArgRS, AsnRS, AspRS, GluRS*, GlyRS, HisRS, IleRS, LeuRS, MetRS, PheRS*, ProRS*, SerRS, TyrRS*, ValRS*, AspRS*, SerRS*
	<i>Aeropyrum pernix</i>	ProRS TyrRS
	<i>Aoglobus fulgidus</i>	TyrRS
	<i>Methanocaldococcus jannaschii</i>	TyrRS
	<i>Methanocaldococcus jannaschii</i>	ProRS TyrRS*
	<i>Methanosarcina barkeri</i>	SerRS
	<i>Methanosarcina mazei</i>	PylRS
	<i>Methanothermobacter thermautotrophicus</i>	ProRS
	<i>Pyrococcus abyssi</i>	SerRS, ThrRS
	<i>Pyrococcus abyssi</i>	MetRS
	<i>Pyrococcus abyssi</i>	ThrRS
	<i>Pyrococcus horikoshii</i>	AlaRS, AsnRS, LeuRS*, LysRS II, MetRS, PheRS, SerRS, TyrRS
	<i>Pyrococcus kodakaraensis</i>	AspRS
	<i>Sulfolobus tokodaii</i>	AspRS
	<i>Thermoplasma acidophilum</i>	HisRS
E	<i>Bos taurus</i>	SerRS, TrpRS*
	<i>Homo sapiens</i>	GlyRS
	<i>Neurospora crassa</i>	TyrRS
	<i>Saccharomyces cerevisiae</i>	AspRS*, ArgRS, GluRS, MetRS, TrpRS, TyrRS*

C

Groupe d'espace	Nombre
P 2 ₁ 2 ₁ 2 ₁	52
p 2 1	27
C 2 2 21	27
P 3 ₁ 2 1	25
C 2	23
P 4 ₃ 2 ₁ 2	15
P 4 ₁ 2 ₁ 2	10
C 2 2 2 ₁	9
P 6 5 2 2	7
P 2 ₁ 2 ₁ 2	7
P 4 ₁ 2 ₁ 2	7
P 4 ₃ 2 ₁ 2	6
P 1	4
P 6 1 2 2	4
I 2 ₁ 2 ₁ 2 ₁	3
P 6 ₃ 2 2	3
P 6 ₅	2
I 2 2 2	2
P 6 1 2 2	2
P 4 ₂ 2 ₁ 2	1
R 3	1
P 4 ₁ 2 2	1
P 4 ₁	1
P 6 ₄	1
P 2	1
P 6	1
I 4 ₁ 2 2	1
C 2 2 2 ₁	1
P 2 ₁	1
P 6 ₃	1
c 2	1

Tableaux I. 13 : Tableaux statistiques des structures d'aaRS de la PDB. Tableau A : les nombres d'aaRS résolues par an, libre ou en complexe (colonnes 1 et 2), les limites de résolution moyenne (Res), minimale (Res_min) et maximale (Res_max) pour les enzymes libres ou en complexe (*-c) sont donnés. Enfin, le nombre de structures par règne sont indiquées dans la colonne A pour archée, B pour bactérie et E pour eucaryote. **Tableau B :** Répartition des organismes dans l'étude de la structure des aaRSs. Le classement est donné par règne (B, A, E). (*) : aaRS résolue en complexe avec son ARNt homologue. **Tableau C :** Statistiques pour la répartition du nombre d'aaRS pour un groupe d'espace.

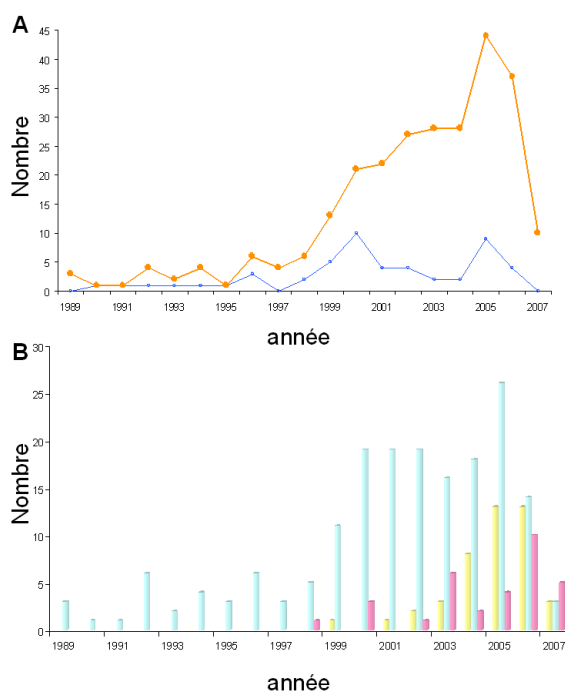


Figure I. 16 - A - Représentation graphique du nombre des structures d'aaRS résolues depuis 1989. Le nombre total de structures résolues par an est représenté en orange (courbe), en complexe avec l'ARNt (bleu). **B- Répartition par règne des structures d'aaRS résolues depuis 1989.** Le nombre d'aaRS bactériennes est représenté en cyan, les archées en jaune et eucaryotes en rose.

Les structures en complexe avec l'ARNt : celles-ci sont bien représentées (seules 6 types d'aaRS n'ont aucune structure en complexe référencée : His-, Gly-, Ala-, Asn-, Pyl-RS et LysRS1) et la présence de l'ARNt est souvent responsable de la stabilisation de domaines qui sont flexibles sous la forme native de la protéine cristallisée (voir § suivant). La résolution de la première structure de complexe aaRS-ARNt, cristallisée au sulfate d'ammonium a également conduit à l'obtention de nombreuses autres structures cristallographiques en complexe, notamment les complexes AspRS-ARNt^{Asp} de levure et d'*E. coli*, PheRS-ARNt^{Phe}, SerRS-ARNt^{Ser} et LysRS-ARNt^{Lys} de *T. thermophilus* (Ducruix et Giegé, 1999a), (Tableau I. 16.B).

Représentativité entre les règnes : La Figure I. 16 B présente la proportion de structures résolues répartie selon les trois règnes du vivant. Les bactéries sont largement surreprésentées comparativement aux archées et eucaryote. Sur la totalité des structures présentes dans la PDB, 70% proviennent de protéines d'origines bactériennes contre 13 et 17% pour les eucaryotes et les archées, respectivement. Le nombre de structures résolues concernant les règnes archée et eucaryote

Empilement cristallin : Certains groupes d'espaces sont largement favorisés pour la cristallisation des aaRS (voir Figure I. 16. B) en particulier le groupe d'espace P2₁2₁2₁. Une étude avait rassemblé les statistiques des groupes d'espaces favoris des macromolécules, pour le cas particulier des monomères. Il en ressortait que les protéines cristallisent préférentiellement dans ce groupe (36%) suivi des mêmes groupes d'espaces retrouvés dans les statistiques restreintes aux aaRS (Wukovitz et Yeates, 1995). On peut admettre à partir de l'ensemble des résultats que les aaRS ne favorisent pas un empilement particulier à leur nature ou à leur propriétés de surface.

2. Quelques leçons tirées des données structurales des aaRS

Nombre : L'augmentation exponentielle du nombre de structures résolues durant cette dernière décennie est applicable à l'ensemble des structures résolues par la cristallographie des rayons X. Cet essor s'explique entre autres par les progrès des méthodes de cristallogenèse, la précision des rayonnements synchrotrons mais aussi des méthodes biochimiques qui permettent d'augmenter les rendements de purification et la pureté du matériel final. Les séquences des génomes d'organismes de plus en plus variés et les possibilités données par la biologie moléculaire (mutations ponctuelles, éliminations de modules désordonnés) sont pour beaucoup dans l'élargissement de l'éventail des structures résolues et dans la diversité de leur source.

Résolution : Les progrès des méthodes de cristallogenèse se reflètent également dans la qualité des structures résolues à haute résolution. Il faut prendre en considération le fait que l'on s'intéresse d'année en année à des espèces de plus en plus complexes au niveau structural (eucaryotes). La cristallisation des protéines eucaryotes dotées de leur domaine additionnel flexible conduit souvent à une population moins stable, moins homogène et par conséquent à des données cristallographiques de qualité inférieure.

Molécules impliquées dans l'assemblage des aaRS : Des modifications conformationnelles d'échelle variable sont souvent induites en réponse à la liaison de ligands aux aaRS. Les molécules impliquées dans ces variations vont des ions (zinc pour la MetRS, Mechulam *et al.*, 1999) à l'ARNt. Dès la première étape de liaison de l'acide aminé, des changements structuraux ont été observés pour les Cys-, Met-, Arg- et TyrRSs. Les structures résolues en présence d'analogues de l'état de transition (amino acyl-adenylate) ont également permis de mettre en évidence le rôle central de la lysine conservée du motif KMSKS comme c'est le cas pour la TyrRS, la MetRS, la TrpRS ou la LeuRS. L'étude structurale relative à l'étape d'activation de la ProRS a permis d'identifier les modifications conformationnelles relatives à la fixation des substrats Pro et ATP mais également à la formation de l'intermédiaire prolyl-adénylate qui semble nécessaire pour les réarrangements conformationnels finaux intervenant dans la liaison de l'extrémité 3' de l'ARNt^{Pro} (Yaremchuk *et al.*, 2000).

La propriété ARNt-dépendante de la réaction d'aminoacylation de certaines enzymes (GluRS, ArgRS, GlnRS, LysRS1) a été expliquée par le réarrangement structural qu'implique la liaison de l'ARNt homologue. Celui-ci contribue à assembler les motifs spécifiquement impliqués dans la reconnaissance des substrats et leur positionnement correct pour la réaction d'activation de l'acide aminé (Sekine *et al.*, 2006, Uter *et al.*, 2006, Terada *et al.*, 2002, Delagoutte *et al.*, 2000). Les modifications conformationnelles vis à vis de la fixation du substrat ne sont pas conservées entre les classes I. Ainsi, le degré de réorganisation en réponse à la fixation de l'ARNt varie largement de la rotation de domaines entiers (domaines d'éditions), à la stabilisation de

boucles (KMSKS, HIGH) en passant par de simples mouvements des chaînes latérales. Pour la Classe II, l'adaptation de l'AsnRS, l'AspRS et la LysRS (classe IIb) se rapporte au domaine en "OB fold" de leur domaine anticodon. En particulier, le domaine anticodon de l'AspRS de levure admet, en réponse à la fixation de l'ARNt, un mouvement de rotation en direction du site catalytique, accompagné d'une translation du module intermédiaire et du passage à une conformation fermée du site actif par l'intermédiaire de 2 boucles mobiles (Cavarelli *et al.*, 1994; Sauter *et al.*, 2000). A plus grande échelle la cristallisation en complexe a également permis de stabiliser des domaines flexibles à l'état libre comme c'est le cas de la PheRS de *T. thermophilus*, la LeuRS de *Pyrococcus horikoshii* et la CysRS d'*E. coli* (voir Figure I. 17).

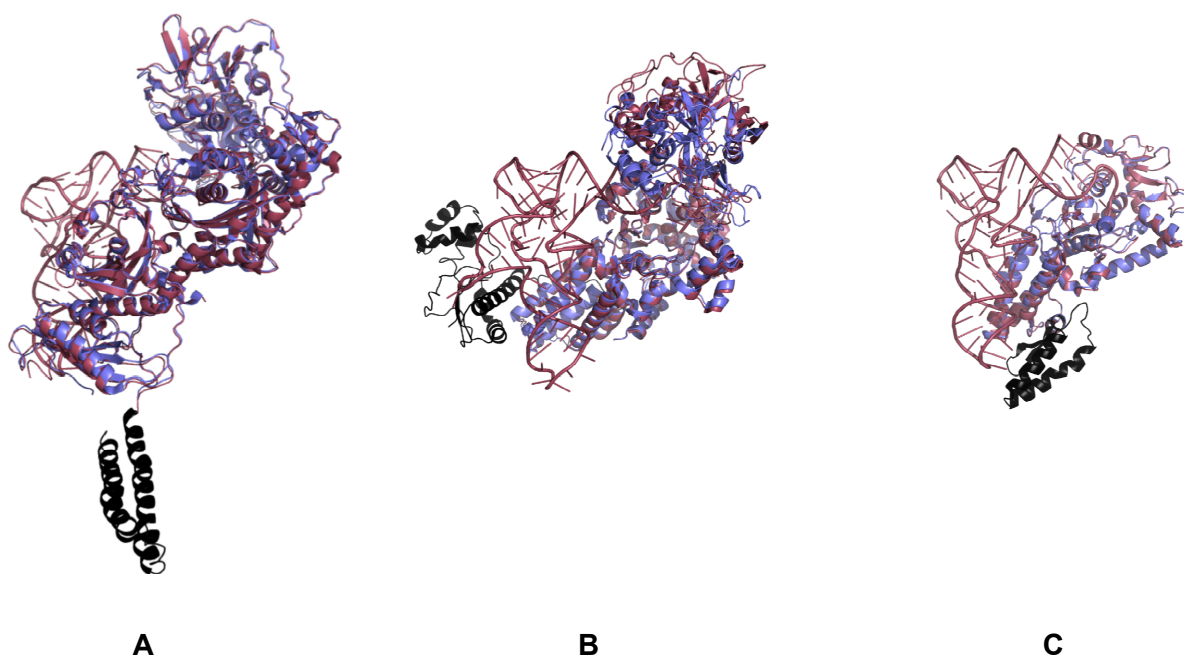


Figure I. 17 : Exemples de structures dont la résolution en présence de l'ARNt correspondant ont permis de stabiliser des domaines flexibles sous forme libre. Structures de **A**- la PheRS de *T. thermophilus*, apo ou complexée à l'ARNt^{Phe} (1JJC et 1EIY, respectivement), **B** - la CysRS d' *E. coli* apo ou complexée à l'ARNt^{Cys} (PDB-ID :1LI5 et 1U0B, respectivement) et **C** - : la LeuRS de *Pyrococcus horikoshii* apo ou complexée à l'ARNt^{Leu} (1WKB et 1WZ2, respectivement). Les structures des enzymes sous forme apo et complexée sont superposées et colorées en bleu et rouge brique, respectivement. Les domaines additionnels stabilisés en présence d'ARNt sont colorés en noir.

.....

Les aaRS eucaryotes - Comme c'est le cas de nombreuses autres protéines, les aaRS d'origine eucaryote présentent souvent des insertions et domaines flexibles en N- ou C-terminal. Comme nous l'avons vu au § I.D.1, elles peuvent interagir avec des macromolécules qui ne sont pas

forcément impliquées dans la réaction d'aminacylation. Il est d'autant plus difficile de les stabiliser qu'on connaît mal les fonctions non canoniques qui leur sont attribués et par conséquent les molécules cibles intervenant potentiellement dans la stabilisation de ces domaines. Ces caractéristiques les rendent plus complexes à cristalliser car bien plus sensibles à la protéolyse et plus difficiles à stabiliser. On a alors souvent recours à la protéolyse ménagée ou à l'élimination complète par mutagenèse dirigée des domaines qui insufflent le désordre potentiel et freinent le processus de cristallisation.

Récemment, par exemple, la TyrRS de mitochondrie humaine a été cristallisée et sa structure résolue sous une version minimale c'est-à-dire sans son domaine additionnel C-terminal (dit S4-like) (Bonnefond *et al.*, 2007), ([Voir annexe II, et articles 2 et 3](#))

Le bilan statistique de cette compilation nous informe sur les disparités des éléments disponibles pour chaque type d'aaRS. L'objectif étant, d'une part, de caractériser les divergences qui existent entre les règnes d'un même système afin de comprendre leur évolution et les variabilités de leur fonctionnement, d'autre part de connaître les adaptations structurales induites lors d'une réaction d'aminacylation et de l'interaction avec les différents substrats qui y sont impliqués. Ainsi, selon le système étudié, on a pu couvrir les trois règnes (Tableau I. 13. B, pour le système Tyr, cf. Tsunoda *et al.*, 2007) et représenter divers états de cette réaction (l'activation, la reconnaissance de l'ARNt, les pré- et post-transferts) ou n'être qu'au commencement du travail (un système, un état pour la LysRS1).

Les progrès des méthodes de cristallogénèse ont contribué, à degré équivalent de complexité, à améliorer la qualité de la résolution des structures et leur nombre. Les séquençages des génomes et les progrès des méthodes de purification ont permis de s'étendre à des organismes plus diversifiés et de compléter peu à peu l'étude des systèmes. Ces structures, combinées aux expériences biochimiques nous apportent l'information nécessaire pour comprendre les divergences fonctionnelles apportées par leur diversité structurale. Grâce à l'étendue des connaissances de ce domaine les pièces du puzzle se rejoignent. Une analyse structurale complète des aaRS est d'autant plus importante dans l'étude évolutive de ces molécules que les structures tridimensionnelles sont plus conservées que leur séquence (O'Donoghue et Luthey-Schulten, 2003). Par exemple, ce type d'analyse a permis de mettre en évidence des couples structuraux interclasses avec des paires isolées spécifiques d'acides aminés stériquement proches, (par exemple la TyrRS, classe Ic, avec la PheRS, classe IIc), qui posent la question de la duplication des gènes donnant lieu à 2 aaRS primordiales d'architectures différentes et spécifiques d'un même substrat (Ribas de Pouplana et Schimmel, 2001b).

III. Ma contribution à l'étude des aminoacyl-ARNt synthétases

Durant mon travail de thèse, je me suis penchée sur le cas d'une aaRS atypique : la glutaminyl-ARNt synthétase de *Deinococcus radiodurans* (GlnRS-Dr). Si cette enzyme présente potentiellement tous les modules nécessaires à son fonctionnement correct, elle se distingue de ses homologues bactériennes par 3 extensions dont la plus frappante est située à son extrémité C-terminale. Celle-ci représente près d'un tiers de sa séquence et est indispensable à son activité catalytique. La dernière moitié de cet appendice présente une homologie de séquence avec un domaine de la sous-unité GatB de l'amidotransférase (Voir § I. C. 4) et à une protéine retrouvée à l'état libre nommée YqeY. L'originalité fonctionnelle de la GlnRS-Dr ayant été préalablement analysée par des études biochimiques, nous nous sommes attachés à comprendre le particularisme structural de cette GlnRS.

Cette étude regroupe les 2 facettes exposées dans cette introduction : l'étude de cette aminoacyl-ARNt synthétase atypique procure un aspect fonctionnel et structural supplémentaire au panel de stratégies employées par la nature pour permettre la réaction d'aminacylation, influencer son efficacité, discriminer ses substrats spécifiques. Cet aspect sera développé dans une première partie. Dans une seconde partie, nous nous attacherons à son étude cristallographique proprement dite. En effet, la GlnRS-Dr n'a pas d'original que sa séquence mais aussi sa propension à cristalliser. La principale particularité de la GlnRS sous sa forme cristalline est que l'ensemble de l'extrémité C-terminale est non visible en densité électronique. Il en résulte que 30% de la protéine est non ordonnée dans le cristal. Cette observation se situe en marge des principes de cristallogénèse selon lesquels une protéine flexible donne des cristaux de mauvaise qualité voir ne cristallise pas. Un autre aspect original réside dans l'élasticité de son empilement cristallin : ainsi, le volume de la maille peut varier de 15% entre absence et présence de ligand tout en maintenant le même empilement.

La cristallisation de la GlnRS en présence d'additifs a permis d'ordonner une extrémité flexible aux abords du site actif lui donnant une conformation fermée. Néanmoins, les structures résolues n'ont pas conduit à la stabilisation de l'extrémité C-terminale qui la rend si originale. Pour connaître les propriétés de repliement de ce module additionnel, des études structurales en solution ont été investies par les techniques de résonance magnétique nucléaire et de dichroïsme circulaire. Nous présenterons ces résultats dans une troisième partie.

Ce projet repose essentiellement sur les aspects structuraux de la GlnRS de *D. radiodurans* et combine des méthodes d'analyse structurale allant de la cristallographie à la RMN. En parallèle de ce projet, j'ai mis à profit les compétences acquises en cristallographie en participant à la résolution de la première structure de TyrRS mitochondriale humaine.

- RÉSULTATS ET DISCUSSION -

I. Préambule

A. Originalités de la GlnRS de *Deinococcus radiodurans*

La majorité des bactérie et des archées utilisent une voie indirecte d'aminoacylation pour compenser l'absence de GlnRS (voir § I. C. 4). La GlnRS est rarement trouvée chez les bactéries et n'a jamais été référencée dans un génome d'archée. *D. radiodurans* est une des rares bactéries qui possède une GlnRS (GlnRS-*Dr*) et donc, utilise la voie directe d'aminoacylation des ARNt (Curnow *et al.*, 1998).

De nombreuses aaRS ont, en plus des modules basiques que sont le domaine de liaison à l'anticodon et le domaine catalytique, des modules additionnels impliqués dans des fonctions diverses (voir § I.C.5 et I.D.1). Aucune GlnRS bactérienne n'a été référencée comme présentant de tels modules. C'est pourtant le cas de la GlnRS-*Dr* et c'est pour cette raison qu'elle a particulièrement retenu notre attention.

B. La GlnRS : une aaRS chimère issue de deux voies d'aminoacylation de l'ARNt

1. Structure primaire

L'organisation structurale de la GlnRS-*Dr* a été établie par alignements de séquence. La GlnRS-*Dr* est hautement homologue aux autres GlnRS à l'exception de quelques insertions et surtout d'un module additionnel qui constituent son originalité : une extension N-terminale (M1-A43), 3 insertions dans la partie conservée de la GlnRS de 17 (412-428 : ***ins1***), 27 (534-560 : ***ins2***) et 5 (615-619 : ***ins3***) résidus, toutes situées au niveau du domaine de liaison à l'anticodon (Figure II.1). La plus surprenante réside à l'extrémité C-terminale de l'enzyme et s'étend sur 220 résidus. En dépit de ces insertions, tous les éléments de séquences caractéristiques d'une GlnRS sont présents et proches de celle d'*E. coli* (GlnRS-*Ec*) avec laquelle elle compte 44% d'identité et la plupart des résidus contactant les substrats (ARNt^{Gln}, ATP, L-Gln) y sont conservés (Deniziak *et al.*, 2007). Un second set d'alignements a été établi pour l'appendice C-terminal de la GlnRS avec le domaine YqeY de 49 bactéries dont ceux des sous unités GatB et GatE des amidotransférases Glu/Asp-AdT (GatCAB) et Glu-AdT (GatDE). Sur les 220 acides aminés constituant l'extrémité C-terminale, les 150 résidus terminaux s'alignent avec les autres protéines avec un meilleur score pour les 30 derniers résidus.

Des approches immunologiques ont permis de confirmer la pérenité structurale de ces protéines avec les protéines YqeY révélées par ces alignements : ainsi, des anticorps anti-AdT spécifiques des sous-unités GatA et GatB de *T. thermophilus* reconnaissent la GlnRS-*Dr* mais pas celle de *E. coli*. Les mêmes anticorps ont été utilisés sur des formes tronquées de la GlnRS-*Dr* : ainsi, un variant tronqué de l'appendice C-terminal (GlnRS-ΔC), n'est pas reconnu par les anticorps mais les formes isolées des appendices, soit sous forme entière (YqeY-λ) soit sans le

linker (YqeY-δ), le sont.

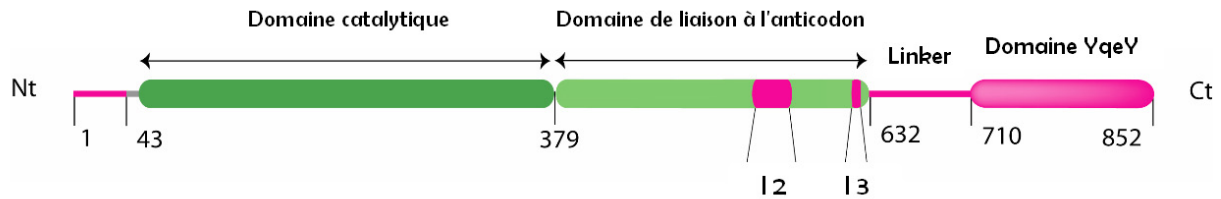


Figure II. 1 : Représentation schématique de la structure primaire de la GlnRS-Dr. Les régions désordonnées de la structure cristallographique sont données en magenta. I2 et I3 indiquent les boucles mobiles *ins* 2 et *ins* 3. Les domaines classiques de l'enzyme sont représentés en vert foncé et vert clair pour le domaine catalytique et le domaine de liaison à l'anticodon, respectivement.

2. Cœur de la GlnRS-Dr

La structure de la GlnRS-Dr a été résolue par cristallographie des rayons X à 2,3 Å de résolution. L'organisation modulaire du cœur de la GlnRS-Dr est proche de celle d'*E. coli*. Les alignement globaux donnent des RMSD compris entre 1,9 et 2,3 Å (structure d'*E. coli* native ou en complexe avec l'ARNt), mais donnent des valeurs plus faibles pour les domaines isolés. Pour cette raison, la nomenclature utilisée pour la GlnRS d'*E. coli* a été appliquée à la GlnRS-Dr. Le site actif (SA) se décompose en un domaine de Rossmann (DNF, dinucleotide fold) et le domaine de liaison à la tige acceptrice (ABD), le domaine de liaison à l'anticodon (ACD) se découpe en un domaine distal (DBB) et proximal (PBB), riche en feuillets β. Les deux modules sont reliés par un domaine intermédiaire (HSD, helical subdomain), (Figure II. 2).

En dépit des bonnes propriétés de diffraction de ces cristaux, l'extension N-terminale, les insertions *ins1* et *ins2* et surtout l'extrémité C-terminale longue de 220 résidus ne sont pas visibles dans la structure représentant 35% de régions mobiles dans le cristal.

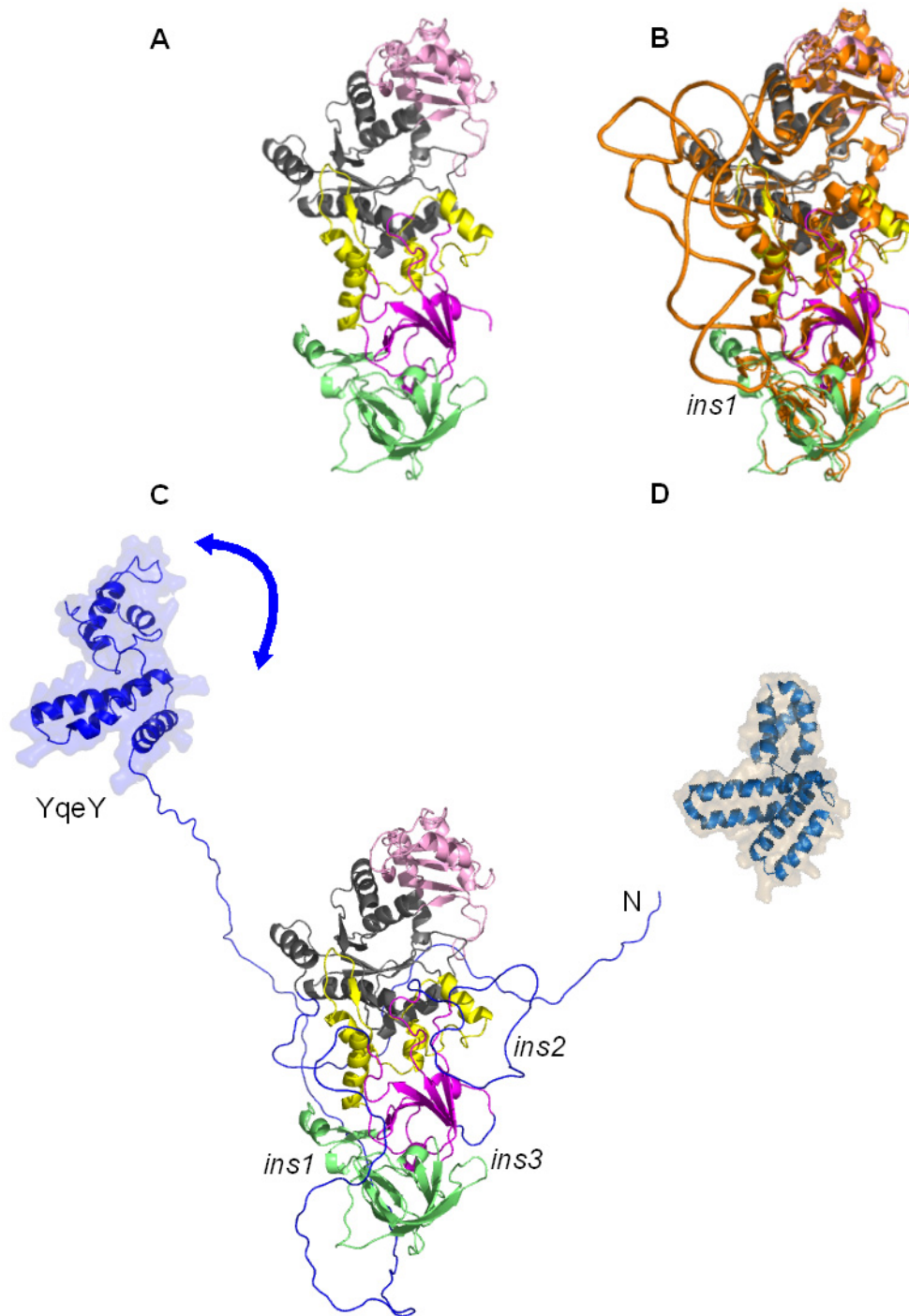


Figure II. 2 : GlnRS-Dr, une enzyme chimérique – A, B : Architecture classique du cœur conservé. **A-** Structure de la GlnRS de *Deinococcus radiodurans* résolue à 2,3 Å de résolution (PDB_id : 2HZ7), (Deniziak *et al.*, 2007). Le code couleur respecte la nomenclature du système *E. coli*. On distingue les 2 modules du site catalytique que sont le domaine de Rossmann (DNF, DiNucleotide Fold, 47-147 :252-305, en gris) et le domaine de liaison à la tige acceptrice (ABD, Acceptor-Stem Binding Domain, 143-251, en rose pâle). Le site de liaison à l'anticodon est composé de 2 sous-domaines : le domaine proximal (PBB, Proximal β -barrel, 380-387 :516-633, magenta) et le domaine distal (DBB, Distal β -barre, 388-515, vert pâle). Les modules catalytiques et de reconnaissance à l'anticodon sont reliés par un module intermédiaire (HSB – Helical subdomain, 306-379, en jaune) La plupart des insertions (sauf *ins1*) ne sont pas visibles dans la structure. **B-** Superposition des structures de la GlnRS de *Deinococcus* et *E. coli* (PDB_id : 1O0B). Les codes couleurs pour la GlnRS-Dr sont les mêmes qu'en A. La GlnRS-Ec est représentée en orange clair. Le RMSD global est de 1,9 Å; l'architecture classique est conservée. **C-** Modèle de la GlnRS-Dr dotée des régions non visibles dans la structure. Les régions correspondant au modèle sont en bleu. Les extrémités N- et C-terminale avec le domaine Yqey (surface), les insertions *ins1* et *ins2* sont précisées. Le modèle a été déterminé à partir des structures de la GlnRS et de Yqey de *Bacillus subtilis* (PDB_id : 1NG6). **D-** Structure de Yqey de *Bacillus subtilis*.

C. Rôle de l'extrémité C-terminale et implications du domaine YqeY

1. Extrémité C-terminale, indispensable à la réaction d'aminacylation

L'extrémité C-terminale additionnelle a fait l'objet d'expériences biochimiques afin de comprendre son rôle dans la protéine. Ainsi, la GlnRS-ΔC est encore capable de charger son ARNt homologue mais la délétion du domaine YqeY induit une perte significative d'activité catalytique. Cette dernière ne se traduit ni au niveau de l'efficacité de charge de l'acide aminé ni pour sa spécificité. En revanche, la perte d'affinité entre la GlnRS-ΔC et la forme entière pour l'ARNt^{Gln} (issu de l'ARNt totale de *D. radiodurans*) est significative ($K_M^{(GlnRS-ΔC)} / K_M^{(GlnRS)} = 24$). La présence du module YqeY accroît donc l'activité catalytique de l'enzyme tant dans sa fonction d'activation de la glutamine, ARNt dépendante, que pour la charge de son l'ARNt homologue.

La GlnRS-Dr peut être décomposée en deux parties : la première correspond au cœur, commun aux GlnRS, proche d'*E. coli* et dont la plupart des résidus impliqués dans sa fonction sont conservés. La seconde est représentée par le domaine YqeY conservé dans les sous-unités GatB de l'AdT. Cette dernière palie chez une majorité de bactéries et archées, l'absence de GlnRS et/ou d'AsnRS par l'utilisation d'une voie indirecte d'aminacylation. Plus précisément, la sous-unité GatB y joue une fonction de reconnaissance de l'ARNt misacylé (Deniziak *et al.*, 2007, voir aussi [Introduction, I. C. 4](#)). Ensemble, ces éléments suggèrent que la GlnRS-Dr est une protéine chimérique de deux acteurs de la voie directe et indirecte d'aminacylation.

2. YqeY : un 'module' libre ou lié

Le domaine YqeY est également retrouvé à l'état libre chez une centaine d'organismes bactériens ou eucaryotes et, notamment chez *Bacillus subtilis* dont la structure tridimensionnelle a été résolue par un projet de génomique structurale (PDB_ID : 1NG6). Il est toujours intégré à la partie C-terminale du domaine GatB de GatCAB (Glu-Asp AdT) ou du domaine GatE de GatDE (Glu-AdT, archée). Comme introduit précédemment, YqeY est également présent dans la partie C-terminale de la glutaminyl-ARNt synthétase de *D. radiodurans* mais aussi chez *D. geothermalis* (bactérie, deinococcus-thermus), *Salinibacter ruber* (bactérie, bacteroidetes) et *Anaeromyxobacter dehalogenans* (bactérie, protéobactérie). Cette architecture est unique pour toutes les GlnRS procaryotes.

De manière générale, sa fonction à l'état libre comme à l'état lié n'a pas encore été clairement définie. De plus, parmi les structures cristallographiques résolues des amidotransférases GatCAB et GatDE (Feng *et al.*, 2005; Nakamura *et al.*, 2006; Schmitt *et al.*, 2005) seul celle du complexe GatDE-ARNt^{Gln} a permis d'en stabiliser une partie (Oshikane *et al.*, 2006). Néanmoins, les analyses biochimiques concordent pour lui attribuer une fonction de liaison aux acides nucléiques et un rôle dans la catalyse des réactions les impliquant : ainsi, la délétion de ce domaine, soit pour la GlnRS-Dr, soit pour les amidotransférases GatCAB ou GatDE provoque une chute de leur activité catalytique. En revanche, les analyses par retard sur gel de deux

variants isolés (YqeY- δ et YqeY- λ , voir [Figure II. 5](#)) du module C-terminal additionnel montrent que celui-ci n'est capable de lier l'ARNt^{Gln} que lorsqu'il est fixé de manière covalente au cœur de la GlnRS.

D. Particularité de l'empilement cristallin de la GlnRS ou la cristallisabilité d'une enzyme désordonnée

L'histoire de la cristallisation des macromolécules biologiques a souvent montré qu'une protéine flexible avait peu de chances de cristalliser. Les premiers succès de la cristallisation de macromolécules biologiques ont souvent été le résultat d'ingénierie génétique conduisant à la suppression de ces régions instables qui perturbent l'empilement du cristal et la formation des contacts intermoléculaires nécessaires à sa croissance. Même limité au domaine des aaRS, les exemples pour lesquelles cette stratégie a porté ses fruits ne manquent pas, depuis les premières digestions par protéolyse ménagée de la MetRS d'*E. coli* (Waller *et al.*, 1971), à plus récemment, l'AspRS de levure ou la TyrRS mitochondriale qui ne cristallisent que sous une forme dépourvue de leur extrémité additionnelle N ou C-terminales (Bonfond *et al.*, 2007; Sauter *et al.*, 2000). Cette règle générale qui consiste à déléter les régions dynamiques ou potentiellement capables d'introduire un désordre dynamique dans le cristal continue indéniablement de faire ses preuves. L'exemple de *D. radiodurans* offre cependant un contre-exemple notoire avec une protéine dont l'importante proportion de régions désordonnées n'altère pas sa capacité à cristalliser. Cette enzyme est, comme nous l'avons vu précédemment, constituée d'un cœur conservé chez son homologue d'*E. coli* et d'une extrémité C-terminale additionnelle de 220 résidus que les logiciels de prédiction de structure présentent comme désordonnée sur 70 résidus. La GlnRS de *D. radiodurans* sous sa forme entière a *a priori* une faible capacité à cristalliser car 30% de sa structure est potentiellement désordonnée en solution. Des tests effectués en DLS ont révélé qu'elle était monodisperse et homogène en solution et ont alors encouragé à engager des tests de cristallisation. Des cristaux de GlnRS ont été obtenus (Deniziak *et al.*, 2004) et la structure tridimensionnelle résolue (Deniziak *et al.*, 2007). De façon surprenante, la densité électronique résultant de l'analyse des données a permis de résoudre seulement deux tiers de la structure de la protéine totale. La possibilité de protéolyse de la partie flexible du module additionnel a été rejetée par analyse des cristaux dissous sur gel de polyacrylamide SDS en conditions dénaturantes, suivie de leur analyse en spectrométrie de masse ([Figure II. 3](#)), prouvant que la quasi totalité de l'enzyme est belle et bien présente dans le cristal. Ce décalage de masse détecté lors des expériences en MALDI ne peut en effet être lié à l'absence de l'extrémité C-terminale additionnelle. Le séquençage de l'échantillon a montré qu'une protéolyse était située à l'extrémité N-terminale. L'espace disponible dans l'empilement cristallin est donc suffisant pour loger une extrémité flexible et dynamique.

Ainsi, la cristallisation de la GlnRS-*Dr* défend le principe selon lequel un cristal peut-être constitué d'une proportion conséquente de la protéine non ordonnée, la seule condition étant

qu'elle puisse maintenir une surface de contact intermoléculaire suffisante pour permettre la stabilité de l'empilement cristallin. Ce résultat nous pousse à tirer une première leçon sur la cristallisabilité d'une protéine d'intérêt. Avant d'en ôter les domaines flexibles, on doit avant tout contrôler les propriétés de monodispersité et l'homogénéité de la protéine. Si c'est le cas, c'est une candidate à ne pas négliger pour la cristallisation. Les détails de cette étude sont donnés dans l'article qui suit.

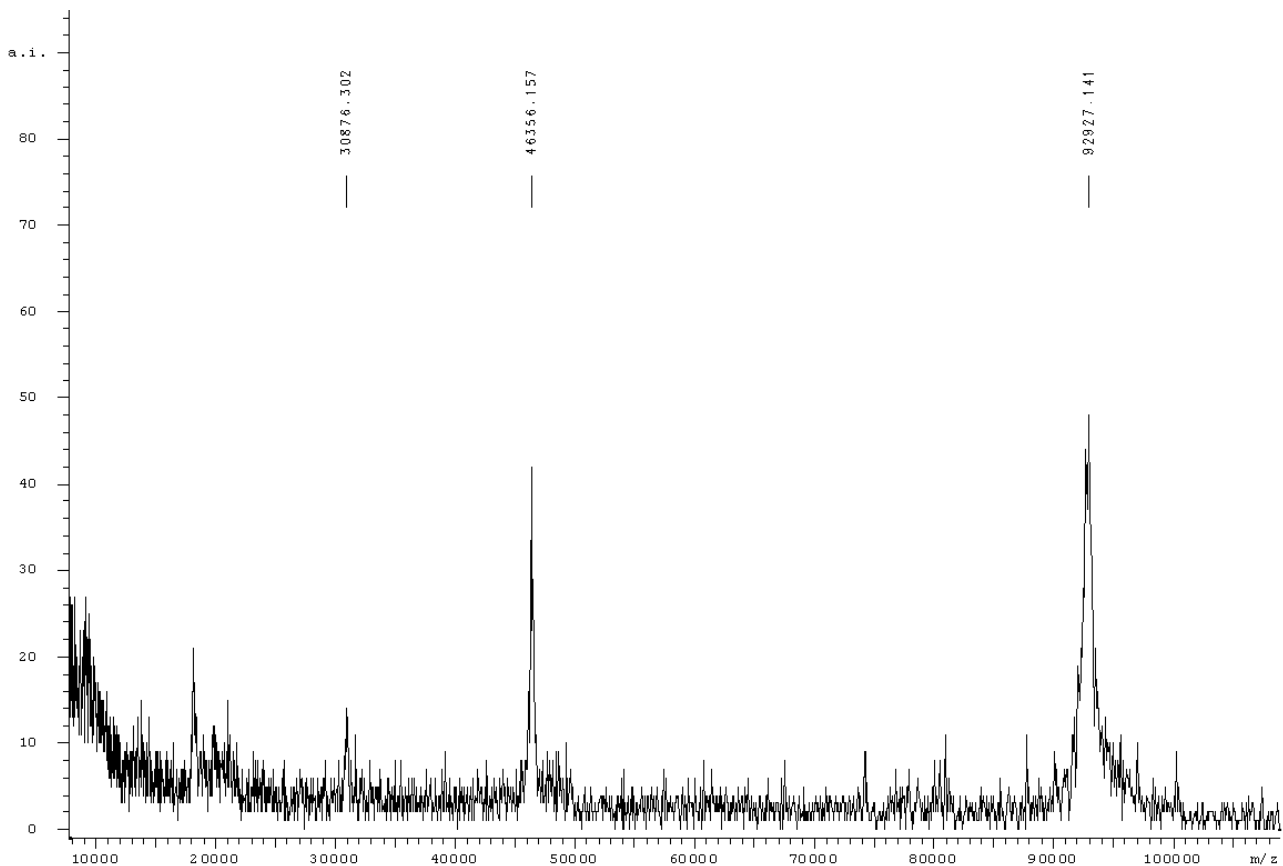


Figure II. 3 : Spectre de masse (MALDI-TOF) effectué sur les cristaux dissous de GlnRS. Il donne une estimation de la masse totale (92927 Da) légèrement décalée de celle de sa masse réelle. Ce décalage ne peut néanmoins être lié à l'absence de l'extrémité C-terminale additionnelle. Le séquençage de l'échantillon a permis de montrer que la protéolyse était située à l'extrémité N-terminale.

Disorder can exist inside well-diffracting crystals

Élodie TOUZÉ, Bernard LORBER, Marzanna DENIZIAK, Hubert Dominique BECKER, Daniel KERN, Richard GIEGÉ, and Claude SAUTER

Crystal Growth & Design, 2007, Vol. 7, N° 11, Pages 2195–2197

Pages 60-62 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1021/cg7007057>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://www-sicd.u-strasbg.fr/services/peb/>

II. Etude cristallographique de la GlnRS de *D. radiodurans*

A. Les objectifs de l'étude

La cristallisation de la GlnRS-*Dr* a conduit à la résolution de sa structure dont seul le cœur conservé est visible en densité électronique. Dans la continuité du travail de cristallogénèse, l'objectif de ce travail était de caractériser les régions non visibles et, en particulier, l'extrémité C-terminale. Pour ce faire, un premier volet de mon travail a suivi différentes stratégies qui visaient soit à stabiliser les extrémités flexibles dans l'empilement cristallin, soit à fixer leur conformation en les cristallisant en présence de leurs substrats naturels ou d'analogues synthétiques impliqués dans l'interaction avec l'enzyme. Dans cette optique, nous avons tenté dans un premier temps de cristalliser la GlnRS dans un empilement cristallin différent de $P2_12_12_1$ par la recherche de nouvelles conditions de cristallisation. Afin d'induire des conformations plus stables, des recherches de conditions de cristallisations ont été engagées en présence de l'ARNt^{Gln} transcrit, de glutamine et d'analogues de l'intermédiaire réactionnel adénylate : le glutaminyl-adénylate (glutamine-sulfamoyl adenosine, QSI, RNA-TEC) ou le glutaminol (fournis par l'équipe de Robert Chênevert, Université Laval, Québec), (Figure II. 4).

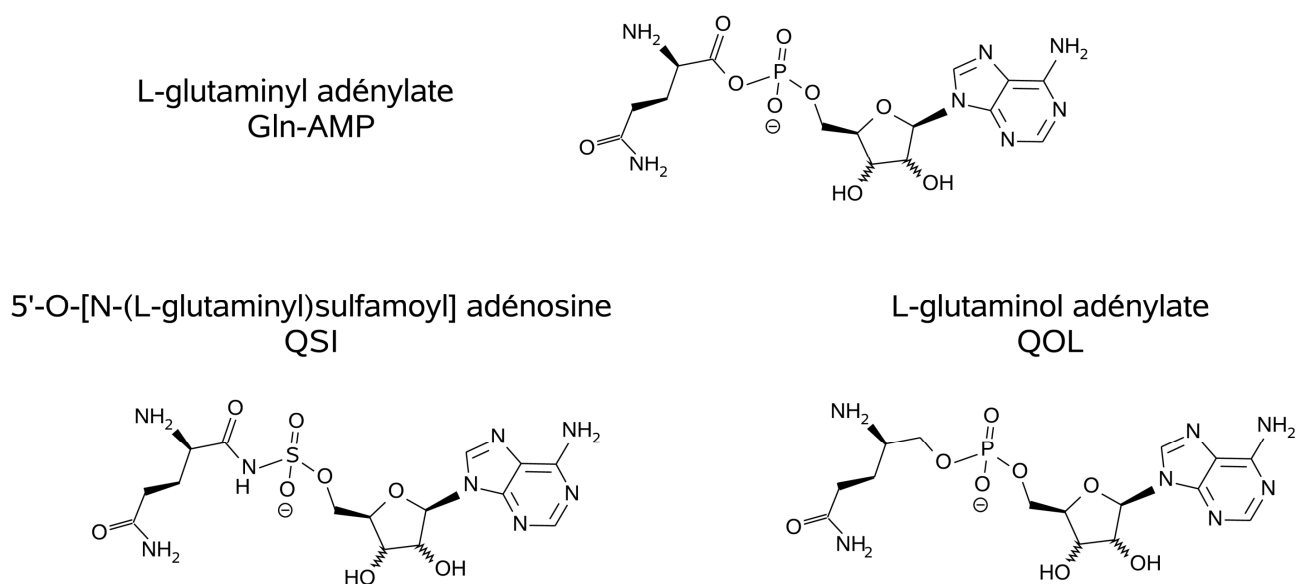


Figure II. 4 : Adénylate et analogues. La figure représente la formule de l'intermédiaire réactionnel naturel de la réaction de glutaminylation de l'ARNt^{Gln} (L-glutamyl adénylate, Gln-AMP) et les deux analogues synthétisés. Le 5'-O-[N-(L-glutaminyl)sulfamoyl]adénosine ou **QSI** et le L-glutaminol adénylate ou **QOL**.

A mon arrivée au laboratoire, 4 constructions dérivées de la GlnRS-*Dr* étaient disponibles (voir Figure II. 5). La GlnRS- Δ , dépourvue de l'extrémité additionnelle, le domaine YqeY- δ , homologue de l'extrémité C-terminale de la sous unité GatB de l'Adt et YqeY- λ , constitué du domaine YqeY- δ et du "linker", qui rejoint le cœur conservé de la GlnRS au domaine YqeY. Chacun de ces dérivés a été purifié d'après le protocole détaillé dans le chapitre *Matériel et Méthodes*, III.

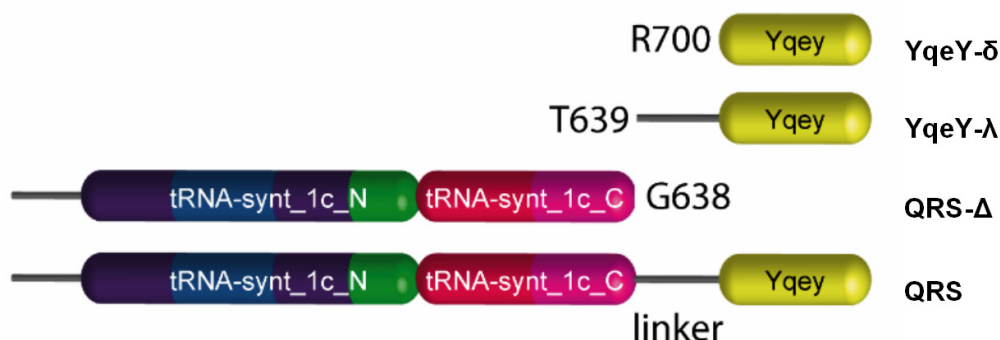


Figure II. 5 : Représentation schématique des 4 variants de la GlnRS : YqeY- δ , YqeY- λ , GlnRS- Δ et GlnRS native. Les résidus terminaux de chacun d'entre eux sont précisés. Le domaine additionnel YqeY est représenté en jaune.

B. Stratégies de cristallisation

1. Recherche de nouvelles conditions de cristallisation

La GlnRS-*Dr*, - Δ , YqeY- δ et - λ purifiées ont été soumises à des tests de cristallisation. Seule la forme YqeY- δ n'a conduit à aucune piste de cristallisation. La recherche de nouvelles conditions de cristallisation de la GlnRS-*Dr* par l'utilisation des cribles commerciaux Qiagen® (PEGs Suite I et II, Classics, JSCG) a conduit à l'obtention de nouvelles conditions de cristallisation qui ont été reproduites et optimisées en batch, en capillaire, en diffusion de vapeur et en présence de gel d'agarose. Des tests de cristallisation en champ magnétique ont également conduit à l'obtention de cristaux orientés (voir Figure II. 6). Les nouvelles conditions ont été utilisées pour la cristallisation en présence de petits ligands : deux analogues de l'adénylate (QSI et QOL) et la glutamine. Des tests de cristallisation en complexe avec l'ARNt^{Gln} transcrit ont également été effectués avec ou sans analogues, reproduites et optimisées en diffusion de vapeur en présence de PEG de haut poids moléculaire (PEG 6000 et 20000). La GlnRS- Δ cristallise dans de nombreuses conditions du kit de cristallisation Qiagen® PEGI Suite et donne de très bonnes bases pour l'obtention de cristaux de qualité. Quelques exemples sont donnés Figure II. 6. Au contraire, les tests de cristallisation des deux dérivés YqeY- δ et YqeY- λ ont été peu fructueux. Dans le premier cas, l'absence de "hits" est probablement due à la forte instabilité de YqeY- δ , aisément dégradé, même lors de l'étape de purification.

	1	2	3	4	5
a	QRS / QSI 16% Peg 3350 – 200 mM NaF	QRS Native 15% Peg 3350 – 200 mM NaF	QRS / QSI 16% Peg 3350 – 200 mM Ammonium formate pH 8,5	QRS / QSI 20% Peg 3350 – 200 mM Ammonium formate pH 7	QRS / gln 20% Peg 3350 – 200 mM Ammonium formate pH 7
b	QRS Native 15% Peg 3350 – 200 mM NaF	QRS Native 15% Peg 3350 – 200 mM NaBr	QRS Native 15% Peg 3350 – 200 mM LiBr	QRS Native 15% Peg 3350 – 200 mM KBr	QRS Native 15% Peg 3350 – 200 mM NaBr
c	QRS - agarose 16% Peg 3350 – 200 mM NaF	QRS + agarose 16% Peg 3350 – 200 mM NaF	QRS + agarose 16% Peg 3350 – 200 mM NaF	QRS Peg 3350 – 200 mM Ammonium formate pH 8,5 CHAMP MAGNETIQUE	QRS / GOL + agarose Peg 3350 – 200 mM Ammonium formate pH 8,5
d	QRS-ΔC / gln 20% Peg 3350 – 200 mM Ammonium formate pH 7	QRS-ΔC / gln 20% Peg 3350 – 200 mM acétate ammonium	QRS-ΔC 12% Peg 20000 – 0,1M Mes pH 6,5	QRS-ΔC 20% Peg 10000 – 0,1M Sodium acetate pH 4,6	QRS-ΔC / GOL 20% Peg 3350 – 200 mM Ammonium formate pH 7
e	QRS-ΔC 20% Peg 3350 – 0,2M diammonium citrate	YqeY-λ 30% Peg 8000 – 0,2M LiSO4 0,1 M Sodium acétate pH 4,5	YqeY-λ 25% Peg 550 – MME 0,1M Hepes	QRS-ARNt 20% Peg 10000 – 8% Ethylene Glycol – 0,1 M Hepes pH 7,5	QRS-ARNt-adenylate 10% Peg 6000 – 2M Chlorure de sodium

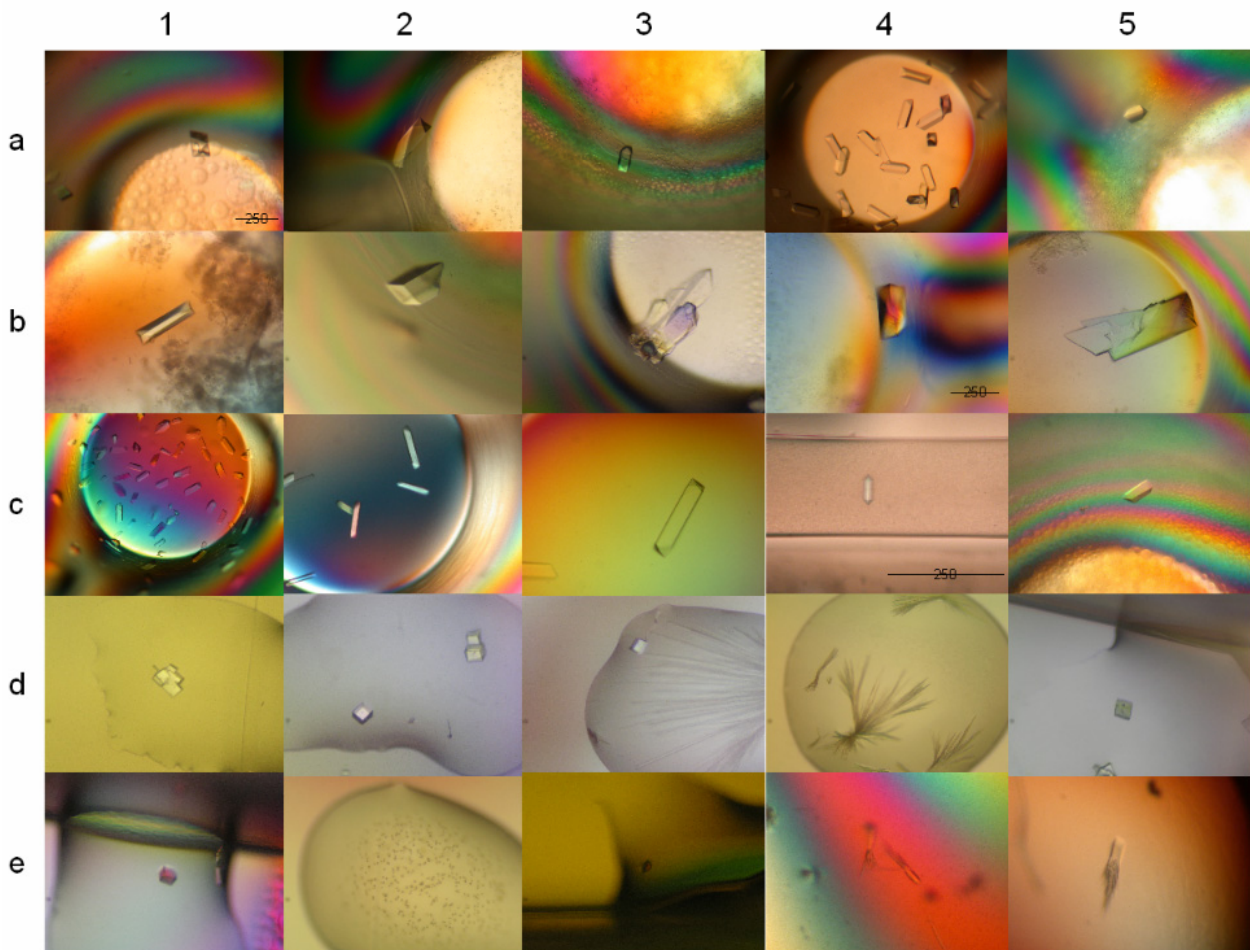


Figure II. 6 : Tableau des conditions de cristallisation et exemples des cristaux obtenus – Lignes a ; b : Exemples de cristaux de la GlnRS native (a1-c5) cristallisée en batch, avec gel d'agarose (b1-5; c2,3), en capillaires (c4), de la GlnRS-ΔC (d1-5), de YqeY-λ ou de la GlnRS en complexe avec l'ARNt^{Gln} transcrit (e4-5). Les cristaux obtenus ont été testés à l'ESRF (European Synchrotron Radiation Facility) de Grenoble. Les co-cristaux de GlnRS et des petits ligands ont permis d'enregistrer 4 jeux complets de données entre 1,9 et 2,5 Å en présence de glutamine, et de 2 analogues d'adénylate (QSI et QOL). Les cristaux obtenus en présence de la GlnRS-ΔC ont conduits à l'enregistrement d'un jeu de données à 1,9 Å de résolution. Les cristaux obtenus en présence d'ARNt^{Gln} ne diffractent pas au-delà de 25 Å. Les données de la GlnRS-QOL et la GlnRS-ΔC ont fait l'objet d'une analyse préliminaire.

2. Cristallisation en gel

La cristallisation en gel d'agarose a souvent permis l'obtention de cristaux de meilleure qualité, tant pour leur taille du cristal que pour leurs propriétés de diffraction. De manière générale, le gel d'agarose, en polymérisant, forme une matrice dans la goutte. La cristallisation en gel d'agarose est susceptible d'améliorer la qualité du cristal puisqu'elle mime l'effet de la microgravité (absence de convection dans la liqueur mère) et évite au cristal de sédimenter dans la goutte (Ng *et al.*, 1997; Lorber *et al.*, 1999) (voir *introduction II.A, Matériel et Méthodes IV.D.1.c*). Dans notre cas, nous avons pu observer des améliorations des qualités physiques du cristal (régularité, durée de vie, taille) comme de ses propriétés de diffraction.

3. Stabilisation par modification des résidus lysine

La composition de l'extrémité C-terminale additionnelle non visible, riche en lysine dans la partie prédite flexible nous a poussé à utiliser la méthode de méthylation des lysines. La méthylation des lysines consiste à modifier l'extrémité radicale des lysines exposées en solvant en ajoutant un groupement méthyl au NH₂ terminal. Ce traitement a déjà donné des résultats encourageants pour augmenter la qualité de cristaux et permettre l'enregistrement d'un jeu de donnée (passage de 7 à 3,2 Å de résolution) (Rauert *et al.*, 2007)

Sur la GlnRS-Dr, le traitement a un effet, puisque la fraction traitée chargée sur un chromatographie d'exclusion de taille conduit à l'élution de deux populations distinctes révélées en spectrométrie de masse : une population hétérogène et une population plus homogène (Figure II. 7). La présence des deux populations est probablement liée au nombre élevé de lysines. Les expériences de MALDI-TOF ont montré une différence de taille de 1500 Da entre la fraction éluee et traitée (Figure II. 7, fraction 11-12), et le témoin. Même si la méthode n'est pas particulièrement adaptée à ce type de mesures, la différence est significative et nous permet de dire qu'il y a effectivement eu méthylation. Des expériences par ESI (Electro-Spray Ionization) ont été initiées donnant des résultats peu satisfaisants vu la quantité limitée de matériel.

Ce résultat préliminaire laisse également préjuger un changement de conformation notable puisque la population méthylée (homogène, réaction totale, fraction 11-12) et peu ou pas méthylée (réaction partielle, fraction 8-9) peuvent être éluées séparément. Malheureusement, la stabilité de la GlnRS soumise à ce type de réaction chimique est fortement altérée et le manque de matériel en fin de purification ne nous a pas permis de rassembler des quantités suffisantes pour engager les tests de cristallisation. Il apparaît en effet que la réaction, déjà loin d'être totale (30% de rendement) entraîne une perte non négligeable de matériel due à la précipitation de la protéine traitée (50%), (Walter *et al.*, 2006).

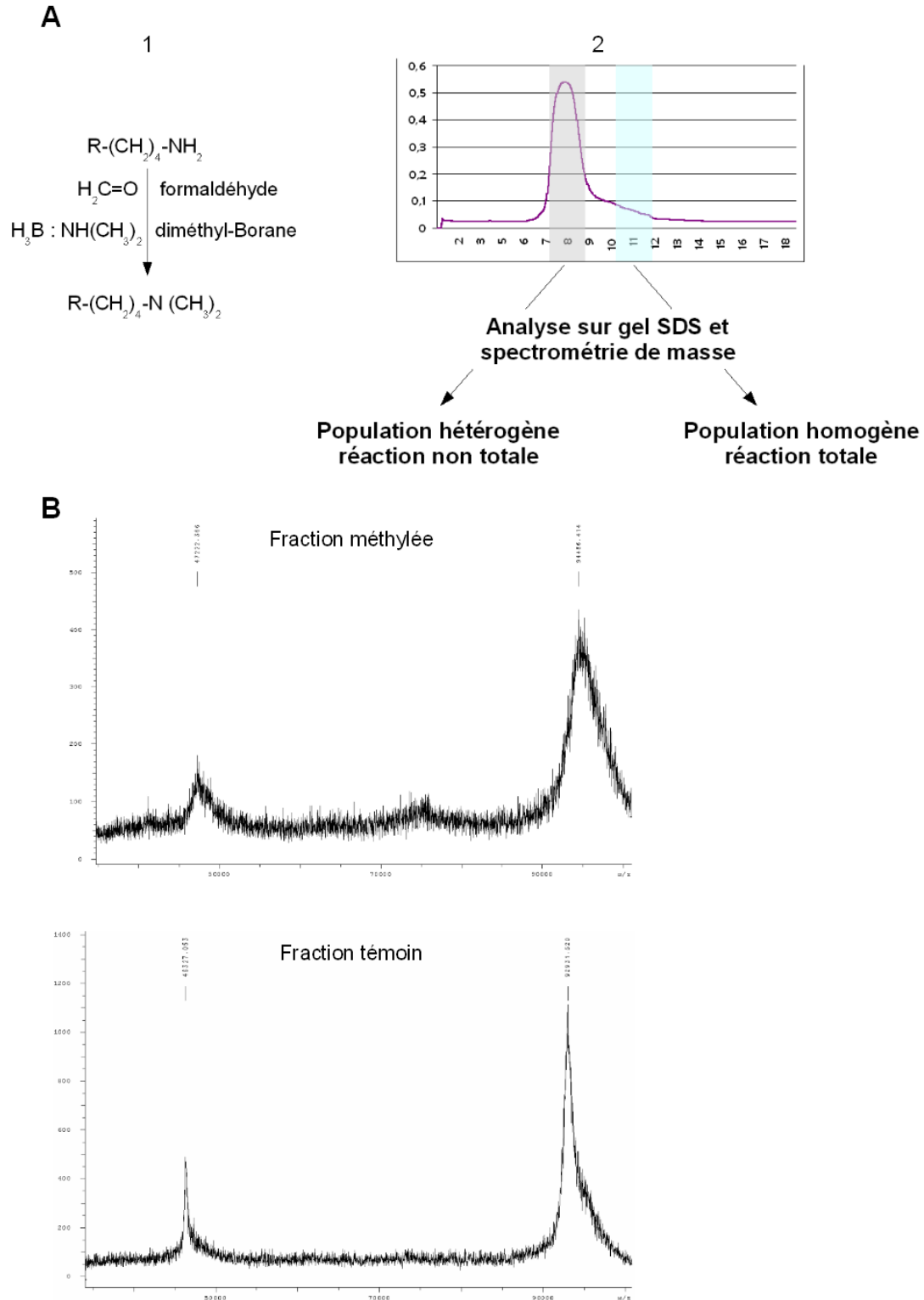


Figure II. 7: Schéma du protocole suivi pour la méthylation des lysines. A : 1 - Réaction opérée sur la chaîne latérale des Lysines ($\text{R}-(\text{CH}_2)_4-\text{NH}_2$). En présence de formaldéhyde et de diméthyl Borane, les NH_2 terminaux sont méthylés. 2. Profil de la chromatographie d'exclusion de taille de la GlnRS traitée. Les fractions sont analysées sur gel SDS en conditions dénaturantes et en spectrométrie de masse. La population majoritaire est très hétérogène (fraction 7 à 9). Une petite part de la fraction traitée (11-12) est effectivement méthylée et peut-être séparée sur chromatographie d'exclusion de taille. B : Spectres MALDI-TOF en spectrométrie de masse de la fraction méthylée de la GlnRS (11-12, Fraction méthylée) et de la GlnRS non traitée (Fraction témoin). On observe un décalage de 1500 Da entre les deux échantillons. Les mesures sont effectuées sur un appareil Brucker Biflex III. Les morceaux de gel correspondant sont découpés et digérés à la trypsine pour les mesures (Voir aussi le chapitre *Matériel et Méthodes*, V. A).

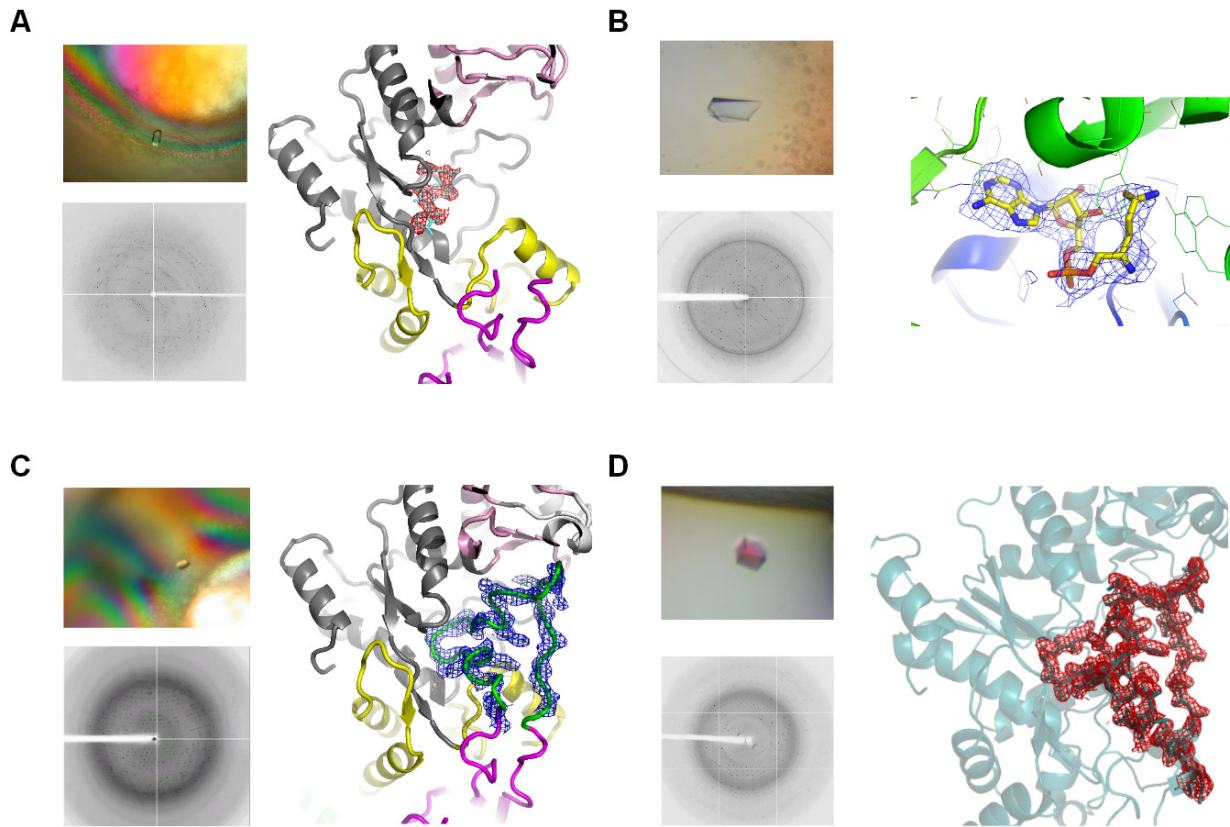
4. Données cristallographiques

L'analyse des jeux de données a conduit à la résolution des 4 structures de la GlnRS-Dr. Les données sont résumées dans la [Tableau II. 8](#) et illustrées [Figure II. 9](#). La cristallisation de la GlnRS en présence d'analogue de l'adénylate (QSI) a conduit à l'enregistrement de données à plus haute résolution que l'enzyme sous forme apo. Quelles que soient les conditions trouvées, la GlnRS native, en présence des petits ligands ou même la GlnRS-ΔC admettent le même empilement cristallin $P2_12_12_1$. Néanmoins, la comparaison des paramètres de maille des structures de la GlnRS entière résolues avec ou sans les petits ligands révèle une variation du volume de la maille dans 2 des 3 dimensions. Ainsi, pour 2 des structures résolues en présence de substrat (QSI et Gln), on observe une augmentation de 9 et 5 Å des paramètres a et c respectivement. L'augmentation des paramètres de maille de la GlnRS co-cristallisée avec le QOL est intermédiaire (2,7 et 1,9 Å sur a et c, respectivement). Cette augmentation s'applique de manière conséquente à la proportion en solvant contenue dans le cristal.

	GlnRS	GlnRS-QSI	GlnRS-Gln	GlnRS-ΔC	GlnRS-QOL
Collection des données					
Source	ID14-4	ID14-4	ID14-1	ID-29	ID14-1
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_1$	$P2_12_12_1$
Paramètres de maille a, b, c (Å)	74,1 - 95,9 - 115,7	83,3 - 95,3 - 120,2	81,7 - 95,3 - 120,6	53,7 - 84,2 - 73,9	76,8 - 95,8 - 117,6
- α, β, γ (°)	90 - 90 - 90	90 - 90 - 90	90 - 90 - 90	90 - 90 - 90	90 - 90 - 90
Limites de résolution (Å)	2,3 - 30	2,16 - 30	2,5 - 30	1,89 - 30	2,37 - 40
No. d'observations	209589	484900	134576	378626	178164
No. de réflexions uniques	36991	51572	31237	51729	35396
Haute résolution (Å)	2,3 - 2,36	2,16 - 2,22	2,5 - 2,54	1,89 - 2	2,37 - 2,51
Complétude (%)	98,1 (92,9) ^a	99,3 (99,4) ^a	99,7 (99,7) ^a	98,2 (89,2)	98,6 (92,8)
Redondance	5,6	9,4	4,3	7,3	5
R _{merge} (%)	6,2 (25,0) ^a	5,6	7,2	5,8 (51,9)	7,6 (31,5)
I/σ (I)	18,1 (5,5) ^a	19,6 (3,5) ^a	12,2 (2,6) ^a	21,5 (3,5) ^a	15,0 (3,9)
Contenu en solvant (%)	44,1	51,4	50,6	51,1	46,9
Mosaïcité (°)	0,5	0,42	0,36	0,17	0,25
Affinement					
Facteur R (%)	20	23,9	23,7	23,5	23,3
Facteur R libre (%)	24,2	26,4	26,5	26,8	28
Proteine	4488	4443	4503	-	-
Molécules d'eau	160	139	130	-	-
Facteur B moyen (Å ²)	41,7	49,4	55,7	-	-
Facteur B ligand (Å ²)	-	-	-	-	-
RMSD liaisons (Å), angles (°)	0,009 ; 1,5	0,01 ; 1,4	0,007 ; 1,3	-	-

^a Limite Haute de résolution

Tableau II. 8 : Données cristallographiques des structures de la GlnRS sous forme native, co-cristallisée en présence des analogues de l'adénylate (GlnRS-QSI et GlnRS-QOL), de glutamine (GlnRS-Gln), ou de la forme tronquée en C-terminal de la GlnRS (GlnRS-ΔC). Les statistiques concernant les données hautes résolutions sont précisées entre parenthèses. Les enregistrements ont été effectués à l'ESRF-Grenoble.



	[Proteine] mg/mL	[Substrat] mM	Agents cristallisants		pH	Gel d'Acarose	Technique
GlnRS + QSI	3.3	QSI 2.5mM	15% PEG 3350	200 mM Fluorure de Sodium	pH 7	0.2%	Batch
GlnRS + QOL	4	QOL 2.5mM	16% PEG 3350	200 mM Ammonium formate	pH 7		Batch
GlnRS + Gln	3.3	Glutamine 2.5mM	16% PEG 3350	200 mM Ammonium formate	pH 8.2	0.2%	Batch
GlnRS- ΔC	5		20% PEG 3350	200 mM di-Ammonium citrate			Diffusion de vapeur

Figure II. 9: Illustration des structures obtenues depuis le cristal au cliché de diffraction à la structure tridimensionnelle. **A ,B :** GlnRS en présence d'analogues de l'intermédiaire réactionnel adénylate (QSI, **A** ou QOL, **B**). **C :** GlnRS cristallisée en présence de glutamine. **D :** GlnRS dépourvue de son module C-terminal additionnel (GlnRS-ΔC).

C. Analyse des nouvelles données structurales

1. GlnRS-*Dr* cristallisée en présence d'analogue d'adénylate, QSI

La structure de la GlnRS co-cristallisée en présence d'adénylate a été résolue par remplacement moléculaire avec la GlnRS sous sa forme apo (PDB_id : 2HZ7). La carte de densité électronique nous a permis de placer l'analogue QSI dans le site actif de l'enzyme, mais la présence de cet additif n'a pas permis de stabiliser les régions flexibles de la GlnRS-*Dr* et, tout comme dans la structure de la protéine sous forme apo, l'extrémité C-terminale n'est pas visible dans la densité électronique donc flexible dans le cristal.

Description générale

La fixation de l'adénylate dans le site actif de l'enzyme provoque des modifications conformationnelles du cœur de la GlnRS. En particulier, le domaine ABD, dont on a observé chez *E. coli* les modifications de conformation en réponse à la fixation de l'ARNt et des petits substrats, subit un mouvement de rotation en direction du site actif. Ce mouvement s'applique également au domaine distal de liaison à l'anticodon. La superposition des structures libre de *D. radiodurans* et complexée d'*E. coli* illustre ce mouvement global de l'enzyme (voir [Figure II. 10](#)). Globalement la structure présente des rmsd comparables pour l'enzyme d'*E. coli* seule et en complexe (1,7 et 1,6 Å), et proche de l'enzyme natif (rmsd 0.8 Å). Pour *D. radiodurans*, les alignements domaine à domaine des structures native et en présence de QSI donnent des scores faibles compris entre 0,25 et 0,45 Å avec un maximum pour le domaine ABD du site catalytique et DBB du site de liaison à l'anticodon. Les détails des mouvements locaux ont été calculés pour chaque domaine isolé (voir aussi [Figures II. 20 et 21](#)).

L'adaptation induite par la fixation d'un petit substrat ou d'un analogue n'est pas particulière à la GlnRS-*Dr* car elle est observée dans la région du domaine de Rossmann pour d'autres synthétases de classe I (GluRS, ArgRS, TyrRS, TrpRS). Plus localement, l'adénylate accueilli dans le site actif ne permet pas de positionner la région du KMSKS dans sa conformation productive. Comme son homologue chez *E. coli* (et les autres synthétases dont l'activation est ARNt dépendante), la GlnRS-*Dr* n'adopterait cette position que lorsque l'ARNt se lie à l'enzyme.

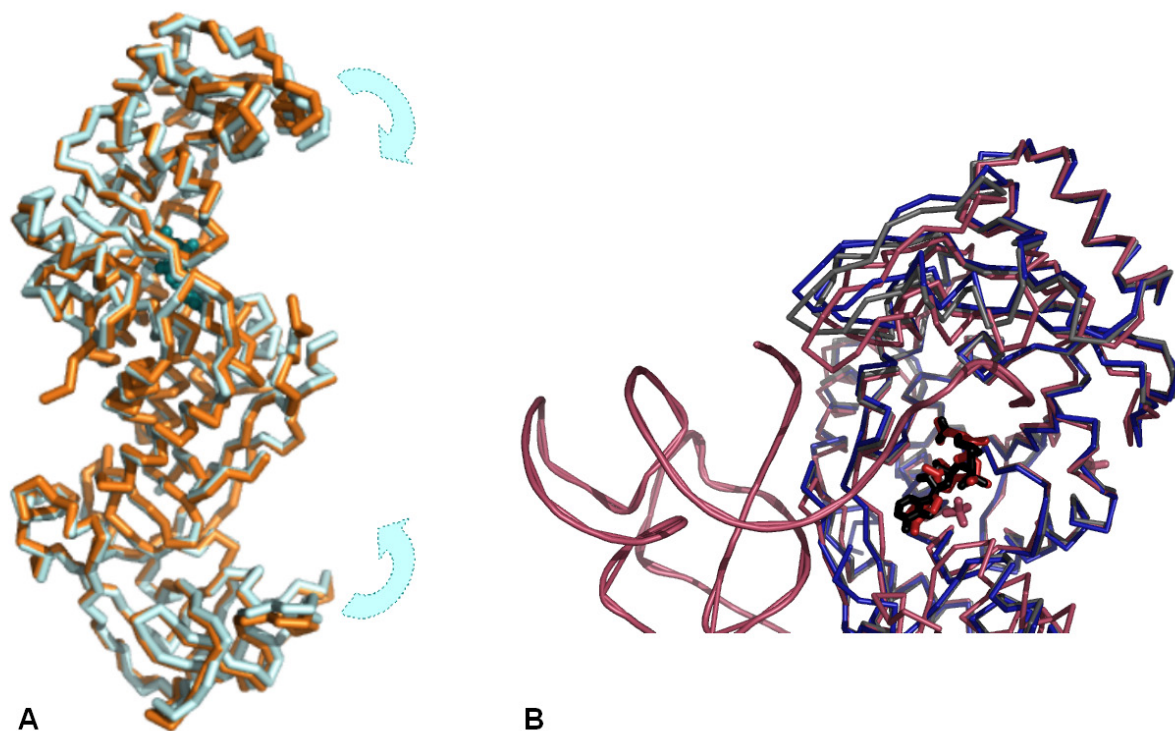


Figure II. 10 : A : Superposition des structures de la GlnRS-Dr sous forme apo ou en présence de QSI. Les régions les plus variables sont situées dans le domaine ABD. Le rmsd global pour l'alignement des deux structures est de 0,8 Å.

B : Vue rapprochée de la superposition des structures de la GlnRS-Dr / QSI (bleu) et la GlnRS-Ec native (gris) ou en complexe avec QSI/ARNt^{Gln} (pourpre). La partie AMP de l'analogue dans la structure de la GlnRS-Dr est décalée par rapport à celui de la GlnRS-Ec en complexe. Le RMSD global est de 1,6 Å pour les 2 structures.

Site actif de l'enzyme et contacts spécifiques avec le substrat

Les résidus en contact avec le QSI ont été comparés à ceux trouvés dans les structures résolues de la GlnRS-Ec/ARNt^{Gln} / QSI (PDB_id : 1EUQ) ou la GlnRS-Ec/ARNt^{Gln}/Gln-AMP (PDB_id : 1O0B). La plupart des résidus fixant l'adénylate dans le site actif sont conservés pour les GlnRS de *E. coli* et *D. radiodurans*. Cependant, et sans doute du fait de l'absence d'ARNt, certains contacts observés entre les substrats QSI ou ATP-Gln et la GlnRS-Ec, en présence d'ARNt, manquent dans le complexe GlnRS-Dr /QSI. En particulier, la zone correspondant à l'AMP "perd" certains contacts. Cet effet est majoritairement lié au fait que la boucle du KMSKS ne soit pas dans sa conformation productive car elle est impliquée dans la stabilisation de la région AMP. L'absence de deux interactions spécifiques de cette région contribue à déstabiliser la région AMP de l'analogue dans le site actif (voir Figure II. 11). La base A76, quant à elle, est responsable chez *E. coli* du maintien de la glutamine dans le site actif.

Position de l'adenylate dans le site actif

Afin de poursuivre l'analyse, nous avons comparé la position de l'adénylate dans le site actif des GlnRS-Dr et -Ec. Comme nous l'avons précisé précédemment, en dehors des domaines d'insertions de la GlnRS-Dr, les domaines catalytiques des 2 enzymes libres sont proches (Les RMSD du domaine ABD pour la GlnRS-Ec libre ou en complexe sont de 0,75 et 0,85 Å, respectivement). L'orientation des substrats dans le site actif de l'enzyme d' *E. coli* en complexe avec l'ARNt^{Gln} (GlnRS-Ec / ARNt^{Gln}/ - QSI ou - AMP-Gln) ont été comparés avec celle observée chez *D. radiodurans* car on ne dispose pas de structure sans l'ARNt. Les positions des chaînes latérales de la région Gln de QSI des structures d'*E. coli* et *D. radiodurans* sont comparables (distances inter-atomiques de l'ordre de 0.8 Å). Celles-ci augmentent au fur et à mesure que l'on atteint la partie AMP de l'analogue : ainsi on observe des distances allant de ~1,5 Å pour le ribose à ~2,5 Å pour la base. En réponse à l'absence de certains contacts due à la conformation relâchée du site actif, les positions des QSI sont voisines dans la région 'glutamine' mais décalées dans la région 'AMP' (voir Figures II. 10, II. 11, Tableau II. 12).

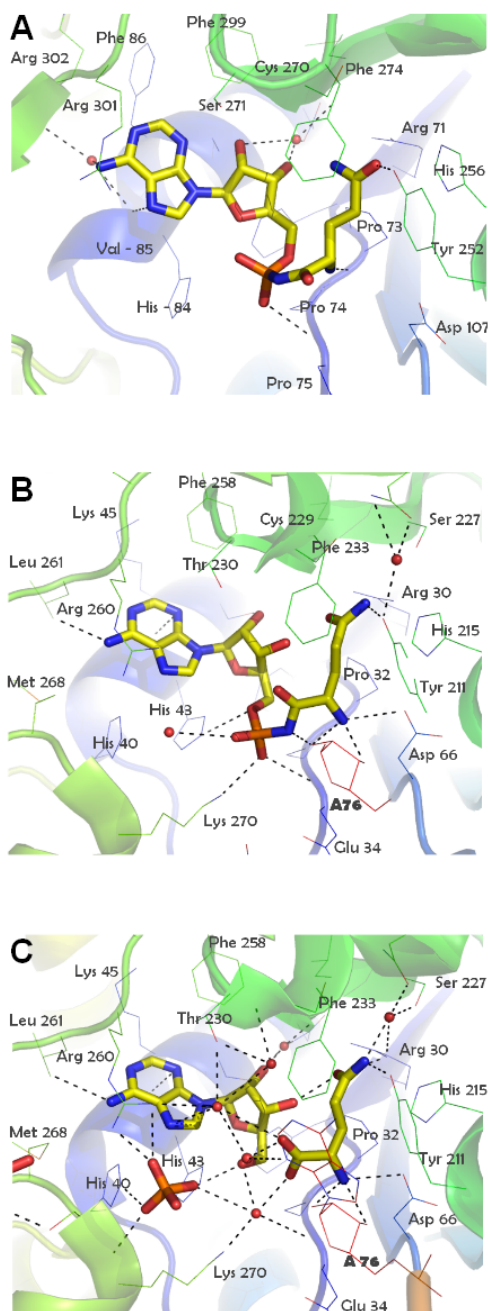


Figure II. 11 : Comparaison des contacts entre le ligand et le site actif de la GlnRS-Dr et Ec :
 (A) GlnRS-Dr et QSI , (B) GlnRS-Ec, ARNtGln et QSI (PDB_id : 1EUQ), (C) GlnRS-Ec, ARNtGln, Gln et AMP (PDB_id : 1O0B). Les résidus impliqués sont donnés et les contacts illustrés en pointillés. La base A76 de l'extrémité 3' de l'ARNtGln intervenant dans la stabilisation des substrats est indiquée en gras.

Des observations comparables ont été faites entre les structures de la GluRS de *T. thermophilus* résolues sous forme libre, co-cristallisée soit avec du glutamate et l'ATP (PDB_id : 1J09), l'ATP (seul/ARNt^{Glu}, PDB_id : 1N75/1N77) ou l'analogue du Glu-AMP et ARNt^{Glu} (PDB_id : 1N78). Dans ce cas, la chaîne latérale de l'acide aminé est maintenue dans une position similaire qu'il soit lié ou non à l'ARNt homologue. Par contre, l'ATP admet deux conformations différentes selon que l'ARNt est présent ou non. Les auteurs ont proposé l'existence d'une position non productive de l'ATP en réponse à la conformation du site actif. Dans ce cas, la GluRS, l'ATP et le glutamate forment un complexe qui prévient la formation du Glu-AMP en absence d'ARNt (Sekine *et al.*, 2003). La proximité conformationnelle des GlnRS-*Ec* et -*Dr* et la conservation des résidus impliqués dans fixation des substrats d'aminoacylation nous autorisent à établir un parallèle entre les conformations des sites actifs en présence des différents ligands. Comme il l'a été proposé pour la GluRS de *T. thermophilus* et la GlnRS de *E. coli* (Sherlin et Perona, 2003), la GlnRS de *D. radiodurans* pourrait présenter deux conformations distinctes : une conformation productive en complexe avec l'ARNt et une conformation non productive qui empêche la formation du Glu-AMP. Il a été proposé que ce mode de structuration en réponse à la fixation des substrats a été acquis pour éviter l'éventuel manque en source d'énergie dans un contexte où le Glu-AMP est instable chimiquement et où la concentration intracellulaire de glutamate (et de glutamine) est élevée chez de nombreux eucaryotes et procaryotes (Danbolt, 2001).

Comme c'est les cas pour la GluRS de *T. thermophilus*, la chaîne latérale de l'acide aminé peut-être positionnée correctement dans le site actif de l'enzyme en absence d'ARNt homologue. La base A76 de la tige acceptrice jouerait néanmoins un rôle primordial chez la GlnRS d'*E. coli* : le site actif est en partie formé par l'empilement de la base A76 avec 2 résidus (*Ec* : Phe233 et Tyr211 / *Dr* : Phe274 et Tyr252) impliqués dans l'interaction avec la glutamine (Bullock *et al.*, 2003). En absence d'ARNt chez *D. radiodurans*, ces 2 résidus interagissent avec la chaîne latérale de la glutamine et sont impliqués dans son maintien. La base A76 intervient probablement, comme pour la GlnRS-*Ec*, dans le positionnement du groupement carboxyle de la glutamine (effectivement en décalage par rapport à celui de la GlnRS-*Ec*/QSI/ARNt^{Gln}).

En conclusion, le passage à la conformation productive du site actif devrait être provoquée, comme pour les autres GlnRS, par la présence d'ARNt et serait destinée, comme pour la GluRS de *T. thermophilus* et pour la GlnRS complexées d'*E. coli*, à orienter l'ATP dans sa position correcte pour autoriser le déroulement de la réaction d'aminoacylation, en particulier par le changement de conformation de la boucle du KMSKS directement impliquée dans sa stabilisation. Remarquons que la forme GlnRS/QSI est un état artificiel de l'enzyme puisque l'acide aminé activé n'existe normalement que lorsque l'ARNt est complexé à l'enzyme.

2. Comparaison avec la position du glutaminol dans le site actif

Les données cristallographiques préliminaires montrent que ce second analogue de l'intermédiaire réactionnel est également décalé par rapport à l'orientation de l'analogue chez *E. coli*. Le groupement AMP est ici aussi dans une conformation 'non productive' tout comme la boucle du KMSKS. La position des atomes de la chaîne latérale de la glutamine est globalement similaire à celle du QSI mais on observe une modification de la position de la base de l'AMP, qui présente une orientation encore plus relâchée (voir [Tableau II. 12](#)). Cette observation peut-être expliquée par la conformation globale de l'enzyme, qui, quant à elle est moins contractée en direction du site actif. Elle est, en effet, plus proche de la position de l'enzyme natif dans la région du domaine ABD et PBB : le rmsd global entre les 2 structures cristallisées avec des analogues est de 0,87Å contre 0,34 Å avec la structure de l'enzyme native.

	Distances (Å)		Moyenne	
	QOL – 1EUQ	QSI – 1EUQ		
N1	2,54	2,41	Base	QSI+QOL 1,98
C2	2,49	2,74		
N3	2,25	2,34		
C4	2,05	1,53	QOL	2,15
C5	2,08	1,43		
C6	2,34	1,80		
N6	2,40	1,90	QSI	1,82
N7	1,84	1,53		
C8	1,68	1,32		
N9	1,79	1,21	Ribose	QSI+QOL 1,53
C1'	1,67	1,37		
C2'	1,2	1,48		
O2'	1,41	2,39	QOL	1,33
C3'	1,19	1,53		
O3'	0,94	1,85		
C4'	1,31	1,32	QSI	1,72
O4'	1,61	2,17		
C5'	1,29	1,95		
O5'	1,34	1,44	Phosphate	1,51
P/S	1,29	1,72		
O/N 10	0,96	0,97		
C	0,75	1,28	Chaîne latérale	QSI+QOL 0,77
CA	0,55	0,56		
CB	0,48	1,19		
CG	0,52	1,26	QOL	0,66
CD	0,63	0,61		
OE1	0,75	0,73		
NE2	0,66	0,49	QSI	0,89

Tableau II.12 : Tableau des distances interatomiques (en Å) des analogues d'adénylate chez *D. radiodurans* (QSI et QOL) et chez *E. coli* (QSI, 1EUQ). La moyenne des distances par région du substrat est donnée dans le panneau de droite pour la partie AMP (**Base**, **Ribose**, **Phosphate**) et la partie glutamine (**Chaîne latérale**).

3. GlnRS-Dr en présence de glutamine : la surprenante conformation du site actif

La recherche de nouvelles conditions de cristallisation a conduit à l'obtention de cristaux en présence de glutamine, substrat naturel de la GlnRS-Dr *in vivo*. Ceci a permis de stabiliser une insertion (**ins2**) d'une trentaine de résidus branchée au domaine de liaison à l'anticodon et qui, dans la protéine repliée, est située à proximité du site actif de l'enzyme. La résolution de la structure de l'enzyme native ne permettait pas de distinguer cette boucle probablement en mouvement dans le cristal, ou adoptant un ensemble de conformations distinctes (Figure II. 2).

Description générale

En présence de glutamine dans le milieu de cristallisation, la boucle **ins2** est clairement résolue et induit une conformation encore jamais observée du site actif que l'on appellera fGlnRS (f pour fermée). Lorsqu'elle est sous cette forme, la boucle supplémentaire s'insinue profondément dans la poche du site actif, comme si elle le verrouillait (Figure II. 13.A).

La densité électronique de la région du site actif n'offre pas d'indication formelle de la présence de glutamine. Seule une région en marge du site de fixation d'*E. coli* donne une densité électronique (Fo-Fc) qui pourrait être le site effectivement occupé par l'acide aminé. Les superpositions de la GlnRS-Dr en conformation fermée (fGlnRS-Dr) avec la structure de la GlnRS-Ec/ARNt/Gln-AMP (PDB_id : 1O0B), d'une part, et la GlnRS-Dr complexée à l'analogue de l'adénylate (GlnRS-Dr/QSI), d'autre part, nous ont permis d'évaluer l'espace disponible pour les différents substrats dans la région du site actif. La fGlnRS-Dr ne laisse qu'une place restreinte dans la région d'activation de l'acide aminé et n'autorise pas, sans gêne stérique, l'incorporation de la glutamine dans sa position active. Celle-ci ne pourrait occuper le site que s'il s'opère un changement de position de la boucle (fermée à ouverte) ou qu'elle occupe une autre place dans le site actif (Figure II. 13.B).

En complément, sa conformation fermée interdit toute liaison de l'ATP dont le site de fixation est entièrement occupé par les résidus S543-A546. De manière intéressante, nous avons pu observer qu'une autre partie de la boucle semble mimer l'extrémité CCA de l'ARNt^{Gln}. En effet, les **résidus 553 à 558** sont positionnés de façon similaire à l'extrémité 3' CCA de l'ARNt. A l'exception de la région d'interaction du CCA, les sites de liaison à l'ARNt sont disponibles. L'insertion **ins2** de la GlnRS-Dr adopte donc 2 conformations distinctes : une conformation ouverte ou dynamique, en absence de glutamine, et fermée et stable en sa présence. Notons également que la boucle **ins2** pourrait représenter un lien supplémentaire, qui s'ajoute au module intermédiaire (HSD ou Helical subdomain), entre les modules fonctionnels de la GlnRS (modules catalytique et de liaison à l'anticodon. En interagissant directement, elle renforcerait la communication entre eux.

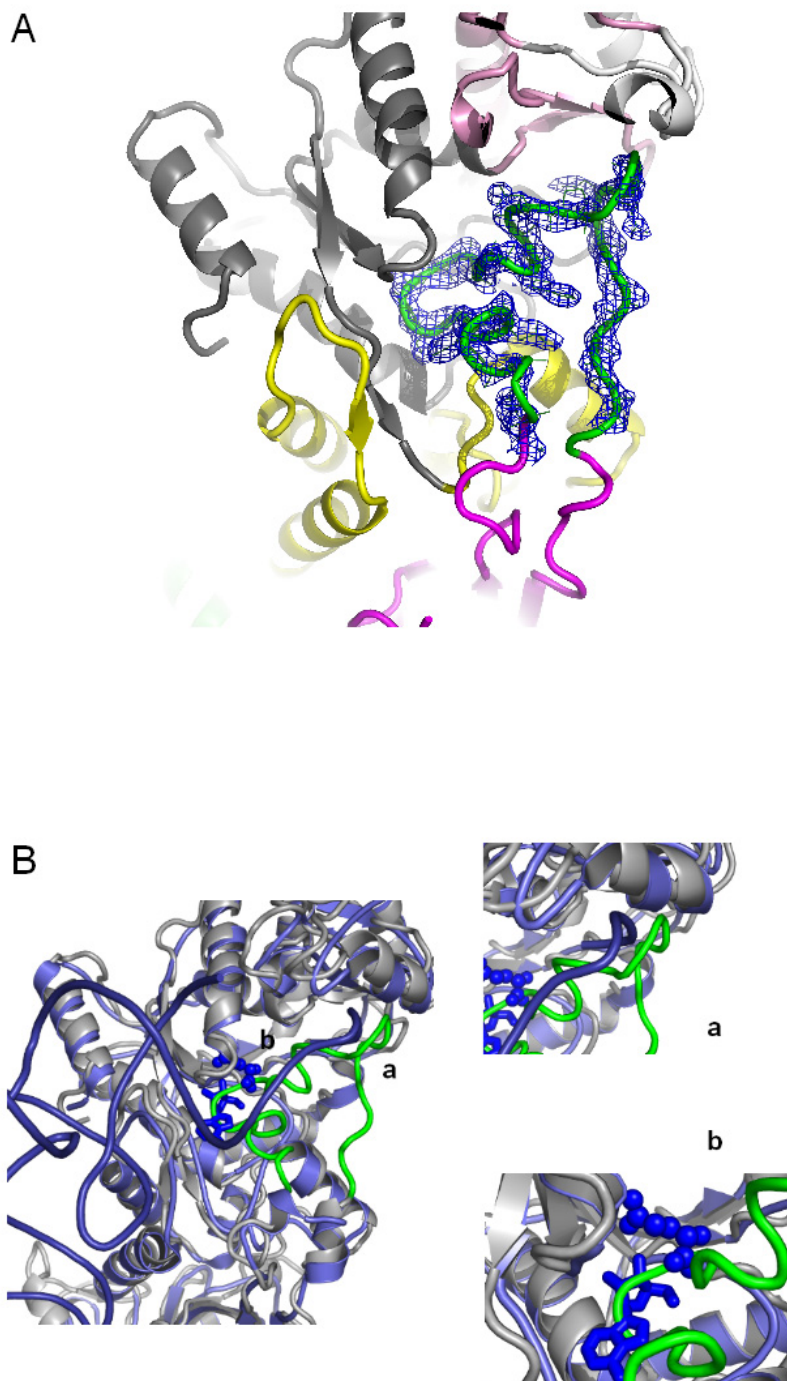


Figure II. 13 : Vue rapprochée du site actif de la GlnRS-Dr dans sa conformation fermée. Celle-ci est observée lorsque la GlnRS-Dr est co-cristallisée en présence de glutamine. Les codes couleurs des domaines de la GlnRS sont identiques à ceux de la **figure II. 2**. La boucle d'insertion est représentée en vert clair dans toutes les figures. **(A)** Représentation de la conformation fermée du site actif de l'enzyme. La carte différence (Fo-Fc) de la densité électronique de la boucle est contourée à $2,5 \sigma$ et représentée en bleu; **(II)** superposition de la structure de la GlnRS-Dr (gris clair) avec celle d'*E. coli* (bleu électrique) complexée à l'ARNt^{Gln}, la glutamine et l'AMP. **(B, a)** : vue rapprochée de la zone de l'extrémité 3' CCA de l'ARNt (bleu foncé) et de la boucle du site actif GlnRS-Dr (vert clair). **(B, b)** : vue rapprochée du site de reconnaissance de l'acide aminé et visualisation de la boucle d'insertion dans cette zone (PDB_id : 1O0B).

Boucle		En contact avec GlnRS-Dr			Chez Dr	Chez Ec	3'CCA-ARNt ^{Gln} – Ec					
Résidu	Atome	Résidu	Atome	Distance	Résidu	Résidu		N	A	Dist		
GLU 533	N	PRO 532	CA	2,43								
			CB	3,27	MET 224	ILE 183	CD1	A 72	O4*	3,44		
			C	1,33			CG2			3,4		
			O	2,24	ARG 172	ARG 133	NE		O3*	3,01		
		HIS 531	C	3,16			NH2			3,44		
			PRO 532	N			2,81	CD	O2*	3,13		
				CG			3,38	NE		2,89		
			HIS 531	O			3,16	GLU 169	ARG 130	CD	G 73	O2P
	CA	PRO 532	C	2,42			CG			3,3		
	CG	LYS 313	NZ	3,44	ARG 172	ARG 133	NH2	C 74			2,91	
			NZ	3,37	GLU 168	ILE 129	CG2		O1P	3,38		
			NZ	2,89	ARG 172	ARG 133	NH2				2,86	
			ILE 308	CA	3,35				NH1	O2P	3,25	
		THR 309	N	3,03	NH2					3,44		
		C	PRO 532	C	3,12				ALA 251	MET 210	CE	C5*
				O	3,33	ARG 233	ARG 192		CB	O2	3,5	
				C	3,48	GLY 209	GLY 168		CA	N3	3,15	
		O	HIS 531	O	3,03				C			3,48
GLU 533				N	3,45				O		3,25	
ALA 534	N		CA	2,43	PRO 167	PRO 126	CG		C4	3,46		
			CB	3,4	GLY 209	GLY 168	O			3,38		
			C	1,33	VAL 165	LEU 124	O		N4	2,67		
			O	2,24	GLY 209	GLY 168	O			2,71		
			CA	C	2,42	VAL 235	LYS 194	NZ	C 75	O2P	2,87	
				O	2,72	ARG 233	ARG 192	NE			2,74	
			C	C	3,06			CZ			3,26	
				O	2,99			NH2		3,02		
			O	C	2,97			VAL 235	LYS 194	CE		
					O	2,53	ALA 251	MET 210	CE		C5*	3,44
			ALA 536	O	LYS 313	NZ	3,39	ASN 110	ASN 69	ND2		O3*
ASP 540	OD2		CE	3,22	ARG 233	ARG 192	NH2			2,97		
			NZ	3,39	ASN 110	ASN 69	ND2		O2*	2,61		
			CA	THR 309	O	3,4	ARG 233	ARG 192	NH2	A 76	P	3,34
			CG		CB	3,47	THR 109	THR 68	CB		O1P	3,43
			OD1		OG1	3,29			OG1			2,89
					CB	3,11	ARG 233	ARG 192	NH2			2,83
					OG1	3,39	TYR 252	TYR 211	N		O2P	2,98
					O	3,09			CB			3,46
			OD2	ARG 301	NE	2,86	ASP 75	GLU 34	OE1		C4*	3,3
			N	THR 309	O	3,01	PHE 274	PHE 233	CE1		N1	3,37
						CG	LYS 311	NZ	3,29			CZ
OD1	HIS 81	NE2				2,83	CE1		3,25			
	HIS 84	NE2				3,2						
THR 309	O	3,22										
OD2	LYS 311	NZ				2,93					N3	3,32
						N	2,89					
						CB	3,44					
						CG	3,34					
		O				HIS 84	CD2	3,23				
							NE2	3,02				
SER 543	CB	GLY 83	O	3,06								
			OG		CA	3,2						
					C	3,35						
					O	3,27						
			O	HIS 84	CD2	3,28						
ALA 546	CB	TYR 252	CE2	3,39								
			O	ASP 75	OD1	3,17						
GLU 548	CG	GLU 273	CB	3,49								
GLU 550	OE2	ARG 312	CD	3,36								
			NH1	2,77								
GLU 552	OE1	LYS 225	CE	3,47								
PRO 553	CG	ARG 233	NH2	3,44								
ALA 555			CZ	3,37								
			NH2	2,87								
GLU 558		ASN 110	ND2	3,15								
			CG	3,41								
THR 559		GLU 112	OE1	2,98								
		ASN 110	OD1	2,89								
ALA 560		ARG 312	NH1	3,02								
			NH2	3,28								

Tables II. 14 : A - Tableau des contacts intramoléculaires entre la boucle *ins2* et le reste de la GlnRS-Dr. Les distances sont données en Å avec un seuil de 3,5 Å de distance. **B - Tableau des contacts intramoléculaires entre l'extrémité 3' de l'ARNt^{Gln} et le site actif de la GlnRS-Ec.** L'extrémité CCA (C74 - C75 - A76) est notée en gras.

Contacts spécifiques de la boucle *ins2* et variation des positions des résidus impliqués dans la reconnaissance des substrats

Les contacts entre la boucle *ins2* et le reste de la protéine ont été comparés aux contacts établis entre les différents substrats des GlnRS-*Ec* et -*Dr*. Les structures de la GlnRS-*Ec* seule (PDB-id : 1NYL) ou en complexe avec l'ARNt (PDB-id : 1EUQ) ainsi que la structure de la GlnRS-*Dr* seule (PDB-id : 2HZ7) ou liée à l'analogue de l'adenylate (GlnRS-*Dr* /QSI) ont été utilisées.

Une première approche a été engagée suite à l'observation de la ressemblance d'une partie de la boucle avec la conformation de l'extrémité 3'CCA de l'ARNt^{Gln} dans la structure en complexe d'*E. coli*. Les résidus contactés chez *E. coli* en présence d'ARNt^{Gln} sont tous différents des résidus contactés lorsque la boucle ferme le site actif à l'exception de la Tyr252 (-*Dr* / Tyr211-*Ec*) responsable, chez *E. coli*, du positionnement de A76 de l'extrémité 3' de l'ARNt et qui contacte la chaîne latérale de la glutamine chez *E. coli* et *D. radiodurans*. La plupart des résidus impliqués dans la discrimination de l'ARNt ne sont pas en contact direct avec *ins2* (Tableau II. 14).

Les 2 motifs consensus KMSKS et HIGH entrent directement en contact avec la boucle *ins2*. La comparaison des positions des atomes avec ceux des structures sus-citées nous a permis d'observer que les positions des Cα du KMSKS sont plus proches de celles de la GlnRS-*Ec* complexée à l'ARNt que celles de la GlnRS-*Ec* native, GlnRS-*Dr* native ou en présence d'adénylate. Il en est de même pour les chaînes latérales qui adoptent la même orientation à l'exception de l'Arginine 312 (TSKRK) (voir Figure II. 15 et 16, Tableau II. 17). Les chaînes latérales de 3 des 4 résidus du motif HIGH sont également impliquées dans les contacts directs de *ins2*. Le motif HIGH de la GlnRS-*Ec* (et des classes I en général) est spécifiquement impliqué dans la reconnaissance et l'interaction avec l'ATP. Le passage de la conformation ouverte à fermée induit un déplacement du motif HIGH en direction du site actif. De plus, des résidus en contact direct avec *ins2* sont, pour leur équivalent chez *E. coli*, impliqués dans la reconnaissance de l'ATP comme R301, K311, D75 et opèrent un changement de position par rapport à la conformation ouverte. D'autres modifications concernent la région 271-276 où S271 positionne l'ATP, et où les résidus 274,275 et 276, conservés chez *E. coli* sont impliqués dans la reconnaissance ou la discrimination de l'ARNt. Dans cette région, le résidu E273 présente une position retrouvée uniquement dans la structure native de la GlnRS d'*E. coli*. E273 est le seul résidu de cet enchaînement qui, chez *E. coli*, n'est pas impliqué dans des fonctions de discrimination ou de reconnaissance. Il est à la charnière de 2 régions de reconnaissance des substrats impliqués dans la réaction d'aminacylation. E273 pourrait être impliqué dans le changement de conformation de la boucle avec laquelle il est en contact direct. Les résidus impliqués dans la reconnaissance spécifique de l'ARNt^{Gln} sont eux disponibles et prêts à l'accueillir pour l'activation de l'acide aminé.

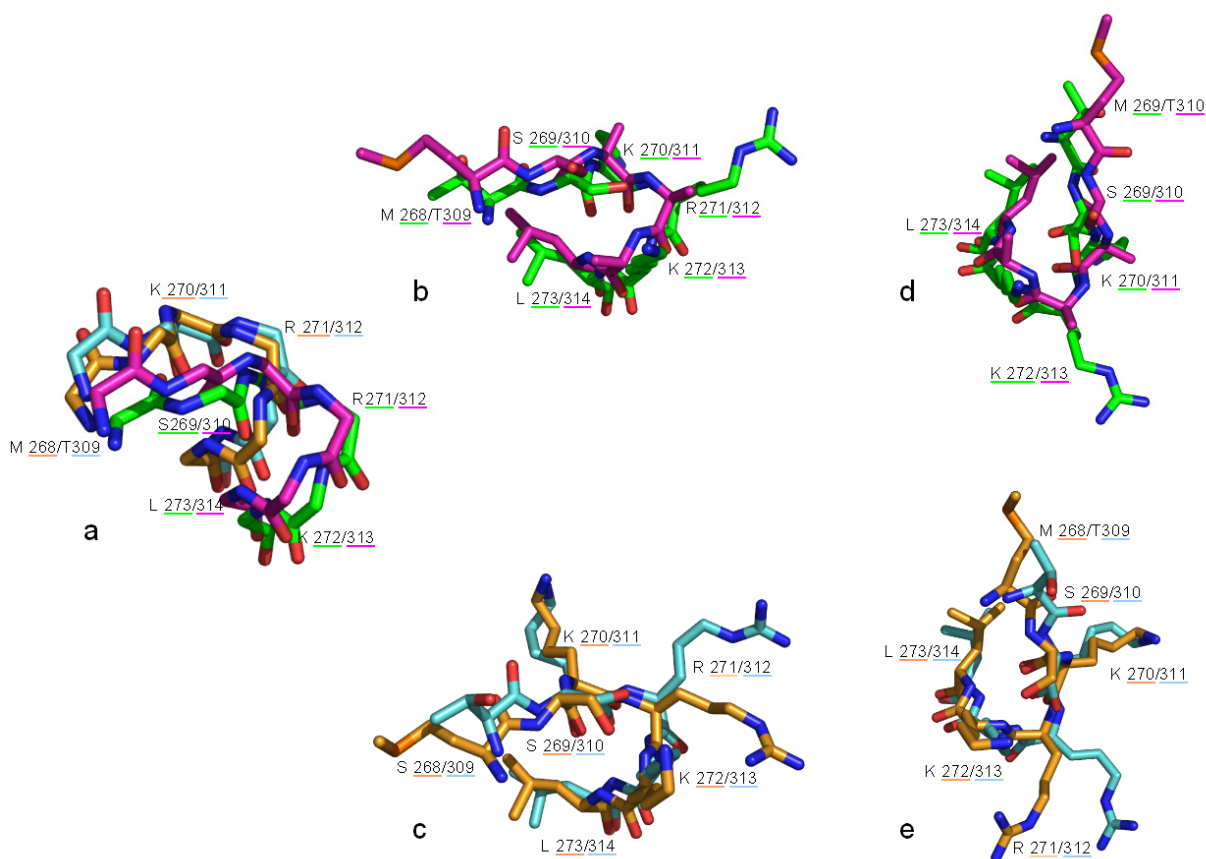
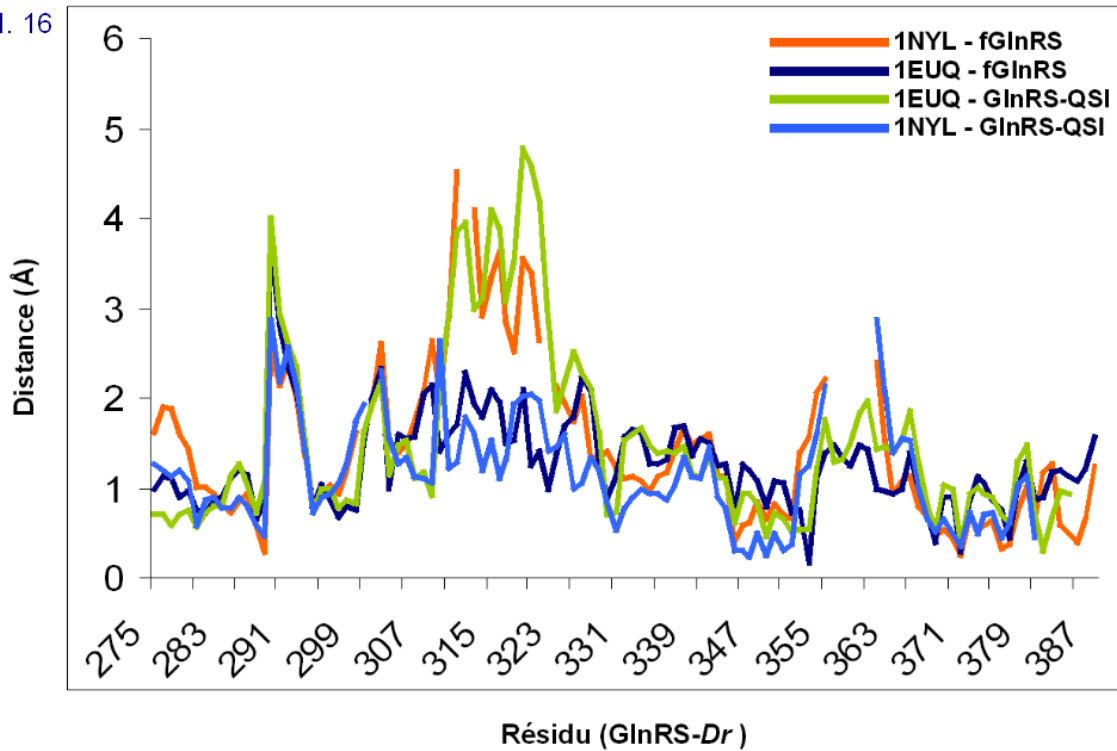


Figure II. 15 : Superposition des KMSKS des GlnRS-*Dr* et *Ec* sous forme libre ou complexée à leurs substrats. Les structures représentées sont : en rose GlnRS-*Ec* native (PDB_id : 1NYL), en orange clair GlnRS-*Ec*/ARNt^{Gln}/QSI (PDB_id : 1EUQ), en bleu clair fGlnRS-*Dr*, en vert GlnRS-*Dr*/QSI – a : Représentation des Cα de la région ^{268/309}-**MSKRKL**-^{273/314} dont les 3 premiers résidus représentent la séquence consensus des GlnRS pour les 4 structures. – b : Représentation de la région **MSKRKL** pour la GlnRS-*Ec* native (rose) et la GlnRS-*Dr*/QSI (vert) avec leurs chaînes latérales. – c : Représentation de la région **MSKRKL** pour la GlnRS-*Ec*/ARNt^{Gln}/QSI (orange clair) et la GlnRS-*Dr*/QSI (bleu clair) avec leurs chaînes latérales – d/e. Même représentation qu'en b/c avec une rotation de 90° sur l'axe des z.

II. 16



II. 17

		Distance (Å)						Distance (Å)			
		EC:NAI - Dr:Gln	EC:ARNtGln - Dr:Gln	EC:ARNtGln - Dr:OSI	EC:Dr:OSI			EC:NAI - Dr:Gln	EC:ARNtGln - Dr:Gln	EC:ARNtGln - Dr:OSI	EC:Dr:OSI
GUN 234	VAL 275	1,62	0,98	0,7	1,27	PRO 291	PRO 332	1,12	1,65	1,59	0,89
ASP 235	ASP 276	1,9	1,12	0,7	1,19	THR 292	THR 333	1,08	1,63	1,68	1
ASN 236	ASN 277	1,88	1,1	0,6	1,13	ILE 293	LEU 334	0,97	1,28	1,48	0,95
ARG 237	ARG 278	1,6	0,9	0,7	1,2	SER 294	ARG 335	1,13	1,27	1,38	0,93
ARG 238	ALA 279	1,44	0,97	0,76	1,09	GLY 295	ALA 336	1,18	1,32	1,42	0,87
LEU 239	ILE 280	1,01	0,75	0,66	0,6	LEU 296	GLN 337	1,43	1,66	1,39	1,04
TYR 240	TYR 281	1,01	0,74	0,72	0,88	ARG 297	ARG 338	1,67	1,69	1,46	1,34
ASP 241	ASP 282	0,85	0,67	0,79	0,9	ARG 298	ARG 339	1,49	1,36	1,15	1,12
TRP 242	TRP 283	0,82	0,9	0,8	0,77	ARG 299	LEU 340	1,56	1,55	1,12	1,11
VAL 243	LEU 284	0,74	1,1	1,16	0,78	GLY 300	GLY 341	1,59	1,51	1,33	1,46
LEU 244	MET 285	0,84	1,21	1,26	0,89	TYR 301	VAL 342	1,13	1,24	1,1	0,89
ASP 245	ALA 286	0,94	1,16	1,04	0,79	THR 302	THR 343	1,14	1,28	1,1	0,78
ASN 246	ALA 287	0,6	0,65	0,74	0,58	ALA 303	PRO 344	0,41	0,76	0,61	0,31
ILE 247	LEU 288	0,29	0,82	1,08	0,46	ALA 304	GLU 345	0,6	1,27	0,95	0,31
THR 248	ASN 289	2,55	3,66	4	2,87	SER 305	ALA 346	0,61	1,21	0,93	0,23
ILE 249	PHE 290	2,14	2,82	2,93	2,18	ILE 306	VAL 347	0,84	1,08	0,82	0,5
PRO 250	ASP 291	2,4	2,35	2,62	2,56	ARG 307	ARG 348	0,67	0,8	0,48	0,25
VAL 251	PRO 292	1,99	2,07	2,35	2,15	GLU 308	ALA 349	0,82	1,08	0,74	0,5
HIS 252	ARG 293	1,37	1,45	1,54	1,42	PHE 309	PHE 350	0,7	1,07	0,65	0,3
PRO 253	PRO 294	0,8	0,8	0,72	0,73	CYS 310	ALA 351	0,66	0,72	0,62	0,37
ARG 254	HIS 295	0,95	1,03	1	0,9	LYS 311	ALA 352	1,39	0,75	0,53	1,15
GUN 255	GLN 296	1,03	0,86	0,94	0,94	ARG 312	GLN 353	1,57	0,16	0,53	1,25
TYR 256	TYR 297	0,93	0,88	0,77	1,06	ILE 313	ILE 354	2,08	1,12	1,23	1,62
GLU 257	GLU 298	1,19	0,79	1,28	1,28	GLY 314	GLY 355	1,4	1,4	1,76	2,13
PHE 258	PHE 299	1,63	0,82	1,73	1,73	VAL 315	VAL 356	2,22	1,48	1,3	
	GLY 300		1,5	1,62	1,84				1,33	1,32	
SER 259	ARG 301	1,96	2	1,93		LYS 317	SER 357		1,24	1,5	
LEU 261	ARG 302	2,62	2,32	2,19	2,31	GLN 318	ARG 358		1,48	1,84	
ASN 262	GLY 303	1,52	1	1,15	1,55	ASP 319	ASN 360		1,44	1,96	
LEU 263	LEU 304	1,4	1,81	1,26	1,26	ASN 320	ARG 361			1,44	
GLU 264	GLU 305	1,48	1,55	1,53	1,33	THR 321	THR 362	1,54	0,97	1,46	2,06
TYR 265	TYR 306	1,75	1,57	1,1	1,14	ILE 322	VAL 363	0,95	0,94	1,41	1,39
THR 266	THR 307	2,06	2,05	1,18	1,1	GLU 323	ASP 364	1,08	0,99	1,52	1,55
VAL 267	ILE 308	2,63	2,15	0,92	1,07	MET 324	ILE 365	1,13	1,39	1,86	1,54
MET 268	THR 309	2,94	2,94	2,16	2,83	ALA 325	ALA 366	1,01	1,32	1,32	1,02
SER 269	SER 310	2,92	1,81	2,94	1,23	SER 326	VAL 367	0,89	0,74	0,79	0,67
LYS 270	LYS 311	4,51	1,71	3,87	1,3	LEU 327	TYR 368	0,46	0,39	0,55	0,51
ARG 271	ARG 312		2,29	3,95	1,8	GLU 328	GLU 369	0,55	0,89	1,03	0,65
LYS 272	LYS 313	4,1	1,94	3	1,81	SER 329	ASN 370	0,44	0,89	0,98	0,5
LEU 273	LEU 314	2,92	1,79	3,1	1,19	CYS 330	ALA 371	0,25	0,29	0,41	0,36
ASN 274	ARG 315	3,36	2,1	4,09	1,52	ILE 331	VAL 372	0,7	1,67	0,93	0,72
LEU 275	GLU 316	3,62	1,95	3,89	1,1	ARG 332	ARG 373	0,53	1,14	1,02	0,5
LEU 276	LEU 317	2,85	1,5	3,09	1,31	GLU 333	ASP 374	0,6	1,06	0,94	0,7
VAL 277	VAL 318	2,51	1,52	3,96	1,92	ASP 334	ASP 375	0,64	0,87	0,89	0,73
THR 278	GLN 319	3,55	2,09	4,78	2,02	LEU 335	LEU 376	0,32	0,76	0,7	0,45
ASP 279	GLY 321	3,39	1,24	4,58	2,04	ASN 336	ASN 377	0,37	0,44	0,62	0,62
LYS 280	ALA 322	2,63	1,42	1,98	1,98	GLU 337	HIS 378	0,75	0,98	1,29	1,04
VAL 282	VAL 323	2,14	1,39	1,95	1,46	ASN 338	ARG 379	1,02	1,3	1,49	1,14
GLU 283	SER 324	1,95	1,7	2,19	1,59	ALA 339	ALA 380	0,63	0,87	0,71	0,44
GLY 284	GLY 325	1,73	1,78	2,51	1	PRO 340	PRO 381	1,18	0,9	0,31	
TRP 285	TRP 326	2,02	2,24	2,38	1,07	ARG 341	ARG 382	1,27	1,17	0,69	
ASP 286	ASP 327	1,35	2,06	2,11	1,34	ALA 342	VAL 383	0,58	1,21	0,96	
ASP 287	ASP 328	1,31	1,29	1,63	1,19	MET 343	MET 384	0,5	1,12	0,95	
PRO 288	PRO 329	1,41	0,9	0,93	0,93	ALA 344	ALA 385	0,41	1,08		
ARG 289	ARG 330	1,22	1,12	0,74	0,54	VAL 345	VAL 386	0,65	1,23		
MET 290	MET 331	1,1	1,57	1,52	0,78	ILE 346	LEU 387	1,25	1,57		

Figure II. 16 : Représentation graphique des distances entre les Ca des résidus équivalents de la région du KMSKS des GlnRS-Dr en conformation ouverte ou fermée et des GlnRS-Ec native ou en complexe avec l'ARNt^{Gln}. On distingue : les distances interatomiques (Å) (i) entre la GlnRS-Ec native (PDB_id : 1NYL) et la fGlnRS sont représentées en orange; entre la GlnRS-Ec/ ARNt^{Gln}/QSI (PDB_id : 1EUQ) et la GlnRS-Dr/Gln en bleu foncé; entre la GlnRS-Ec/ ARNt^{Gln}/QSI et la GlnRS-Dr/QSI en vert clair; entre la GlnRS-Ec native et la GlnRS-Dr/QSI en bleu clair, en fonction des résidus tels qu'ils sont numérotés chez *D. radiodurans*.

Tableau II. 17 : Tableau des distances correspondantes. Les résidus équivalents chez *E. coli* sont donnés dans la première colonne. Le cadre rouge indique la région **MSKRK**.

Position de la boucle du KMSKS

Tous les contacts spécifiques de l'interaction avec l'ATP sont engagés dans la boucle d'insertion mais l'enzyme reste disponible pour contacter l'ARNt. En plus des modifications conformationnelles directement causées par la position de la boucle on observe une variation de la position de la boucle du KMSKS. Les comparaisons avec les autres structures de GlnRS-*Dr* et *Ec* nous montrent qu'elle est plus proche de l'orientation adoptée par la GlnRS-*Ec* complexée à l'ARNt^{Gln} que celle de l'enzyme libre ou en présence d'adénylate (GlnRS-*Ec* native, GlnRS-*Dr* native, GlnRS-*Dr*/QSI ou /QOL) (Figure II.16 et Tableau II.17).

Conclusions

D. radiodurans est, avec *D. geothermalis*, la seule GlnRS référencée présentant une insertion dans cette région de l'enzyme (Figure II.18). En présence de glutamine dans le milieu de cristallisation, la boucle **ins2** admet une conformation fermée. Ces observations montrent l'existence, chez *D. radiodurans*, d'un état intermédiaire durant l'étape d'activation ou à l'état natif. En effet, cette conformation peut-être soit engendrée par l'environnement chimique de la protéine, lié aux conditions de cristallisation, soit une conséquence directement liée à la présence de l'acide aminé spécifique. Rappelons qu'il est difficile de trancher quant à la présence ou non de la glutamine dans le site actif de l'enzyme. En effet, s'il est clair que la glutamine ne se place pas dans la conformation active retrouvée chez *E. coli*, un espace est disponible pour l'y accommoder dans une autre position au sein du site actif, mais la densité électronique observée en marge du site productif de fixation ne permet pas de conclure sur cette question.

Il est probable que cette conformation ne soit pas liée à l'empilement car aucun contact intermoléculaire n'intervient aux alentours de la boucle. On observe, dans ce cas, que la seule présence de glutamine dans le milieu entraîne un changement de conformation. Celle-ci est particulièrement illustrée par l'orientation "active" de la boucle du KMSKS que l'on retrouve dans la même position que celle observée dans le complexe GlnRS-*Ec*/ARNt^{Gln}. Le site actif est dans une conformation propice à l'activation de la glutamine mais ne peut fixer l'ATP tant que l'ARNt n'est pas reconnu spécifiquement par l'enzyme : l'absence d'ARNt dans le milieu couplée à la présence de glutamine induirait une conformation intermédiaire, en attente du substrat spécifique qu'est l'ARNt comme si on préparait le site actif à cette réaction tout en bloquant le site de fixation de l'ATP.

En effet, la majorité des résidus impliqués dans la reconnaissance spécifique de l'ARNt^{Gln} ne sont pas impliqués dans des contacts intramoléculaires de la fGlnRS. On peut supposer que l'ARNt peut-être reconnu spécifiquement par l'enzyme dans cette conformation à condition qu'elle entraîne l'ouverture coordonnée du site actif. Le passage d'une conformation à l'autre pourrait être facilitée par la position stratégique de la boucle à l'intermédiaire direct du site de reconnaissance de l'anticodon et du site actif de l'enzyme.

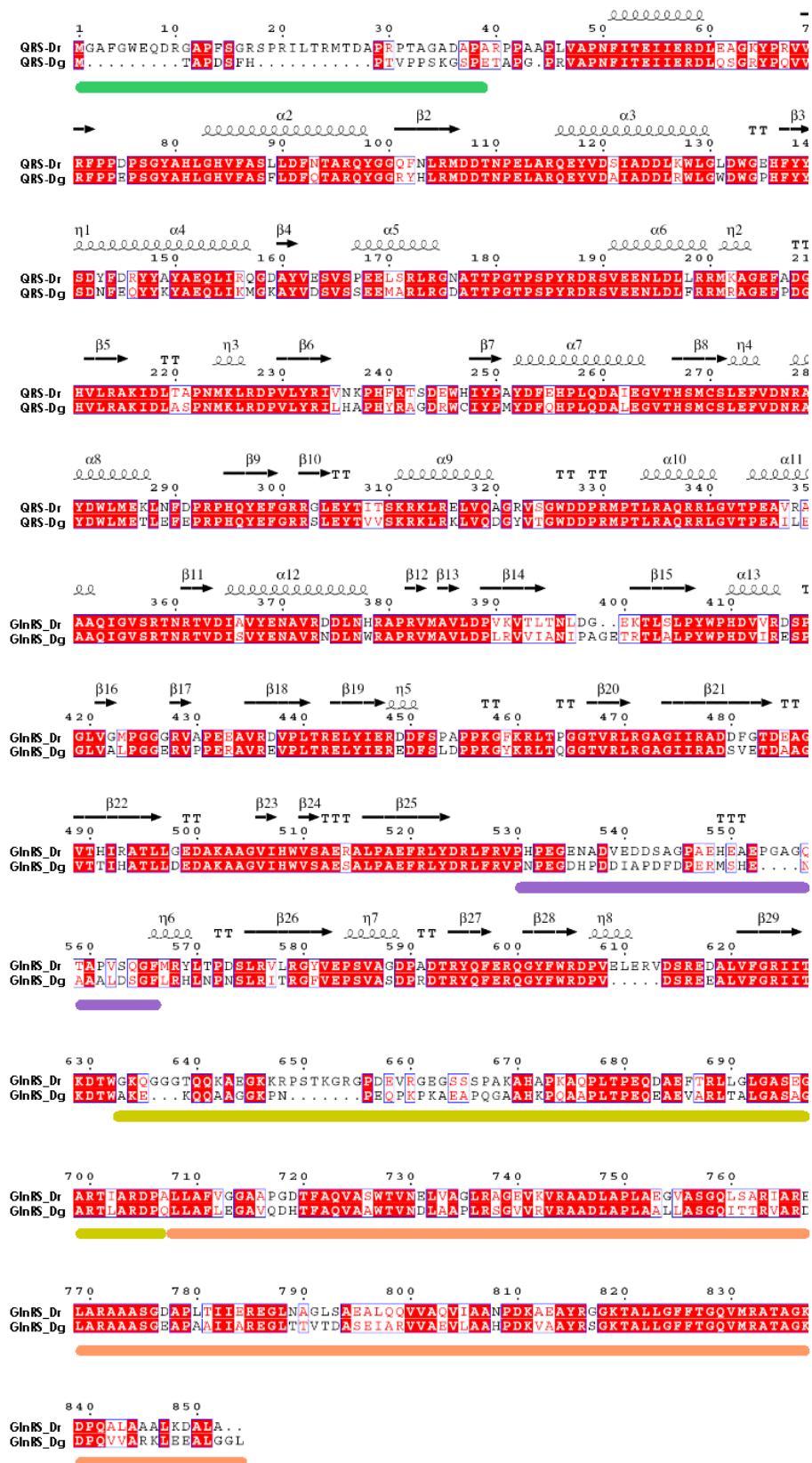


Figure II. 18 : Alignement des séquences entre les GlnRS entre *D. radiodurans* et *D. geothermalis*. Les limites des insertions et extensions sont représentées par des lignes de couleur : en vert, l'insertion N-terminale (1-43), en pourpre, la boucle *ins2*. Le module C-terminal additionnel est indiqué en jaune (linker) et en orange (domaine YqeY).

Une première hypothèse est qu'une telle conformation est celle employée, en l'absence d'ARNt, soit pour empêcher la fixation de l'acide aminé au niveau du site actif, soit pour l'y maintenir en attendant l'ARNt. L'ARNt serait la clé nécessaire pour le déverrouiller et autoriser l'entrée d'un ou deux petits substrats. On peut proposer le mécanisme suivant : la glutamine liée au site actif induit sa fermeture et interdit l'accès de l'ATP. Lorsque l'ARNt homologue se fixe au site de liaison à l'anticodon, il induit, par une étroite communication entre les modules facilitée par la boucle, l'ouverture du site actif, le positionnement du CCA et la réorientation de la glutamine. La fixation de l'ATP étant autorisée, la réaction d'activation peut avoir lieu.

Une seconde hypothèse est que la boucle représente un appendice supplémentaire responsable de la discrimination de l'acide aminé. En effet, la GlnRS-*Dr*, contrairement à *E. coli*, est incapable de lier le glutamate (Deniziak *et al.*, 2007). Au contraire, chez *E. coli*, la discrimination du glutamate s'opère au moment de l'activation par le positionnement de l'acide aminé dans le site actif (Bullock *et al.*, 2003), (Introduction § I.C.1). Chez *D. radiodurans*, cette aptitude à discriminer ces acides aminés stériquement proches pourrait être liée à la présence de cet élément structural additionnel. Si la conformation est induite par la présence de la glutamine dans le site actif et qu'elle est reconnue spécifiquement, il pourrait, une fois l'acide aminé spécifique reconnu, adopter une conformation fermée.

Une dernière hypothèse ferait intervenir le domaine YqeY pour interagir à la boucle **ins2** libérant ainsi le site pour autoriser la fixation de l'ARNt. YqeY aurait alors un rôle de chaperonne pour former le complexe actif en présence d'ARNt. Ainsi, même si la plupart des résidus impliqués dans la reconnaissance des substrats de la GlnRS d' *E. coli* sont conservés, **ins2** comme YqeY- λ pourraient avoir le rôle d'additifs structuraux nécessaires (ou indispensables) au bon fonctionnement de l'enzyme.

4. Comparaison avec la GlnRS-ΔC

Les données récemment enregistrées pour la GlnRS-ΔC ont permis d'obtenir des résultats préliminaires de la structure de cette enzyme tronquée. La GlnRS-ΔC cristallise selon un empilement similaire à la GlnRS entière, mais avec des paramètres de maille bien distincts, dans des conditions de cristallisation 0,2 M di-ammonium citrate et 20% PEG 3350. La GlnRS-ΔC ne semble pas opérer de modifications de conformations significatives et l'architecture globale du cœur chez l'enzyme entière ou tronquée est très proche. Ces résultats confirment les données biochimiques et structurales : en effet, la GlnRS-ΔC conserve ses propriétés de catalyse, est capable de lier la glutamine et l'ATP et tous les résidus essentiels à la réaction sont conservés chez *E. coli* et dans une conformation active. Cependant un élément structural diffère de la GlnRS-*Ec* : la GlnRS-ΔC observe une conformation fermée (fGlnRS-ΔC) illustrée par la présence de la boucle d'insertion **ins2** positionnée comme la f-GlnRS. Aucun contact cristallin ne semble intervenir dans la stabilisation de la boucle (Figure II. 19).

Ce résultat nous confirme tout d'abord l'existence de 2 conformations de l'enzyme : la f-GlnRS et la o-GlnRS, (f- pour fermée, o- pour ouverte). On ne note, en préliminaire aucune modification significative du cœur visible et cela nous confirme que c'est l'extrémité C-terminale additionnelle qui est responsable des particularités fonctionnelles de l'enzyme. L'absence de ce domaine ne provoque pas d'instabilité supplémentaire de l'une ou l'autre région de la protéine tronquée.

Enfin, la boucle en position verrouillée ne nécessite pas la présence du module YqeY. Ce résultat nous montre également que la GlnRS-ΔC peut observer cette conformation à l'état natif, sans ajout de glutamine. Le domaine YqeY confère-t-il un degré de flexibilité supérieur à l'enzyme native qui lui permettrait de présenter 2 conformations en équilibre dynamique ? Si c'est le cas, la GlnRS-ΔC observe une certaine rigidité conformationnelle qui expliquerait pourquoi celle-ci présente si peu d'affinité pour l'ARNt^{Gln}. Nous l'avons vu précédemment, le site actif en position fGlnRS est extrêmement encombré et laisse peu de place pour le positionnement de l'extrémité CCA : l'enzyme reconnaît moins facilement l'ARNt et la présence d'YqeY conditionnerait sa charge. Le module YqeY augmente l'efficacité catalytique de l'activation (X1,7) comme de la liaison de l'ARNt (X170). Celui-ci pourrait jouer un rôle dans la flexibilité conformationnelle de la boucle et donc dans son efficacité à reconnaître ses substrats. Connaissant ces éléments on peut penser que le domaine YqeY est le domaine supplémentaire nécessaire à la reconnaissance préliminaire de l'ARNt qui induit par coopérativité l'ouverture de la boucle pour l'activation de l'acide aminé.

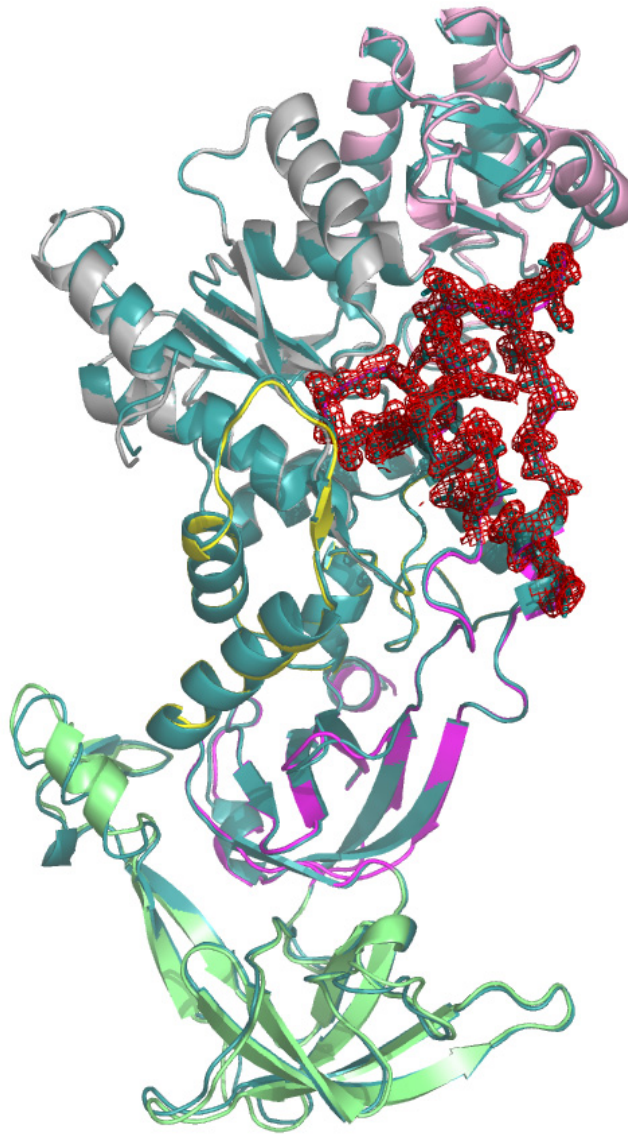


Figure II. 19 : Superposition de la fGlnRS et fGlnRS- Δ C avec l'*ins2* en conformation fermée. Les 2 structures du cœur sont identiques (RMSD = 0,64 Å). La GlnRS- Δ C est représentée en bleu clair, la fGlnRS avec le même code couleur qu'en figure A. La densité électronique (Fo-Fc) correspondant à la boucle ins2 est donnée en rouge (contour à 2,5 σ). Les chaînes latérales des résidus de la boucle sont représentées.

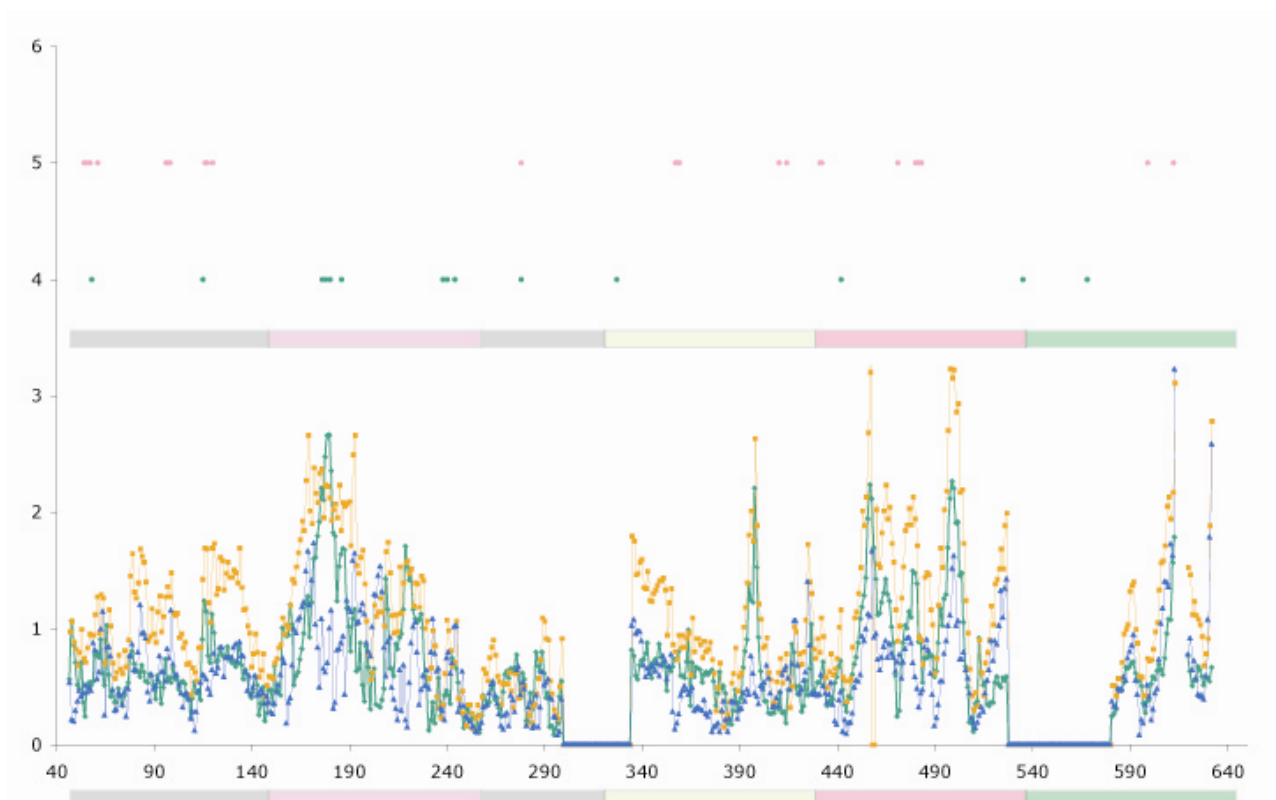


Figure II. 20 : Représentation graphique de la distance entre les résidus des structures 2 à 2.

La superposition des structures a été effectuée domaine à domaine. Les variations des distances a été calculée entre les C α des résidus correspondants. Les distance interrésidus sont représentées comme suit : entre fGlnRS et GlnRS-Dr en orange; entre GlnRS-QSI et GlnRS-Dr en vert; entre fGlnRS et GlnRS-QSI en bleu. La bande colorée représente la localisation des différents domaines dans la séquence primaire. Les codes couleurs des modules suivent la nomenclature de la GlnRS d'*E. coli* : domaine de Rossmann en gris, domaine ABD en rose pâle, domaine intermédiaire en jaune, domaines de liaison à l'anticodon de l'ARNt en rose et vert (domaines distal et proximal, respectivement). Les contacts intermoléculaires sont précisés par des points, de couleur rose pour les contacts communs aux 3 structures et de couleur verte pour les contacts supplémentaires retrouvés dans l'empilement de la GlnRS native. Les insertions visibles dans la structure n'ont pas été prises en compte pour l'alignement.

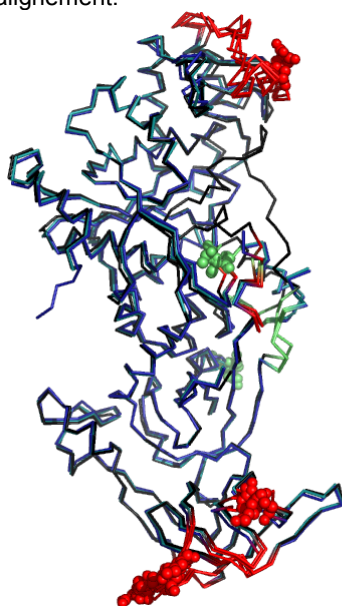


Figure II. 21 : Mise en évidence des régions comportant les variations de distances C α -C α pour l'enzyme Native la fGlnRS et la GlnRS/ QSI. Les distances illustrées sont celles comprises entre 2 et 3 Å, sans tenir compte de la boucle d'insertion ins2. Les régions colorées en rouge correspondent à celles présentant des distances GlnRS native-GlnRS/fGlnRS ou native-GlnRS/QSI > 2Å, celle en vert des distances entre fGlnRS et GlnRS/QSI > 2 Å.

III. Plasticité de l'empilement cristallin

Une autre approche de mon travail concerne l'étude de la géométrie de l'empilement cristallin de la GlnRS-Dr. Comme cela a été observé dans le cas de l'enzyme libre (*Article 1 - Touzé et al., 2007, pages 59-62*). Les cristaux de la GlnRS-Dr combinent une proportion élevée de désordre et une bonne qualité des données de diffraction. Une seconde propriété notable de cet empilement est liée à sa plasticité. La géométrie de l'empilement cristallin de l'enzyme reste le même pour la GlnRS-Dr native ou en présence des petits substrats. La comparaison des paramètres de maille des enzymes cristallisées en présence ou non de petits substrats nous a permis d'observer la variabilité des paramètres de maille : ainsi, le volume de la maille augmente jusqu'à 14% selon que l'enzyme est cristallisée en présence de ces ligands. En effet, lorsque les petits substrats sont dans le milieu de cristallisation, les changements conformationnels induits sont tels qu'elle entraîne une modification notable des paramètres de maille pour 2 de ses 3 dimensions (a et c) et un réarrangement des contacts cristallins. Ces modifications conformationnelles ont un effet sur la surface globale accessible de l'enzyme. L'enzyme cristallisée avec ces ligands présente une conformation plus compactée en particulier dans la région du site actif de l'enzyme d'où une réduction de la proportion de résidus disponibles en surface, et donc des zones de contacts intermoléculaires. C'est ainsi que se traduit l'augmentation des paramètres de maille qui rend l'empilement plus aéré sans toutefois modifier sa géométrie et sa symétrie. La dilatation de la maille est également illustrée par le contenu en solvant des cristaux qui passe de 44 à 51 % et de la taille des canaux.

A. Les contacts cristallins

La géométrie de l'empilement cristallins dépend des contacts directs protéine - protéine. Les forces impliquées dans l'empilement des macromolécules sont faibles (ponts salins, liaisons hydrogènes, interactions de Van Der Waals, la cohésion dipôle-dipôle) et ce sont les interactions non spécifiques qui gouvernent les interactions intermoléculaires dans un cristal. La fragilité des cristaux de macromolécules est aussi liée au fait que seules de petites zones de la surface macromoléculaire sont impliquées dans les contacts intermoléculaires (Carugo et Argos, 1997; Janin, 1997). Si les interactions intervenant dans l'empilement cristallin sont de même nature que les interactions protéine - protéine retrouvées *in vivo* (complexes, oligomères, ...), elles diffèrent par le fait qu'elles ne sont pas sujettes à la pression sélective, et non spécifiques (Bahadur *et al.*, 2004). Ainsi, la cristallisation de protéines dépend d'interactions non prédictibles qui se distinguent du processus de reconnaissance entre protéines dans les conditions physiologiques. La nature des contacts cristallins est très sensible à l'environnement chimique de la molécule. C'est pourquoi les sels, agents cristallisants ou additifs sont susceptibles de modifier les interactions protéine – protéine.

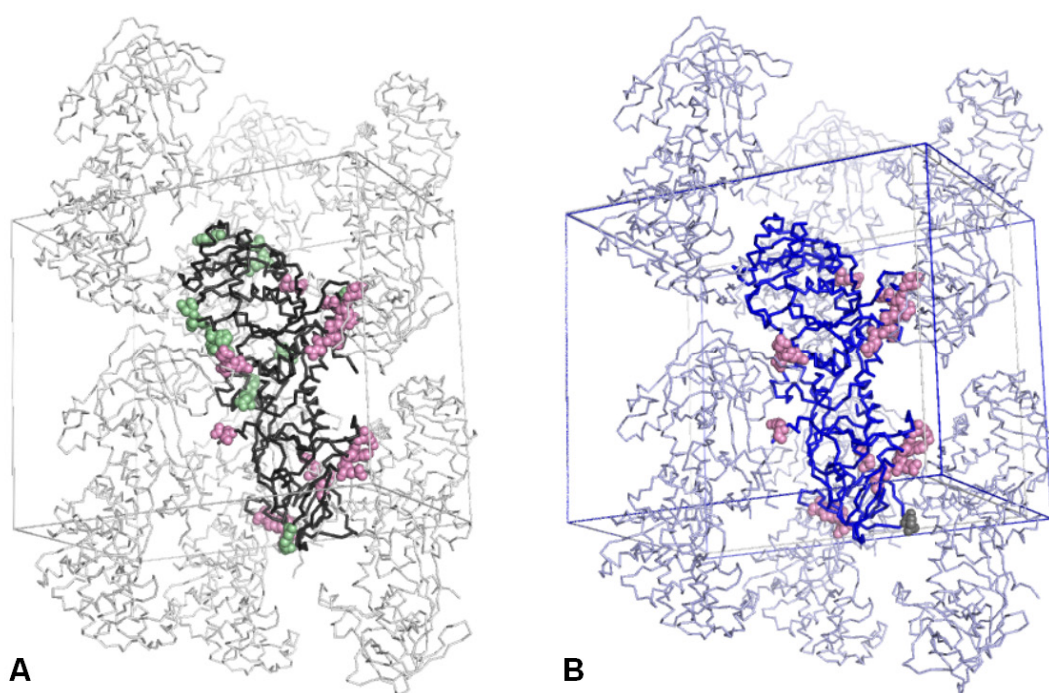


Figure II. 22 : Illustration des contacts cristallins impliqués dans l'empilement pour la GlnRS-Dr native (gris, **A**) ou co-cristallisée en présence de petits substrats. (bleu, **B**) Les contacts sont représentés par des sphères roses pour les contacts conservés dans les trois structures, et vertes et grises pour ceux uniquement trouvés dans la structure native ou en présence de petits substrats, respectivement.

La recherche de nouvelles conditions de cristallisation peut donc engendrer un empilement cristallin différent et stabiliser des régions désordonnées dans d'autres empilements. On peut citer l'exemple d'une enzyme pancréatique (Ribonucléase A pancréatique bovine ou RNase A) pour laquelle la quasi-totalité de la surface a pu être impliquée dans ces interactions ou des différents empilements trouvés pour la thaumatine (Charron *et al.*, 2004; Crosio *et al.*, 1992) .

Ce n'est pas le cas de notre protéine qui, quelles que soient les conditions de cristallisation ou les substrats en présence, cristallise selon un même groupe d'espace ($P2_12_12_1$) : les 3 structures résolues en présence de petits substrats et la GlnRS native. Néanmoins, d'importantes variations des paramètres de maille dans 2 des 3 dimensions sont engendrées par la présence de ces substrats, celles en présence de QOL étant intermédiaires entre la protéine native et la GlnRS /QSI. Ceci a des conséquences directes sur l'étendue des contacts cristallins et la largeur des canaux de solvant. L'analyse des contacts cristallins est illustrée [Figure II. 22](#). La surface de contact accessible pour tous les résidus de la protéine native est de $25600 \text{ \AA}^2$ contre $24500 \text{ \AA}^2$ lorsque la protéine est cristallisée en présence de substrat. La surface de contact totale impliquée dans l'empilement cristallin s'élève à $7700 \text{ \AA}^2$ et $7400 \text{ \AA}^2$ pour la protéine native et complexée respectivement. Ces modifications résident, comme nous l'avons vu plus haut, dans la compaction générale de l'enzyme en particulier pour le domaine ABD et le domaine de liaison à l'anticodon. En ce qui concerne l'architecture du cristal, la majorité des contacts utilisés pour les protéines co-cristallisées avec des ligands sont conservés dans l'empilement de la protéine native. S'ajoutent néanmoins dans le cas de la forme native, des contacts supplémentaires qui contribuent à sa stabilité. La majorité d'entre eux se situent aux environs du domaine catalytique de l'enzyme. En particulier, tous les contacts à proximité de la poche du site actif et impliqués dans l'empilement de la protéine native ne sont pas maintenus en présence des petits substrats.

B. Les canaux de solvant

En termes d'encombrement moléculaire, un cristal de protéine peut être une bonne approximation de la cellule. En effet, l'empilement des molécules délimite des cavités ou canaux de solvant qui peuvent aller jusqu'à 100 Å de large. Une autre cause de la fragilité d'un cristal est alors liée à la forte proportion en solvant contenu dans le cristal pouvant atteindre 80%. La modification de l'empilement cristallin peut changer drastiquement les taux de solvant comme l'exemple observé pour la protéine TcAChE (acétylcholinestérase de *Torpedo californica*) cristallisée sous forme de cristaux trigonaux avec 68% de solvant, ou orthorhombiques (56% de solvant). Dans le cas d'un même empilement cristallin, la modification des paramètres de maille d'un cristal et de son taux de solvant peuvent être induits par la congélation qui entraîne la contraction des canaux de solvant ou leur élargissement par la formation de glace, ou aussi par la déshydratation qui confine les protéines dans des canaux de solvant plus petits (Weik, 2003). Les variations observées sont alors de l'ordre de 1 à 2%. Dans notre cas, le volume occupé par les canaux de solvant est de 14% supérieur à celui occupé en présence de protéine native. La variation des paramètres de maille n'est donc pas engendrée par l'influence de l'environnement mais par des variations de conformation suffisamment importantes pour engendrer un défaut de surface disponible qui entraîne lui-même un empilement plus dilaté. L'addition des petits substrats laisse une place suffisante pour autoriser les mouvements flexibles de l'extrémité C-terminale additionnelle. La nouvelle configuration de l'empilement offre une plus grande marge de flexibilité pour la protéine et plus particulièrement pour l'extrémité additionnelle C-terminale. Quelle que soit la taille des canaux de solvants, elle suffit pour accommoder l'extrémité C-terminale non ordonnée comme présenté dans les modèles Figure II.23.

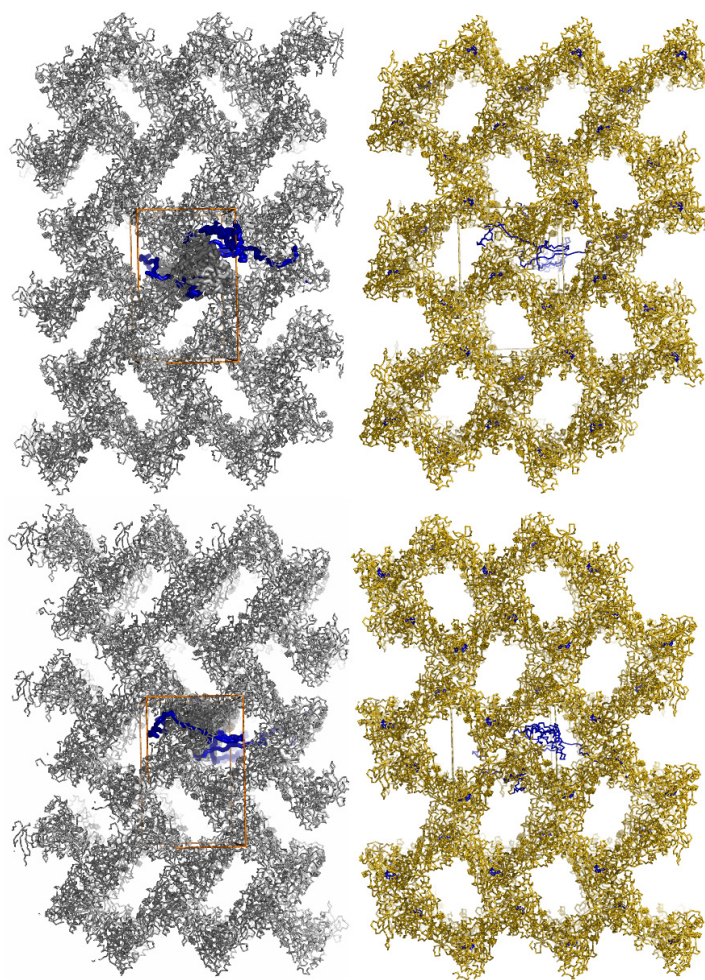


Figure II. 23 : Variation de la taille des canaux de solvant dans les cristaux de GlnRS native (gris) ou complexée (or). Le modèle de la GlnRS pourvue de ses extrémités N et C-terminales (présenté en figure A) y est accommodé en bleu foncé. Les côtés a et c de la maille cristalline sont présentés par la boîte. La taille des canaux de solvants passe d'une surface de $800 \text{ \AA}^2$ à $1200 \text{ \AA}^2$.

La dilatation des paramètres de maille engendre la disparition de nombreux contacts cristallins qui n'est pas compensée par la formation de nouvelles interactions intermoléculaires. Les cristaux comportant des canaux de solvant de volume élevé sont souvent plus fragiles en particulier lors de leur congélation. Les propriétés de diffraction des cristaux obtenus en présence de GlnRS et de ses petits substrats semblent cependant ne pas être altérées par l'augmentation des paramètres de maille. Le remplissage de l'espace libre par le module additionnel C-terminal peut-être une explication. La présence de gel d'agarose en est une autre. Une hypothèse peut-être émise au sujet des contacts "manquants" dans l'empilement de la protéine cristallisée en présence des petits ligands. Ceux-ci se situent essentiellement dans la région du site actif. Or la co-cristallisation de la GlnRS avec la glutamine a permis de stabiliser une boucle à proximité du site actif. De plus, les résidus impliqués dans les contacts intermoléculaires des 2 structures co-cristallisées sont les mêmes. Celles-ci adoptent une conformation contractée en direction du site actif qui empêcherait les résidus impliqués dans les contacts cristallins formés dans l'empilement de la protéine native de s'établir. Les modifications intermédiaires des paramètres du glutaminol (QOL) peuvent également être mises en parallèle : les paramètres de maille dans lesquels elle cristallisent sont à l'image de sa conformation, à l'intermédiaire entre la GlnRS native et la GlnRS / QSI.

IV. Etudes en solution de la GlnRS

A. Résonance magnétique nucléaire

Avec la collaboration de l'équipe de Bruno Kieffer à l'ESBS, Illkirch, et avec le soutien de Marc Vitorino, des études en solution ont tout d'abord été engagées afin d'évaluer qualitativement l'état structuré des 2 constructions YqeY- λ et YqeY- δ . Un spectre RMN dit « à 1 dimension », mettant à profit la résonance des protons ^1H peut-être effectué sur des protéines de taille modérées (10 à 30 kDa). Les deux constructions de 16 et 20 kDa ont été analysées. Une protéine repliée peut-être distinguée d'une protéine non repliée par la largeur spectrale de ses déplacements chimiques. Une protéine désordonnée aura un spectre caractéristique de faible largeur spectrale (6 à 10 ppm) avec un faible étalement des déplacements chimiques pour chaque type de proton (proton amide, proton H α , proton méthyles...). Au contraire, une protéine structurée présente des déplacements chimiques "signature" représentés en général par la présence de protons « méthyle » aux alentours de 0 ppm caractéristiques d'un cœur hydrophobe. Le spectre de la protéine YqeY- λ est clairement celui d'une protéine structurée (Figure II. 24). On y distingue 2 pics typiques du tryptophane structuré (W726) situé au cœur du domaine YqeY, les signatures des groupements méthyles entre 0 et -1 ppm ainsi qu'une bonne définition des protons amide (N-H) dans la zone 6-9 ppm. La même expérience a été effectuée sur la construction YqeY- δ , (dépourvue du linker de l'extrémité additionnelle). Contrairement aux prédictions de structure secondaire, le domaine YqeY- δ est moins bien structuré en solution que le domaine YqeY- λ . Un spectre protons ^1H a en effet également été effectué dans les mêmes conditions sur le domaine YqeY- δ . On retrouve les signatures d'un des 2 protons tryptophane et les protons méthyles. Malgré cela, la fenêtre spectrale des méthyles est plus étroite et les protons amide (CO - N-H) sont moins bien définis. Il s'agit là de signatures de chaînes latérales structurées et déstructurées. Ce résultat préliminaire sous-entend que le domaine YqeY- δ présente un degré de structuration inférieur au domaine YqeY- λ (spectre non présenté).

La stratégie de coupure pour les différentes constructions a été élaborée en fonction des alignements avec les domaines YqeY liés ou libres et de la structure tridimensionnelle de YqeY de *Bacillus subtilis*. Les résultats préliminaires de RMN à 1 dimension penchent vers une forte probabilité que la région prédite comme "non structurée" (linker) de l'extrémité additionnelle de la GlnRs-Dr soit, en réalité, ordonnée dans sa plus grande partie.

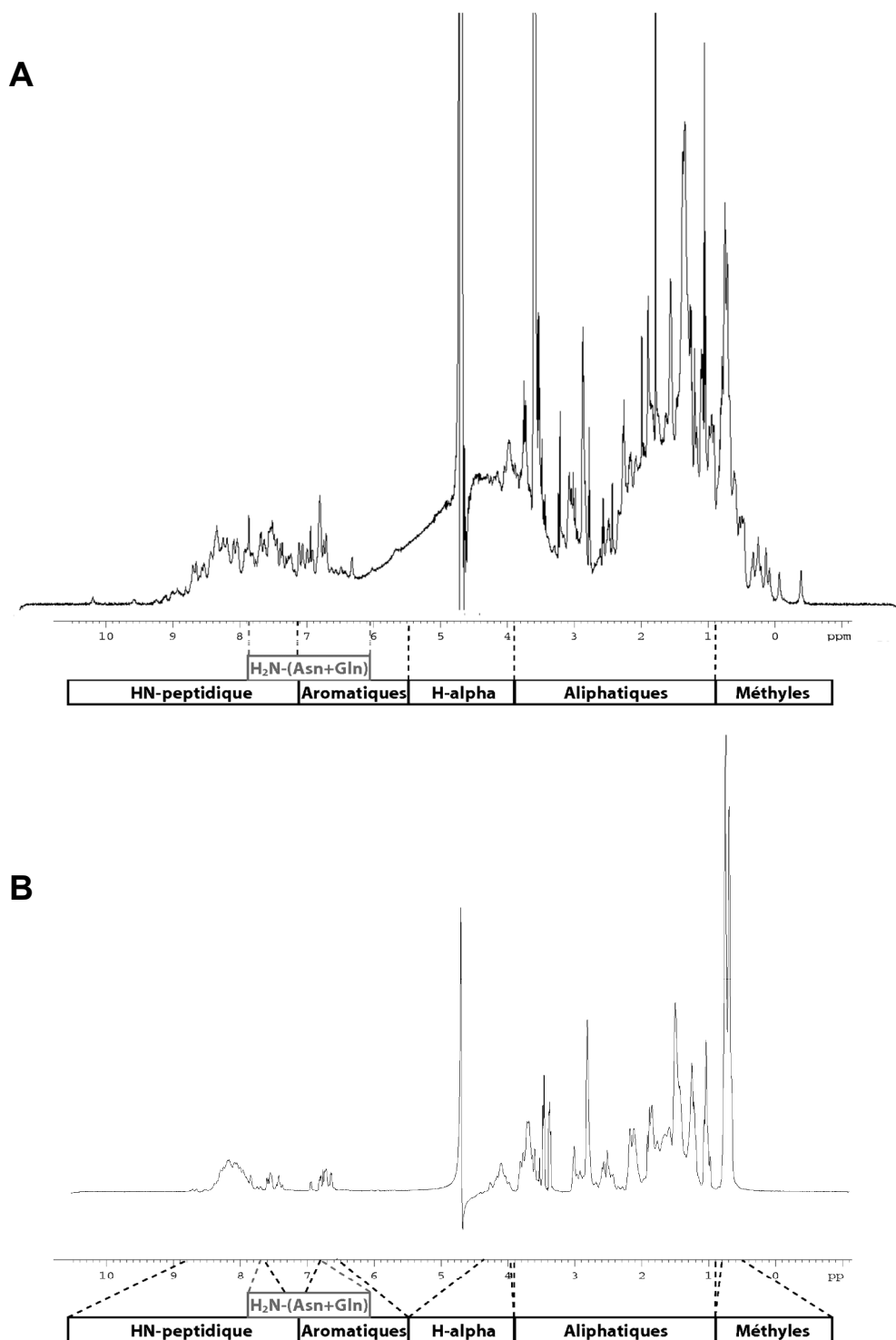


Figure II. 24 : Spectre 1D ^1H à 600MHz, 298K, d'une protéine repliée, YqeY- λ (A) et non repliée, SBP2 (B). Le spectre A, (ce travail) montre un large étalement et une nette distinction des pics individuels dans le région des protons HN-peptidique, des pics caractéristiques d'un cœur hydrophobe dans la région (-1 – 0 ppm), 2 pics représentatifs du tryptophane (9 – 10 ppm). Les valeurs des déplacements chimiques sont indiquées sur l'axe des abscisses et l'emplacement selon la nature des protons précisé en gras.

Les difficultés de cristallisation du domaine YqeY sous forme libre (YqeY- λ , δ) ou liée au cœur de l'enzyme (GlnRS-Dr) et la qualité de résolution du spectre 1D du domaine YqeY- λ nous a conduit à effectuer des expériences de RMN en 2 dimensions. La HSQC (Heteronuclear Single Quantum Coherence) est une méthode rapide et relativement peu coûteuse pour évaluer si une protéine est une bonne candidate pour la détermination ultérieure de sa structure tridimensionnelle. Une ^1H - ^{15}N HSQC nécessite le marquage de la protéine par l'isotope rare ^{15}N . Chaque résidu de la protéine (excepté la proline) possède un "proton amide" attaché au $\text{N}\alpha$ de la liaison peptidique. Dans une expérience HSQC, on cherche à identifier le signal de chaque proton $^{15}\text{NH-CO}$ du squelette polypeptidique sachant que chaque résidu admet un déplacement chimique propre à son environnement. Plus une protéine est structurée, plus son environnement électronique est variable. Ainsi, si la protéine est repliée, les pics sont dispersés et la plupart des pics individuels peuvent être distingués. Le nombre de pics d'un spectre doit correspondre au nombre de résidus de la protéine, auxquels s'ajoutent des "protons amine" des résidus dont les chaînes latérales contiennent des protons liés à un atome d'azote (par exemple la glutamine ou l'asparagine).

Les résultats de l'expérience sont présentés [figure II. 23](#) et nous ont conduit aux conclusions suivantes : (i) l'étalement des tâches et l'excellente résolution du spectre laissent à penser que la présence d'une protéine repliée dans sa quasi-intégralité, (ii) la localisation des déplacements chimiques donnent une bonne estimation pour une protéine à majorité de structure secondaire en hélices α , la résolution est de très haute qualité puisque le comptage des pics visibles sur le spectre HSQC ^1H - ^{15}N de YqeY- λ a donné un total de 204 pics sur les 215 résidus présents. L'ensemble de ces données fait de YqeY- λ une bonne candidate pour la résolution de sa structure 3D par RMN.

Des expériences similaires ont été effectuées en présence de GlnRS- ΔC soit en quantité équimolaire soit en excès (5X). Deux expériences ^1H - ^{15}N HSQC successives, en absence et en présence de ligand (ici le cœur de la synthétase) permettent de visualiser les interactions entre les 2 populations du mélange. Les résultats préliminaires d'interaction des 2 protéines ont permis de confirmer l'absence d'interaction de YqeY- λ avec la GlnRS- ΔC , tronqué en C-terminal. En effet, en présence et en absence de GlnRS- ΔC , les pics des déplacements chimiques et les intensités sont identiques. Aucune interaction en trans n'a pu être identifiée dans chaque condition. Nous admettons ici que les deux modules de la GlnRS-Dr doivent, s'ils interagissent, être liés pour entrer en interaction. Si la présence de l'extrémité C-terminale additionnelle est indispensable pour le maintien de l'activité catalytique de l'enzyme, elle l'est probablement pour permettre l'interaction des modules qui, en solution ne présentent pas d'affinité naturelle l'un vis-à-vis de l'autre. Ce résultat confirme les expériences biochimiques précédentes qui admettaient précédemment que l'appendice YqeY- λ n'interagit pas à l'état libre avec son ARNt homologue et rejoint les expériences de retard sur gel (Deniziak *et al.*, 2007).

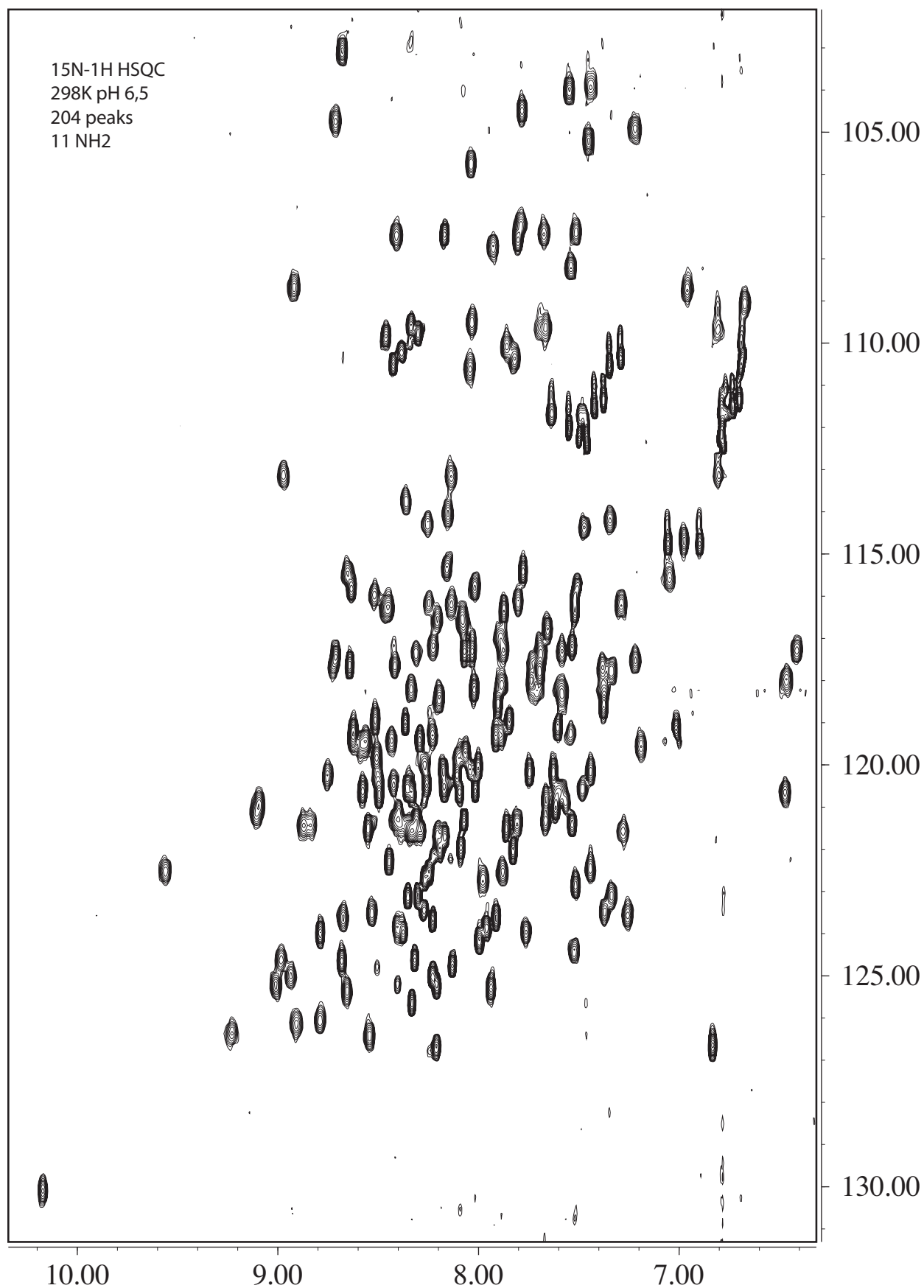


Figure II. 23 : Spectre ^1H - ^{15}N – HSQC du domaine YqeY- Δ marqué en ^{15}N à 25°C pH 6.5. Les taches individuelles représentent les déplacements chimiques de chaque “proton amide” avec en abscisse ceux des ^1H et en ordonnée ceux des ^{15}N . On décompte 204 pics dont 11 potentiels pour les NH2 des chaînes latérales des asparagines et glutamines.

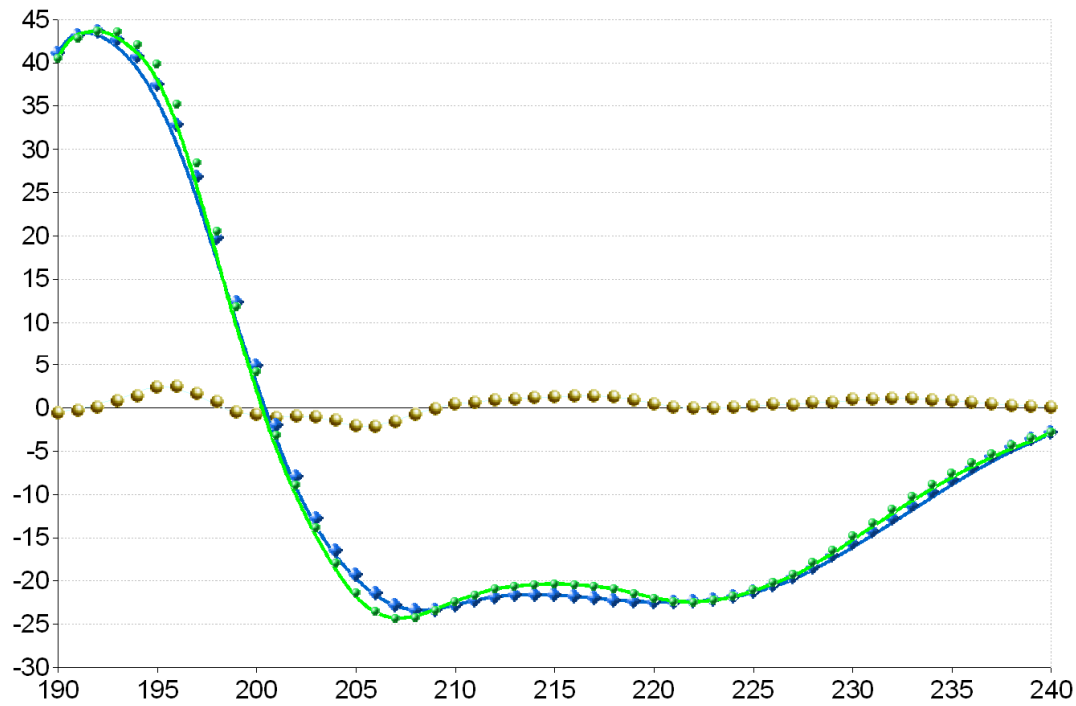
L'ensemble de ces résultats d'analyse structurale en solution par RMN nous indiquent qu'il est peu probable que le "linker" ou extension prédite comme flexible, riche en lysines et probablement impliqué dans l'interaction avec l'ARNt soit désordonné. Il faut néanmoins prendre en compte le fait que cette construction n'est pas sous sa forme native, c'est à dire liée au cœur de la GlnRS. Les données structurales du module YqeY- λ vont dans le sens d'un module totalement structuré, "linker" compris.

Pour le confirmer, nous avons souhaité compléter les données obtenues en soumettant le domaine YqeY- λ à des expériences de dichroïsme circulaire.

B. Analyse par dichroïsme circulaire de l'extrémité YqeY

Le dichroïsme circulaire (ou DC) est une technique spectroscopique d'analyse largement utilisée pour étudier la conformation des acides nucléiques et des protéines. Dans le cas des protéines, il permet d'obtenir des informations sur la nature des différents types de structures secondaires. L'allure d'un spectre DC dépend de l'organisation locale de la liaison peptidique, ce qui permet de distinguer hélices α , feuillets β , ou coiled-coil (hélice hydrophobes).

L'objectif d'une étude en dichroïsme circulaire est d'obtenir une estimation quantitative de la proportion structurée de YqeY- λ . La bonne répartition des déplacements chimiques obtenus en RMN 1 et 2D nous indiquaient que le module C-terminal additionnel semblait très majoritairement replié. L'obtention d'un spectre de bonne qualité est très importante pour une bonne estimation de la proportion des structures secondaires. Les mesures sont en effet extrêmement sensibles au "bruit de fond" qui dépend de la nature des tampons, de leur degré d'oxygénation, du pH mais également de concentration du produit testé. Pour cela, les logiciels de traitement des données procurent un facteur de qualité, NRMSD (Normalized Root Mean Square Deviation), basé sur la correspondance entre les données théoriques et expérimentales. Les résultats sont présentés en [Figure II. 24](#).



N résultat	Hélice 1	Hélice 2	Feuillet 1	Feuillet 2	Tours	Désordonné	Total
1	0,293	0,21	0,016	0,041	0,181	0,259	1
2	0,291	0,221	0,02	0,044	0,162	0,264	1,002

Moyenne 0,292 0,2155 0,018 0,0425 0,1715 0,2615 1,001

NRMSD : 0,043

Figure II. 24 : Spectres expérimental et théorique du domaine YqeY-λ enregistré à 37°C, pH 6,5. Les données expérimentales et théoriques et leur différence sont données en vert, bleu et jaune, respectivement. Les estimations des proportions de structures secondaires sont données dans le tableau ci-dessus. La qualité des statistiques est donnée par la valeur du NRMSD qui doit-être d'une valeur inférieure à 0.1 (Whitmore, *et al.*, 2004).

L'analyse des données en dichroïsme circulaire donne une estimation de la répartition des structures secondaires très majoritairement composés d'hélices α . Or le domaine YqeY de *B.subtilis* dont la structure tridimensionnelle a été résolue, est composé uniquement d'hélices α (78% d'hélices, soit 9; 116 résidus). Les estimations donnent 50% de la structure totale en hélice α . Cette valeur correspondrait au pourcentage d'hélice α dans le noyau YqeY soit environ ... 78%. Dans ce cas, le domaine C-terminal ne présenterait de repliement secondaire que dans sa région homologue au domaine YqeY.

Les mesures ont permis de confirmer la nature des structures secondaires qui constituent l'appendice C-terminal mais ne permettent pas de conclure clairement sur la part structurée de celui-ci. Il est toutefois probable que le domaine présente effectivement une structure secondaire couvrant uniquement la part homologue de YqeY et dont les 25% restants (correspondant à la part "désordonnée" des mesures de DC), seraient non repliés. Il n'en reste pas moins que les spectres obtenus par HSQC sont très encourageants pour conduire à la résolution de la structure du domaine.

- CONCLUSIONS ET PERSPECTIVES -

Conclusions et Perspectives

La détermination de la structure de la glutaminyl-ARNt synthétase de *D. radiodurans* (GlnRS-*Dr*) sous forme native et les questions posées par cette synthétase atypique ont constitué le point de départ de mon projet de thèse. L'un de mes objectifs était d'étudier la structure des séquences additionnelles, en majeure partie désordonnées dans les cristaux de cette GlnRS. Au terme de ce travail, j'ai obtenu la structure cristallographique de l'enzyme en présence de deux analogues de l'intermédiaire réactionnel de la réaction d'activation (QSI et QOL). Je suis également parvenue à stabiliser l'insertion située à la charnière des domaines catalytique et de liaison à l'anticodon de l'ARNt^{Gln} et qui vient fermer le site actif. En outre, ce travail m'a permis d'obtenir la structure de l'enzyme tronquée de son module C-terminal additionnel (GlnRS-ΔC) dont la structure globale n'est pas affectée par la délétion et, présente, la même conformation fermée du site actif.

Le module C-terminal additionnel n'a pas pu être stabilisé dans les différentes formes cristallines de GlnRS que j'ai obtenues, ni être résolu dans sa forme isolée. Néanmoins, des études en solution (RMN et dichroïsme circulaire) nous ont apporté des éléments sur la nature de ses structures secondaires et la proportion de régions potentiellement ordonnées de ce module.

L'étude cristallographique de la GlnRS-*Dr* nous a également permis de mettre en évidence son étonnante aptitude à cristalliser malgré l'importante proportion de régions mobiles dans le cristal mais aussi la surprenante plasticité de l'empilement cristallin.

Petits substrats fixés en absence d'ARNt^{Gln}

La comparaison des positions des atomes des analogues d'adénylate dans le site actif de la GlnRS-*Dr* et de la GlnRS-*Ec* en complexe avec l'ARNt^{Gln} nous a montré que la région AMP de l'adénylate était dans une orientation non productive pour la réaction d'activation. Ce décalage est aussi bien observé pour le QSI que le QOL et est en particulier expliqué par la position, ici en marge de la région AMP, de la boucle du KMSKS. Cette dernière est habituellement impliquée dans la stabilisation de l'AMP et conditionnée par l'intervention de l'ARNt homologue chez les GlnRS.

Fixation de l'acide aminé et boucle ins2

La boucle *ins2* est stabilisée lorsque la GlnRS-*Dr* est cristallisée en présence de glutamine ou lorsque le module C-terminal additionnel est délété, mais est désordonnée dans les cristaux natifs et lorsque les analogues d'adénylate sont dans le site actif. Cette dernière propriété est

avant tout expliquée par l'impossibilité de positionner la boucle lorsque les analogues de substrat occupent une partie de la place qu'elle prendrait si elle fermait le site.

Rappelons que nous n'avons pas d'évidence que la glutamine soit présente ou non dans le site actif de l'enzyme. Quoi qu'il en soit, ni l'acide aminé, ni l'ATP ne sont en capacité d'adopter leur orientation productive pour la réaction d'activation lorsque le site actif est fermé. La conformation de la boucle est-elle directement induite par la présence de glutamine dans le milieu ? Est-elle impliquée dans la discrimination de l'acide aminé ? Prépare t-elle le site actif à opérer la réaction d'activation en adoptant sa conformation active ? Ces éléments qui ajouteraient une étape supplémentaire à la réaction d'activation, confèreraient à cette boucle une fonction protectrice pour préparer l'enzyme à accueillir de l'ARNt^{Gln}.

Synergie entre boucle ins2 et extrémité C-terminale

La forme QSR-ΔC, qui constitue une enzyme très proche de celle qui est trouvée chez *E. coli*, n'est pas déstabilisée par l'absence du module C-terminal alors que les conséquences fonctionnelles causées par cette délétion sont réelles. C'est essentiellement la boucle *ins2* qui distingue les deux enzymes et cet élément supplémentaire interroge, d'une part sur le rôle de la glutamine dans l'acquisition de cette conformation, d'autre part sur celui du module C-terminal dans la dynamique de cette région. Une hypothèse serait que le module intervient dans l'ouverture du site actif qui pourrait expliquer pourquoi l'activité catalytique de la forme GlnRS-ΔC est réduite.

En plus de cela, le module C-terminal pourrait également faciliter l'interaction de l'ARNt homologue avec le domaine de liaison à l'anticodon de la GlnRS et jouerait un rôle de chaperon pour l'enzyme. La communication, renforcée par la position stratégique de la boucle à l'interface des domaines catalytique et de liaison à l'anticodon, faciliterait l'ouverture du site actif pour autoriser l'entrée du ou des petits substrats.

Propriétés structurales de l'extrémité C-terminale

Les données expérimentales en RMN apportent des indications qualitatives sur la proportion structurée du module additionnel (forme YqeY-I). L'étalement et la position des déplacements chimiques du module C-terminal laissent à penser qu'il serait replié dans sa majorité. Le domaine 'YqeY' (forme YqeY-d) seul présente une structure bien moins ordonnée ce qui nous indique que la partie 'linker' participe à la structuration du domaine. Les résultats des tests en dichroïsme circulaire ne permettent pas de trancher sur la part structurée de ce module. Celui-ci serait effectivement, tout comme le domaine YqeY de *B. subtilis*, composé exclusivement d'hélices α. Il reste la possibilité que le 'linker' présente une structuration dépendant du domaine YqeY (par exemple en s'y enroulant) mais sans présenter de structure secondaire. On tend d'autant plus vers cette hypothèse que l'on peut observer la quasi-totalité des résidus du module

en RMN 2D ce qui suggère habituellement une bonne organisation structurale. Quoi qu'il en soit ces informations nous indiquent que le domaine 'YqeY' ne conserve son intégrité structurale que lorsque le 'linker' est présent. Ceci nous indique qu'il devra être utilisé dans son intégralité pour des études ultérieures.

Propriétés d'empilement cristallin

La capacité de la GlnRS-Dr à former des cristaux de bonne qualité, dont 35% de la structure est désordonnée dans l'empilement, constitue un contre-exemple pour les nombreuses protéines dont les capacités de cristallisation sont effectivement inhibées par la présence de modules ou de régions désordonnées. Cette capacité préconise une bonne homogénéité ainsi qu'une surface de contacts suffisamment étendue pour maintenir l'architecture cristalline. La proportion en solvant et la taille des canaux doivent être suffisamment élevés pour permettre d'accommoder le module en mouvement ou adoptant des conformations diverses.

Perspectives

L'ensemble de ces résultats nous ouvre deux voies d'études de la GlnRS de *D. radiodurans*, l'une dirigée vers l'étude du fonctionnement du site actif et de la boucle d'insertion, l'autre vers la module C-terminal. Ces perspectives impliqueront la poursuite d'une approche pluridisciplinaire alliant les techniques de biochimie et de biologie moléculaire, de cristallographie aux rayons X et de résonance magnétique nucléaire.

Fonction de la boucle ins2

L'étude structurale de la conformation de la boucle *ins2* en présence d'ARNt sera effectuée en présence d'ARNt^{Gln} seul, avec la glutamine, avec la glutamine et l'ATP ou avec un analogue de l'adénylate. Ces études nous permettraient de sonder toutes les étapes de la réaction d'activation et de comprendre s'il existe chez *D. radiodurans* un état intermédiaire supplémentaire conditionné par la présence de la boucle *ins2*. Le rôle du module additionnel C-terminal dans cette étape pourra également être caractérisé. Afin de savoir comment la GlnRS est capable de discriminer la glutamine du proche glutamate, la cristallisation de l'enzyme en présence du glutamate pourrait être envisagée.

Le rôle de la boucle *ins2* dans la discrimination et/ou la capacité d'activation de l'acide aminé pourra être caractérisé par une approche de mutagenèse en la remplaçant par une séquence plus courte comme celle trouvée chez *E. coli*. Inversement, l'impact fonctionnel de l'implantation de l'insertion de *ins2* de *Deinococcus* chez *E. coli* pouvant potentiellement se

traduire par un blocage de l'activité enzymatique, pourra être évalué. Enfin, l'activité inhibitrice des analogues d'adénylate sur l'enzyme devra être caractérisée en présence d'ARNt^{Gln}.

Rôle et structure du module C-terminal

Une approche est de résoudre la structure du domaine par cristallographie aux rayons X en le stabilisant soit avec son substrat d'interaction potentiel, l'ARNt^{Gln}, soit en passant par la cristallisation dans un autre empilement cristallin ou encore en réduisant la taille des canaux de solvant par déshydratation des cristaux. Les résultats préliminaires en RMN 2D sont suffisamment prometteurs pour envisager la résolution du module à l'état libre. Pour évaluer les régions d'interaction, des tests en présence du module marqué en ¹⁵N et de la GlnRS-ΔC ont déjà été engagés et pourront être élargis à l'ARNt^{Gln}. Enfin, afin de comprendre l'importance de la partie 'linker' du module, une série de mutants de délétion de tailles variables seront testés pour observer l'effet de la longueur du peptide de connexion sur la fonctionnalité du domaine YqeY dans la réaction de glutaminylation.

- MATÉRIEL ET MÉTHODES –

I. Clonage des protéines étudiées

Les clones des 4 constructions utilisées étaient déjà disponibles à mon arrivée au laboratoire : (i) la GlnRS de *D.radiodurans* entière (GlnRS-Dr) et (ii) tronquée de son extrémité C-terminale additionnelle (GlnRS-ΔC); (iii) l'extrémité C-terminal additionnelle seule composée du module YqeY et de la séquence reliant le module Yqey au coeur de la GlnRS (YqeY-λ); (iv) le module YqeY seul (YqeY-δ).

Les ORF (Open Reading Frame) des 4 variants ont été amplifiés par PCR à partir du plasmide en utilisant les amorces suivantes :

5'-GGTGGG TGCTCTTCCAACATGGGGGCGTTTGGGTGGGAG-3' (sens pour GlnRS et GlnRS-ΔC); 5'-GGTGGTCTGCAGTCATTAGGCGAGAGCGTCTTTGAGCGC-3' (anti-sens pour GlnRS) et 5'-GGTGGTCTGCAGTCATTAGCCGCCCTGCTTGCCCCAGG-3' (anti-sens pour GlnRS-ΔC); 5'-GGTGGTTGCTCTTCCAACATGACCCAGCAGAAGGCGGAAGGC-3' (sens for YqeY-λ), 5'-GGTGGTTGTTCTTCCAACATGCGCACCATCGCCCGTGACCCG-3' (sens for YqeY-δ) et 5'-GGTGGTCGTCAGTCATTAGGCGACAGCGTCTTTGACGC-3' (anti-sens pour YqeY-λ et YqeY-δ).

Les inserts ont été clonés dans le plasmide pTYB11 contenant une étiquette intéine en N-terminal, un gène de résistance à l'ampicilline et 2 sites de restrictions *SapI* et *PstI*. Ce système se distingue des autres systèmes de fusion de protéines par sa capacité à séparer une protéine recombinante de l'étiquette sans utiliser de protéase. Le plasmide donne naissance à une protéine de fusion composée de l'intéine et du domaine de liaison à la chitine situé en N-terminal de la protéine cible. Le site *SapI* permet au gène cible d'être cloné à un site immédiatement adjacent au site de clivage de l'intéine. Elle induit la purification d'une protéine sans aucun résidu additionnel attaché à son extrémité N-terminale (Deniziak *et al.*, 2007). La souche utilisée ER2566 d'*E. coli* est fournie dans le kit IMPACT-CN System (IMPACT™, New England Biolabs®) et peut-être transformée par le vecteur pTYB11 contenant le gène d'intérêt.

II. Cultures de micro-organismes

A. Milieux de culture

Les bactéries trouvent dans ce milieu les composants indispensables pour leur multiplication rapide et en grand nombre. Sauf indication contraire (*), les milieux et solutions stocks sont stérilisés à 395 K (120°C) par autoclave avant toute culture. La préparation de certains composants du milieu M9 nécessite l'utilisation d'eau stérile et le filtrage des solutions stocks. Les cultures sont effectuées à partir de solutions fraîchement préparées. Les quantités utilisées pour la préparation des milieux sont indiquées dans les tableaux récapitulatifs suivants et valables pour un volume de 1L.

1. Milieu LB (Luria Broth)

Composants	Quantité
Bacto-Tryptone	10 g
Extrait de levure	5 g
NaCl	5 g

2. Milieu M9 (minimum)

Ces milieux sont utilisés entre autres pour les cultures en milieu modifié et en particulier pour l'expression de protéines marquées en ^{15}N (pour application en RMN) ou pour les cultures en présence de sélénométhionine (cristallographie des rayons X).

Solutions stock [10X]

“M9 salts”

Composants	Quantité
Na_2HPO_4	60 g
KH_2PO_4	30 g
NaCl	5 g

Milieu de culture

Composants	Quantité
$^{15}\text{NH}_4\text{Cl}^*$	1 g
M9 salts [10X]	100 mL
Glucose*	4 g
1M MgSO_4	2 mL
Thiamine* 10mg/mL	1 mL
Biotine* 1 mg/mL	1mL
Ampicilline * (100 mM)	1mL

1. Conduite des cultures

Quelle que soit la protéine exprimée ou le type de milieu, les cultures sont effectuées selon le protocole suivant : une préculture de 100 à 150 mL de LB Ampicilline est effectuée durant la nuit. Les cultures sontensemencées en milieu stérile à partir de 50 mL de préculture riche puis cultivées en incubateur à 37°C jusqu'à la $DO_{600}=0.8$ (DO, densité optique). Pour diminuer les risques de toxicité liés à la surexpression, les bactéries sont soumises à un choc thermique et la surexpression est induite à 18°C à l'IPTG (isopropyl-beta-D-thiogalactopyranoside) à 1mM final.

III. Purification des protéines

A. Purification d'affinité sur matrice de chitine

Principe : La matrice de chitine est composée de polysaccharides greffés d'un domaine de liaison à l'intéine. Cette protéine est fusionnée à la protéine d'intérêt en N- ou C-terminal et joue le rôle d'étiquette pour la purification par affinité. La technique permet un clivage autocatalytique de la protéine par simple réaction d'un agent réducteur (β -mercaptoéthanol, DTT ou Dithiothréitol) et permet d'obtenir un produit d'élution sans étiquette.

La construction initiale est constituée de la protéine cible fusionnée à l'intéine qui permet d'une part d'éliminer les risques de toxicité due à l'accumulation de la protéine surexprimée dans la cellule hôte, et, d'autre part, d'obtenir une protéine native sans étiquette à ses extrémités. Cette méthode a été utilisée pour les quatre variants de la GlnRS de *D. radiodurans* (GlnRS-Dr, GlnRS- ΔC , YqeY- λ , YqeY- δ) (Figure III.1).

Protocole : L'extrait brut est chargé à 1mL/min sur les billes de chitine préalablement équilibrées avec 20 mM Tris HCl pH 7,5, 100 mM NaCl, 0,5 mM EDTA (inhibiteur de métalloprotéase), 0,2 mM AEBSF (inhibiteur protéases à sérine). Les billes sont ensuite lavées par 3 volumes de tampon de lavage (20 mM Tris HCl pH 7,5) et incubées en présence de 50 mM DTT pendant 24h. Une première élution est effectuée après 24h avec 20 mM Tris HCl pH 7,5, 0,5 mM EDTA, 0,2 mM AEBSF puis une seconde après réincubation de 24h supplémentaires DTT. Le produit d'élution consiste en deux populations de GlnRS dont la plus courte est obtenue durant l'incubation par autoclivage de la partie C-terminale de la protéine. Bien que les 2 formes soient actives en aminoacylation, il est nécessaire de les séparer lors d'une seconde étape de purification sur colonne héparine.

Le protocole est le même pour les protéines tronquées GlnRS- ΔC , YqeY- λ , YqeY- δ et leur produit d'élution consiste en une forme de taille unique. L'étape suivante est également effectuée sur colonne d'héparine.

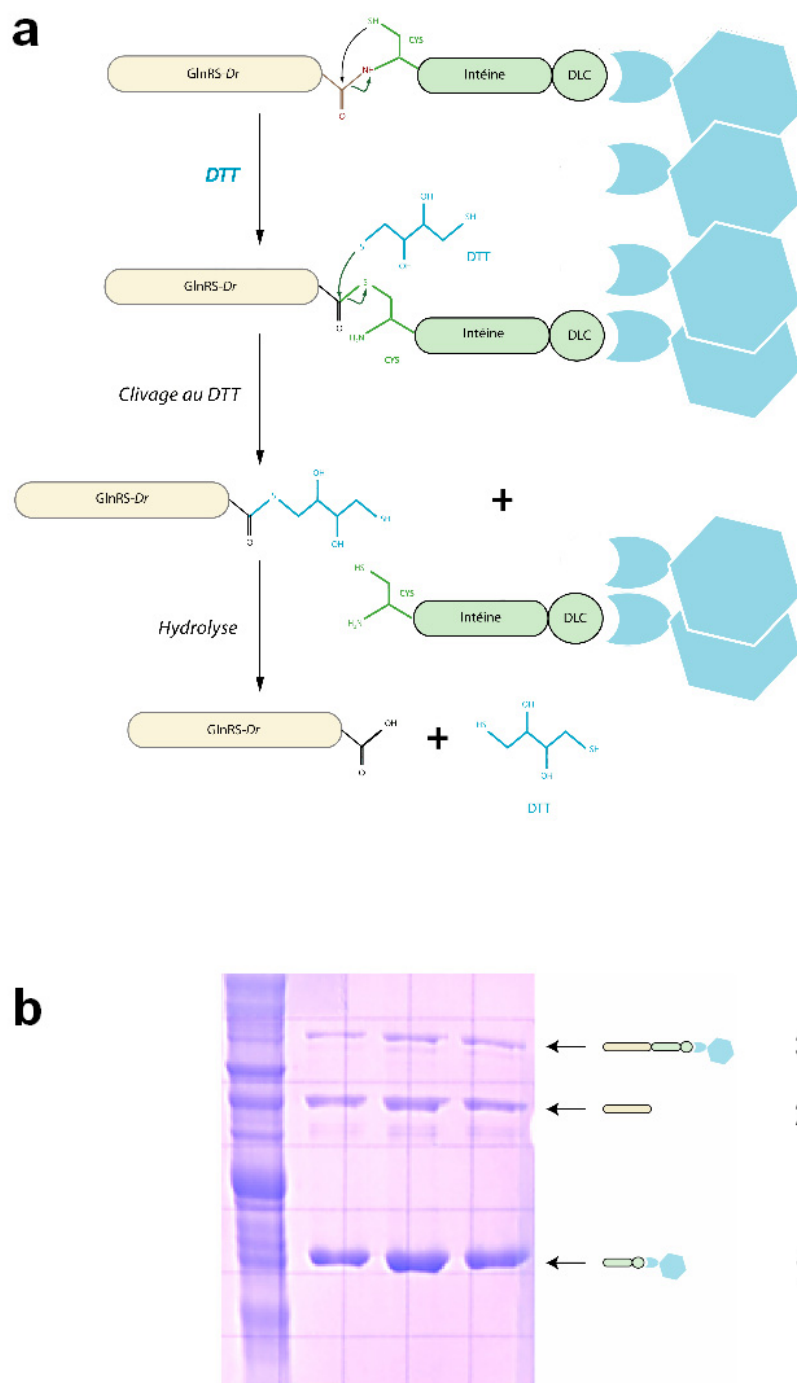


Figure III. 1 : a : Schéma de la réaction d'autoclavage entre la protéine cible (ici GlnRS-Dr) lors de l'incubation en présence de Dithio-thréitol (DTT).- b : Gel SDS en conditions dénaturante d'un extrait de matrice de chitine chargé par la construction fusionnée GlnRS-Intéine et incubée durant la nuit à 50mM DTT. Le produit de la réaction d'autoclavage est constitué de l'intéine qui reste accrochée aux billes de chitine (intéine en vert, billes en bleu, (1.) et de la protéine d'intérêt. Ici, la GlnRS est décrochée de la matrice et élue par simple lavage (GlnRS-Dr, en jaune, 2.). La réaction n'étant pas totale, il subsiste parfois sur la matrice un produit non encore clivé (annoté 3., représentant GlnRS-intéine-chitine).

B. Purification sur matrice substituée à l'héparine

Principe : lors d'une seconde étape, les protéines sont purifiées sur colonne héparine (HiTrap™ heparin HP) échangeuse de cation. L'héparine est formée d'unités disaccharides comportant des groupements sulfates chargés négativement et greffées sur une matrice d'agarose chargée. Le support est caractérisé par une importante densité de charges régulièrement espacées. Elle est connue pour mimer les acides nucléiques et est capable d'établir de fortes interactions avec le domaine cationique d'une protéine. Cet échangeur permet donc de séparer les protéines en fonction de la distribution des charges positives.

Protocole : la colonne est préalablement lavée et équilibrée avec le tampon A (50mM Tris HCl pH 7,5, EDTA, AEBSF, β -mercaptoéthanol). L'extrait protéique est chargé à 0.8mL/min. après lavage de 3 volumes de colonne avec le tampon A, on applique un gradient de force ionique (100 à 300 mM KCl) qui permet de décrocher les protéines graduellement en fonction de leur potentiel électrostatique. Celles-ci sont fractionnées lors de l'élution et un aliquot de chaque fraction est chargé sur gel de polyacrylamide SDS. Les fractions contenant les protéines pures sont rassemblées en vue d'une dialyse puis d'une concentration, ou une dernière étape de purification selon les cas (Figure II. 2).

C. Chromatographie d'exclusion de taille

Principe : une dernière étape peut être nécessaire, en particulier lorsque la protéine est destinée à être utilisée pour les tests de cristallisation. La chromatographie d'exclusion de taille ou "gel filtration" permet de séparer les molécules en fonction de leur poids moléculaire. Ce type de colonne est un bon moyen d'éliminer des "impuretés" ou les différentes formes oligomériques que peuvent prendre la protéine. Elle est également utilisée pour purifier des complexes macromoléculaires de leurs composés non complexés.

Protocole : On utilise une colonne Superdex 200 (Amersham Bioscience) à un débit de 0,5-0,8 mL/min et on étalonne systématiquement la colonne avec un standard de poids moléculaire (Gel filtration standard, Biorad). On utilise un tampon 50mM Tris HCl pH 7,5, 150 mM KCl, EDTA, AEBSF, β mercapto-éthanol) pour l'équilibration. Une fraction de 500 μ L à 2 mL est chargée et l'élution effectuée avec le même tampon. Le produit fractionné de l'élution est chargé sur gel de polyacrylamide SDS et les fractions rassemblées, dialysées, et concentrées.

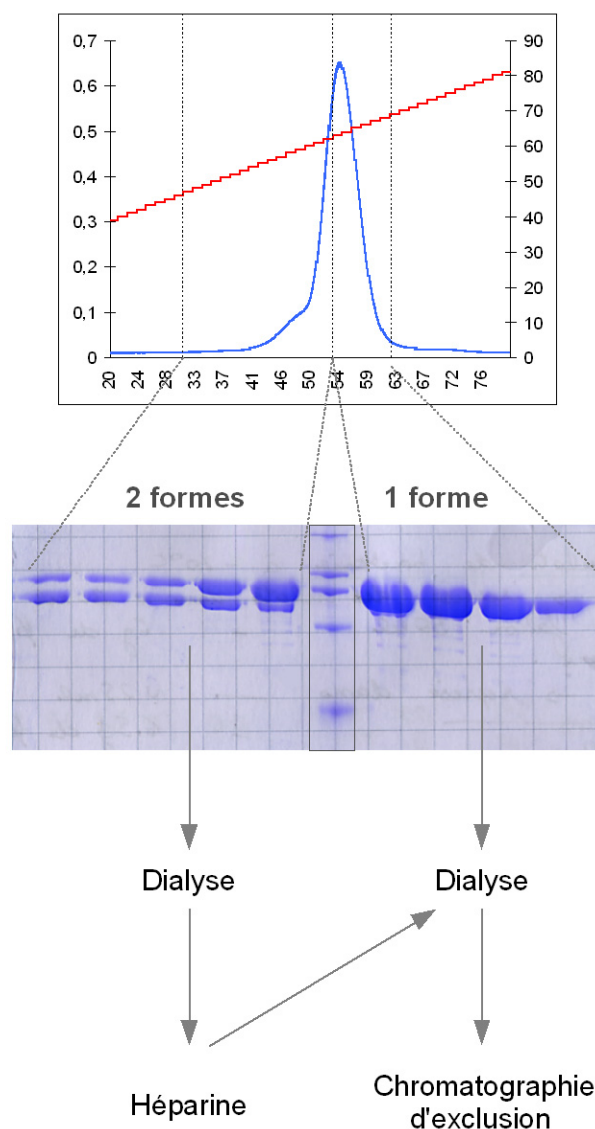


Figure III. 2 : Exemple du profil d'élution obtenu lors de l'étape de purification sur colonne d'héparine. Le produit d'élution sortant en début de gradient (160-175 mM KCl) est constitué de 2 formes de la GlnRS, dont une tronquée en C-terminal (bande inférieure). La forme entière et pure est éluée à une force ionique supérieure (180-200 mM KCl). Les fractions contenant les 2 formes sont regroupées, dialysées et chargées une seconde fois sur héparine. Les fractions pures sont chargées sur chromatographie d'exclusion de taille et concentrées.

D. Séparation des protéines sur gel de polyacrylamide en conditions dénaturantes

Après chaque étape de purification, un aliquot des fractions est chargé sur gel de polyacrylamide en conditions dénaturantes (SDS-PAGE 10% polyacrylamide pour les GlnRS-x, 15% pour les YqeY-x). Le gel migre sous 50 mA d'ampérage avec un tampon 25 mM Tris HCl-Glycine-0,1% SDS. Après révélation par coloration au bleu de Coomassie et décoloration à l'acide acétique, les fractions contenant la protéine d'intérêt sont rassemblées pour une autre étape de purification, dialyse, concentration ou cristallisation.

E. Dialyses

Les fractions purifiées contenant la protéine cible et destinées aux tests cristallisation sont regroupées et dialysées dans des boudins préalablement trempés dans l'eau pure (seuil de coupure de 5000 et 30000 Da selon qu'il s'agit du variant YqeY-x ou GlnRS-x, respectivement) contre un tampon de stockage avec un rapport 100 (50 mM Tris HCl pH7,5, 5 mM β ME, 0,5 mM EDTA) à 4°C pendant la nuit. Les fractions destinées aux expériences fonctionnelles sont stockées à -80°C après deux dialyses contre le même tampon additionné de 10% puis 50%. Pour les expériences de RMN, le tampon Tris HCl est remplacé par le phosphate de potassium à pH 6,5, sans inhibiteur de protéase et évitant tout contact avec le glycérol pour que son signal n'interfère pas avec celui de la protéine. Les expériences de RMN et de cristallographie sont toujours effectuées sur du matériel fraîchement purifié.

F. Concentration

La protéine pure est concentrée par centrifugation dans une cellule de type Vivaspinn 20mL (Sartorius Vivascience) ou Microcon® 500 μ L (Amicon Millipore) suivant les volumes et concentrations ciblés. Les cellules sont préalablement lavées avec du tampon pour ôter les traces de poussière et de glycérol des membranes d'ultrafiltration choisies pour leur faible capacité de rétention des protéines et la taille de leurs pores. Plusieurs cycles de centrifugation à 6000 g à 4°C sont effectués jusqu'à atteindre la concentration souhaitée.

G. Détermination de la concentration

La concentration des échantillons est déterminée par mesure de l'absorbance à 280 nm. Les coefficients d'absorption spécifiques de la protéine ont été calculés à partir de leur séquence par le programme ProtParam <http://expasy.org/tools/protparam.html>. Les mesures sont effectuées sur un spectrophotomètre NanoDrop® ND-1000, dont l'avantage premier est de nécessiter une très petite quantité d'échantillon (1 μ L à 2 μ L) mais également de posséder une zone de linéarité très large (concentrations en protéine 0,1 à 50 mg/mL et en acides nucléiques 2 à 3000 μ g/mL). Les données concernant les 4 formes protéiques de ce travail sont résumées dans le tableau suivant :

Protéines	Nombre de résidus	Masse Molaire (Da)	Concentration mg/mL pour 1 unité A_{280nm}	Coefficient d'extinction molaire ϵ_{280nm}
GlnRS D. radiodurans	852	93470	1,04	97150
GlnRS-ΔC	589	66800	1,32	88240
YqeY-λ	215	21813	0,32	6970
YqeY-δ	165	16771	0,33	5500

IV. Préparation d'ARN de transfert

A. Transcription *in vitro*

Les transcrits d'ARNt^{Gln}(UUG) de *D.radiodurans* ont été obtenus avant mon arrivée au laboratoire par transcription *in vitro* des gènes clonés en aval d'un promoteur de la polymérase du phage T7 selon la technique du transzyme (Fechter *et al.*, 1998).

Le plasmide est linéarisé par BstNI pour introduire dans la séquence un point d'arrêt de la transcription correspondant à l'A76 de l'ARNt. L'ADN linéaire est alors incubé pendant 4h à 37°C dans 500 µL de milieu réactionnel (40 mM Tris HCl pH8,1, 5 mM DTE (dithioerythritol); 22 mM MgCl₂; 1 mM Spermidine; 0,01% Triton X 100; 4 mM A,C,G et UTP et 7,5 µg (120 U) de T7 ARN polymérase). Après 1h il apparaît un précipité de pyrophosphate de magnésium indicateur de l'activité de polymérisation, celui-ci est resolubilisé par l'addition de 10 U pyrophosphatase. La transcription est stoppée par extraction au phénol-chloroforme. Les acides nucléiques sont précipités à l'éthanol, sédimentés et repris dans 150 µL d'eau pure. Les transcrits sont purifiés par électrophorèse sur gel de polyacrylamide dénaturant (8M urée) préparatif.

B. Purification des transcrits sur gel préparatif

Principe : cette étape permet de séparer les ARN transcrits des autres composants du milieu réactionnel à savoir la matrice d'ADN, les nucléotides libres en excès et le pyrophosphate produit lors de l'incorporation, mais aussi le transcrit ciblé des transcrits plus petits, du fait de l'arrêt prématuré de la polymérisation ou plus grands, dû à l'incorporation de nucléotides excédentaires.

Protocole : La solution de transcrit de 250µL, additionnée de 250µL de tampon de dépôt (95% formamide, 1 mM EDTA, 0,025% bleu de bromophénol, 0,025% xylène cyanol) est chargée sur un gel 12% (12% acrylamide, 6% polyacrylamide, 8M urée, tampon TBE ou Tris Borate EDTA) et soumise à un courant de 55 mA. Une pré-electrophorèse "à vide" permet d'éliminer le persulfate oxydant responsable de la polymérisation de l'acrylamide. La bande correspondant au transcrit cible est révélée et découpée sous UV, puis segmentée. Les morceaux de gels sont placés dans une cellule d'électro-élution et mis sous tension dans un tampon TBE stérile à 4°C pendant 2h. Une première fraction de 800µL est collectée après 1h, puis 2h. L'ARN de transfert est précipité à l'éthanol, sédimenté, lavé, séché et repris dans l'eau stérile.

C. Test d'activité des aminoacyl-ARNt synthétases : aminoacylation *in vitro*

Principe : le test d'aminoacylation *in vitro* utilise la faculté qu'ont les des ARNt libres ou aminoacylés à précipiter en milieu acide ce qui permet de les séparer des l'acides aminés libres. La distinction entre les ARNt chargés et les ARNt non chargés est permise par le marquage radioactif des acides aminés. La glutaminylation est suivie en quantifiant la [³H]-glutamine⁷ ou la [¹⁴C]-glutamine.

Protocole : la réaction est initiée par addition de 5 µL d'enzyme diluée dans 100 mM Na-Hepes pH 7,2, (ou 100 mM phosphate de potassium pH 6,5, 0,1 mM EDTA pour les tests préliminaires aux expériences RMN) au milieu réactionnel. Le milieu réactionnel standard (50-100µL) contient 100 mM de Na-HEPES pH 7,2 (ou 100 mM phosphate de potassium pH 6,5, 30 mM KCl), 1mg/mL sérum albumine bovine, 10 mM MgCL₂, 2 mM ATP, 2 mg/mL d'ARNt total de *D. radiodurans* ou 50 µg/mL de transcrit, 20 µM de L-[¹⁴C]-glutamine. Les tests sont réalisés à 37°C dans un volume final de 50 µL. Des aliquots de 10 µL sont prélevés à des temps donnés et déposés sur papier Whatman 3MM (Schleider et Schuell). Pour provoquer la précipitation des ARNt, les carrés de papiers sont plongés dans 5% TCA (acide trichloroacétique) pendant 10 minutes; puis 2 lavages à 5% TCA. Les papiers sont rincés à l'éthanol 100%, séchés et plongés séparément dans des fioles contenant 2,5 mL de liquide de scintillation (Ecoscint OTM, National Diagnostics), La radioactivité portée par les papiers est quantifiée dans un compteur à scintillation (LS6500, Beckman Coulter).

⁷ En cas d'utilisation d'un acide aminé tritié, un facteur de "quenching", caractéristique de l'acide aminé doit être appliqué. Le quenching est dû à l'absorption des acides aminés libres dans le papier, qui diffusent dans ses fibres.

V. Techniques relatives à l'analyse de l'échantillon contenant la protéine

A. Spectrométrie de masse

Cette technique est généralement utilisée pour caractériser la composition d'un échantillon. Son principe est de produire des ions à partir de l'échantillon, de les séparer en fonction de leur masse et de déterminer leur nature à partir des masses observées. Nous avons utilisé les possibilités d'analyse de séquence notamment pour vérifier l'intégrité de la protéine cristallisée et la présence des lysines méthylées. Les 2 méthodes utilisées sont l'ESI (electrospray ionization) ou le MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization – Time Of Fly).

Les expériences sur des cristaux dissous ont été effectuées directement à partir de l'échantillon liquide par MALDI-TOF (voir § VI.A. 4). Le même type d'expérience a été effectué pour les tests de méthylation des lysines (voir § V.C). Des bandes du gel SDS en conditions dénaturantes contenant la protéine ont été préalablement découpées et digérées à la trypsine. Des expériences par ESI ont également été initiées afin d'estimer précisément les variations de poids moléculaires entre les protéines ayant subi le traitement et le témoin.

B. Diffusion de lumière

L'analyse en diffusion de lumière apporte des renseignements sur la taille des particules et leur homogénéité. Les mesures, effectuées sur un appareil dynamique de dispersion de lumière (Protein solution Ltd, High Wycombe, UK), sont collectées à partir d'un échantillon filtré puis centrifugé pendant 5 minutes à 5000 g dans des cuves en quartz.

En admettant que les particules sont des sphères rigides, on peut calculer leur rayon hydrodynamique à partir de leur coefficient de diffusion. En général une protéine pure et homogène est monodisperse, une protéine agrégée ou contaminée est polydisperse. La diffusion de lumière est donc utile pour évaluer la qualité d'un échantillon, son homogénéité et, par conséquent ses chances de cristalliser.

Les expériences de DLS ont été effectuées avant les premiers tests de cristallisation et les tests de dichroïsme circulaire.

C. Méthylation des résidus lysines

Les modifications chimiques peuvent faciliter la cristallisation de protéines ou de domaines solubles ou désordonnés. La méthylation des lysines affecte les propriétés physico-chimiques de la protéine et la rend moins soluble. Elle a été utilisée pour tenter de stabiliser l'extrémité C-terminale désordonnée trouvées dans les cristaux de GlnRS et dont la partie potentiellement flexible est riche en lysines. Les protocoles ont été détaillés récemment (Walter *et al.*, 2006) et un kit commercial est disponible (JBS Methylation Kits).

En pratique, l'échantillon contenant la protéine est dialysé contre un tampon 50 mM Hepes pH 7,5, 250 mM NaCl. On doit éviter tout tampon contenant des groupements amine comme les sels d'ammonium ou le tampon Tris. Un échantillon de 1 mL de protéine concentrée, on ajoute 20 à 100 μ L de réactif A (1 M Diméthyl borane complex) et 40 à 200 μ L de réactif B (1 M formaldéhyde) et on laisse la réaction opérer durant 24h. La GlnRS-Dr étant très riche en lysines (27 résidus), les quantités de réactifs optimales pour 1 mL sont 100 et 200 μ L de réactifs A et B respectivement. Le réactif ne réagit normalement qu'avec les lysines exposées au solvant. Après 24h, l'échantillon est centrifugé afin de retirer les agrégats formés lors de la réaction (parfois jusqu'à 50%) avant de le charger sur une chromatographie d'exclusion de taille. Les fractions éluées contenant le protéine de bonne taille sont recueillies et concentrées pour les tests de cristallisation.

VI. Etudes structurales des protéines

A. Cristallographie des protéines

1. Méthodes de cristallisation

Diverses méthodes de cristallisation des macromolécules ont été développées (Ducruix et Giegé, 1999b). Le principe de la cristallisation consiste à amener la molécule, sans la dénaturer en condition de sursaturation propice à la nucléation puis à la croissance d'un cristal.

a. Diffusion de vapeur

La diffusion de vapeur est la méthode la plus populaire pour la croissance cristalline. Elle consiste à diminuer progressivement la concentration de l'agent cristallisant et de la protéine dans la goutte à partir de la solution du réservoir plus concentrée en agent cristallisant. Les espèces volatiles de la goutte vont diffuser jusqu'à atteindre un équilibre. A ce moment, la pression de vapeur de la goutte est égale à celle du réservoir.

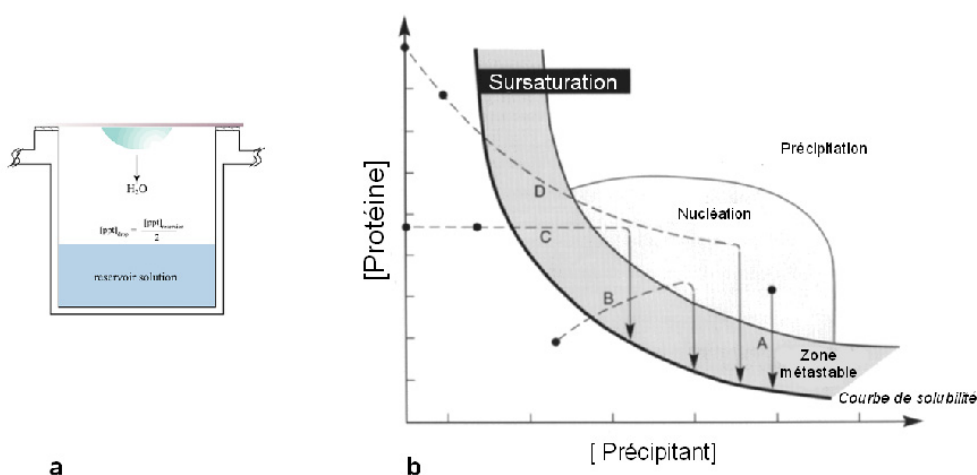


Figure III. 3 : a - Schéma de la technique de diffusion de vapeur. La diffusion de vapeur d'eau conduit, à l'équilibre à une concentration finale de précipitant deux fois plus concentrée dans la goutte que dans le réservoir. **b - Diagramme de phase illustrant les corrélations entre les concentrations de la protéine et de l'agent cristallisant pour quelques techniques de cristallisation.** A : Batch : la concentration en protéine se situe dans la région de sursaturation. B : le point de sursaturation est atteint lorsque la courbe de solubilité est atteinte. En fonction de la concentration finale d'agent cristallisant, on aura précipitation ou cristallisation. C / D : autres techniques : C : Dialyse, D : Diffusion liquide-liquide; Figure d'après Chayen, 1998.

b. Cristallisation en "batch"

La technique du batch consiste à placer la protéine dans des conditions où la sursaturation est instantanément atteinte. Cette technique donne un avantage certain quant à la durée de croissance cristalline. Un autre avantage est la reproductibilité des expériences car les échantillons sont recouverts d'huile de paraffine (imperméable) qui limite l'évaporation et les variations de concentrations non contrôlées, comme c'est le cas dans la diffusion de vapeur. L'utilisation combinée d'huile de paraffine et de silicone (perméable) permet de concentrer graduellement la protéine dans la goutte (Chayen, 1997; D'Arcy *et al.*, 1996).

c. Cristallisation en gel

Principe : la cristallisation en gel d'agarose (extrait d'algue) a été utilisée en particulier en batch pour l'optimisation des conditions de cristallisation. L'environnement gélifié mime l'effet de la microgravité et évite au cristal de sédimenter dans la goutte. Les phénomènes de convection sont largement diminués ce qui conduit à la formation de cristaux de meilleure qualité.

Protocole : l'agarose est un polysaccharide dont les propriétés de polymérisation sont dépendantes de la température. A une concentration de 1 à 2% (m/v), l'agarose est sous forme liquide lorsqu'il est à 90°C, et sous forme de gel à 27°C. En pratique, avant d'être incorporé dans la solution de cristallisation, le gel est liquéfié sur un bloc chauffant à 90°C refroidit à 30°C, puis ajouté à la solution pour une concentration finale de 0,2 à 0,3 % (m/v), la solution devant être maintenue à 30°C. Dans ce cas, l'agarose polymérise en formant des pores de 1µm environ.

d. Cristallisation en champ magnétique

Principe : tout comme la présence de gel dans le mélange de cristallisation, l'application d'un champ magnétique puissant influe sur les courants de convections induits par le champ gravitationnel terrestre. Il peut contrôler les courants et les effets de la sédimentation et serait donc un second outil efficace pour mimer l'environnement spatial induit sous microgravité (Ataka et Wakayama, 2002; Lin *et al.*, 2000). Des effets sur l'orientation du cristal ont été observés par AFM (microscopie à force atomique): le nombre d'îlots individuels à la surface du cristal étant diminuée, celle-ci est donc plus homogène, ce qui améliore la qualité du cristal (Ramachandran et Leslie, 2005).

Protocole : les solutions sont préparées en mélangeant directement l'agent cristallisant et la solution protéique dans une coupelle puis elles sont introduites par capillarité dans des capillaires de 0,3 à 0,5 mm de diamètre. Les capillaires sont ensuite soumis à un champ magnétique dans un spectromètre Bruker DRX 600 MHz pendant la durée de la croissance cristalline. Lorsque l'on obtient des cristaux, ceux-ci sont orientés et peuvent facilement être positionnés pour effectuer la diffraction des rayons X (Figure III. 4).

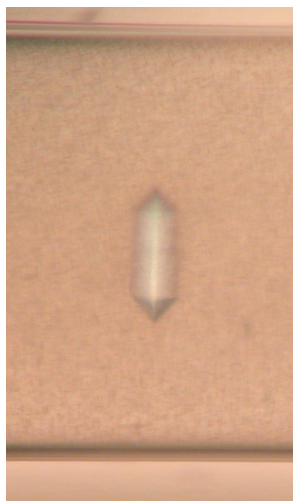


Figure III. 4 : exemple d'un cristal orienté de GlnRS-Dr. Le diamètre du capillaire est de 0,5 mm.

2. Criblage de conditions de cristallisation

Les recherches préliminaires de conditions de cristallisation de la GlnRS et du petit domaine YqeY ont été effectuées par l'intermédiaire de kits de cristallisation commerciaux de type Qiagen (boîtes pré-remplies). Elles utilisent le principe de matrice clairsemée qui rassemble des compositions variées en agents cristallisants, en sels, pH et tampons. Elles ont été réalisées par diffusion de vapeur en gouttes assises d'un volume de 100 à 1000 nL préparées par un robot de cristallisation disponible au laboratoire (Mosquito®, TTP Lab Tech Inc.).

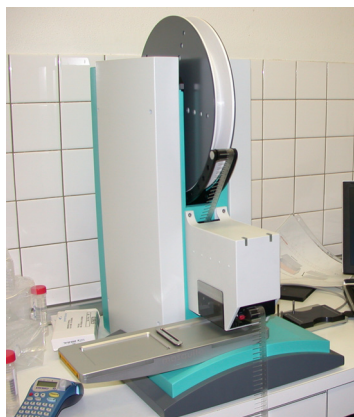


Figure III. 5 : Robot de cristallisation type Mosquito®. Matériel utilisé pour le criblage des conditions de cristallisation.

3. Optimisation des “hits”

Afin d'obtenir des cristaux de bonne qualité, il est souvent nécessaire d'optimiser les conditions de cristallisation trouvées dans le crible. Les conditions les plus prometteuses (les “hits”) sont reproduites puis optimisées en changeant un ou plusieurs paramètres. Il est donc nécessaire de disposer de suffisamment de matériel pour cribler, reproduire, optimiser, de préférence en utilisant le même lot de purification. Les variables les plus souvent choisies sont la concentration en agent cristallisant, le pH, la nature des sels (force ionique) ou des additifs divers (détergents, ligands, substrats...). On peut également jouer sur la concentration en protéine.

4. Observation et analyse du contenu des cristaux

Les cristaux sont observés sur un microscope LEICA M420 et photographiés à l'aide d'un appareil Nikon Coolpix 8400. Le contenu des cristaux peut-être analysé sur gel de polyacrylamide en conditions dénaturantes ou en spectrométrie de masse. Ils sont pour cela d'abord lavés dans plusieurs bains contenant la solution de cristallisation afin d'éliminer les traces de protéine soluble. Ils sont ensuite dissous dans quelques microlitres d'acétate d'ammonium à pH 4,6 avant d'être analysés sur SDS-PAGE ou directement en spectrométrie de masse.

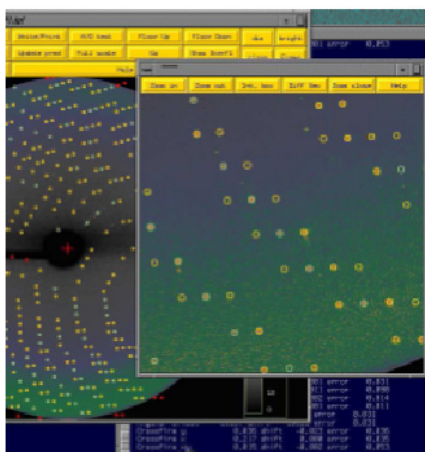
5. Congélation des cristaux

Exposés au rayonnement X le cristal subit des dommages qui peuvent altérer sa qualité de diffraction, en particulier lors de l'utilisation de sources puissantes comme le synchrotron. Pour augmenter sa durée de vie, le cristal est préalablement plongé dans un cryoprotectant (Glycérol, Ethylène glycol, paratone...) avant d'être congelé sous flux d'azote gazeux à -160°C maintenu durant toute l'expérience. La congélation a pour inconvénient d'augmenter la mosaïcité. Des tests de cryoprotection ont été effectués avec les différents produits. La paratone semble être un bon cryoprotectant pour les cristaux de GlnRS.

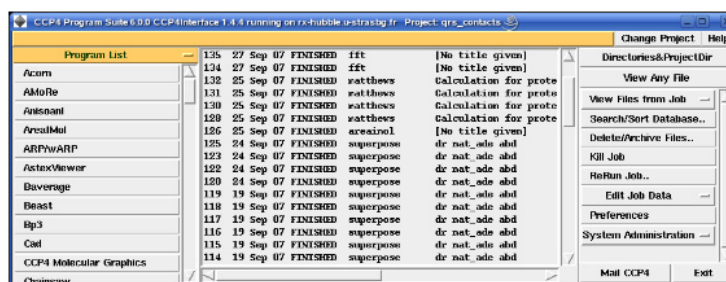
B. Analyse cristallographique

Les cristaux ont été analysés au *European Synchrotron Radiation Facility* (ESRF Grenoble). Les jeux de données ont été collectés à température cryogénique. Les images obtenues ont été traitées à l'aide de la suite HKL (Otwinowski et Minor, 1997) pour les propriétés élémentaires (programme **Denzo** et interface graphique **Xdisp**). Le programme **Scalepack** permet la mise à l'échelle des réflexions : il moyenne les taches équivalents, affine les paramètres de maille du cristal à partir de l'ensemble des données, établit une analyse statistique de la qualité des données (complétude, rapport signal sur bruit ou I/σ , Rmerge). Le calcul des facteurs de structure (Fhkl) à partir des intensités (Ihkl) est effectué à l'aide de **Ccp4** (Collaborative Computational Project, 1994), <http://www.ccp4.ac.uk/>. Le remplacement moléculaire est réalisé à l'aide du programme **MolRep**. Le modèle initial est ensuite affiné dans **Refmac** pour le calcul des premières cartes de densité. L'affinement des coordonnées est effectué par minimisation d'énergie et par recuit simulé avec CNS (Brünger *et al.*, 1998) (<http://cns.csb.yale.edu/>) et la construction avec le programme **Coot** (Emsley et Cowtan, 2004), <http://www.yasbl.york.ac.uk/~emsley/cool/>. voir aussi [Figure III. 6](#).

a



b



c

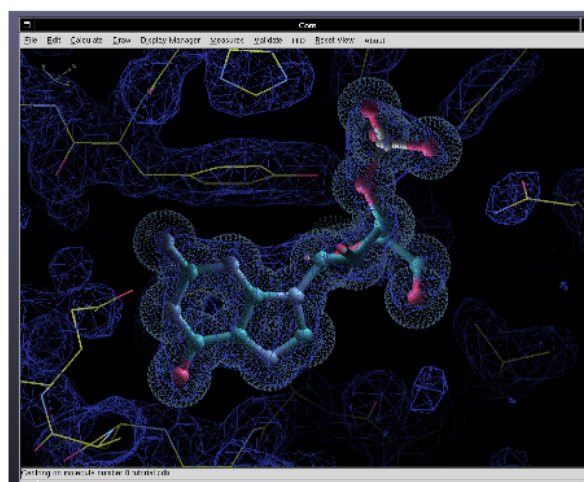


Figure III. 6 : Exemples des résultats obtenus à partir des programmes d'analyse des données cristallographiques. a : Image du traitement d'une image et désignation de propriétés élémentaires (paramètres de maille, groupe d'espace...) par le programme **Denzo** et l'interface graphique **Xdisp**. **b :** Image de la suite de programmes **Ccp4** pour le conversion des intensités (Ihkl) en facteurs de structure (Fhkl), le remplacement moléculaire, le calcul des premières cartes de densité. Il est également utile pour le calcul des contacts cristallins, des surfaces ou des superpositions de structures deux à deux **c :** Carte de densité électronique obtenue à partir du programme **Coot**.

C. Résonance Magnétique Nucléaire (RMN) en biologie structurale

La RMN est une méthode d'étude des macromolécules biologiques en solution. Elle ne nécessite pas l'obtention de cristaux et permet d'étudier des protéines de masse moléculaire allant typiquement de 10 à 30 kDa à des concentrations micro à millimolaires.

Dans le cadre de notre étude une première analyse a été réalisée pour connaître le degré de repliement de la protéine. Un spectre ^1H est effectué pour différencier une protéine repliée d'une protéine non repliée dont le profil est caractéristique. En effet, la dispersion spectrale des fréquences des protons d'une protéine repliée est supérieure à celle de la protéine dénaturée. Les échantillons analysés en RMN sont portés à une concentration de 100 - 300 μM dans un volume de 300 μL . Les spectres sont enregistrés à 285 K avec un spectromètre Bruker DRX 600 MHz équipé d'une cryosonde triple résonance ^1H , ^{15}N , ^{13}C (Figure III. 5).

Une analyse plus poussée consiste à faire des spectres en 2 dimensions. Pour cela, un marquage isotopique est nécessaire. Les isotopes naturellement abondants sont respectivement le ^{14}N et le ^{12}C , mais la plupart des protéines étant obtenues par surexpression bactérienne, il est possible de faire des cultures sur des milieux enrichis isotopiquement. Dans le cas du domaine YqeY, nous avons surexprimé la protéine en milieu marqué ^{15}N après une préculture dans les conditions habituelles (voir § II. A. 2 pour la préparation des milieux). La purification de la protéine marquée suit le protocole habituel.

L'intérêt du marquage de la protéine est d'avoir une estimation plus précise de son degré de repliement. Elle permet également d'estimer *a priori* la qualité des futurs spectres 3D. En effet, l'expression d'une protéine doublement marquée (^{15}N , ^{13}C) nécessite l'utilisation d'un milieu enrichi très coûteux et un spectre préliminaire est donc de mise avant toute purification dans ce type de milieu. Le double marquage de la protéine permet, à partir de la séquence primaire, d'attribuer (ou identifier) un déplacement chimique à un résidu. Une protéine marquée en ^{15}N peut également être soumise à des expériences d'affinité par ajout d'un ligand non marqué à la solution. Une titration peut être faite si la protéine est effectivement en interaction mais les résidus impliqués ne peuvent être attribués que si la structure 3D est résolue.

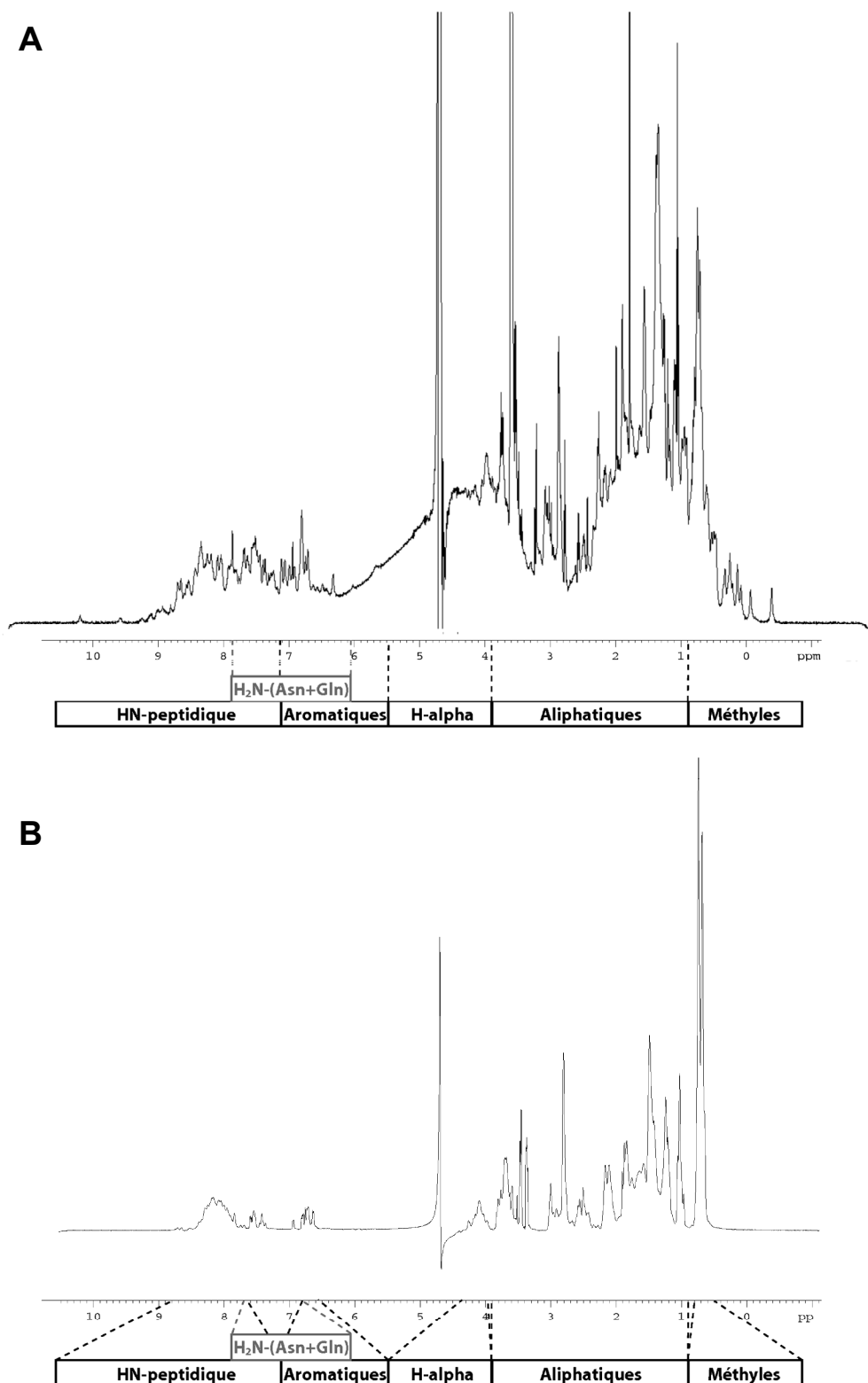


Figure III. 7 : Spectre 1D ^1H à 600MHz, 298K, d'une protéine repliée (A) et non repliée (B). Les valeurs des déplacements chimiques sont indiquées sur l'axe des abscisses et l'emplacement selon la nature des protons indiquée en gras.

D. Dichroïsme circulaire

Principe : le dichroïsme circulaire (ou CD) est une technique spectroscopique d'analyse largement utilisée pour étudier la conformation des acides nucléiques et des protéines. Dans les cas des protéines, il permet d'obtenir des informations qui renseignent sur la nature des différents types de structures secondaires. L'allure d'un spectre CD dépend de l'organisation locale de la liaison peptidique, ce qui permet de distinguer hélices α , feuillets β , "coiled coil" (ou hélices hydrophobes) et structures désordonnées (voir [figure III. 8](#)). Le dichroïsme circulaire donne des informations quantitatives sur la nature des structures secondaires de la protéine étudiée mais ne permet pas de les attribuer à des régions précises.

Protocole : Les tampons utilisés doivent être le plus transparents possible et ne pas contenir de matériel optiquement actif. La région située autour de 200 nm est la plus délicate car c'est dans ces régions que de nombreux tampons absorbent. Une série de protocoles appropriés à l'étude des protéines a été publié récemment (Greenfield, 2006b). L'utilisation d'un tampon Tris HCl est fortement déconseillée. Pour nos expériences, le tampon le plus approprié est 8 mM phosphate de potassium, 40 mM sulfate de sodium, pH 6,5. L'oxygène absorbant à 200 nm, l'échantillon est dégazé à l'Argon pendant 10 minutes. Il est ultracentrifugé à 100 000 g pendant 30 minutes et filtré (0,1-0,2 μ m, Millipore®) pour d'obtenir un mélange homogène et monodisperse en DLS. Les chaînes optiques du spectromètre (Jasco Corp.® J-810) sont également maintenues sous injection d'azote durant l'expérience. L'échantillon est apposé dans des cuvettes de 0,2 cm de parcours optique et les mesures effectuées comme suit : les spectres sont enregistrés de 260 à 190 nm, en vitesse de scan 50 nm/min entre 285 et 298°C. La concentration finale et idéale du domaine YqeY- λ pour ces mesures est portée à 4 μ M.

Traitement des données : le traitement des données par des logiciels comme CD Pro et Dichroweb appropriés permet une estimation de la quantité respective des structures secondaires. Dichroweb est une plateforme de calcul regroupant un ensemble d'algorithmes et 7 bases de données. Le programme utilisé est Contin ou K2D. Le fichier de sortie présente un graphe des données calculées et expérimentales ainsi que 2 facteurs de confiance : NRSMD (Normalized Root Mean Square Deviation) et R-value qui évaluent la corrélation entre les données théoriques et expérimentales (Whitmore et Wallace, 2004).

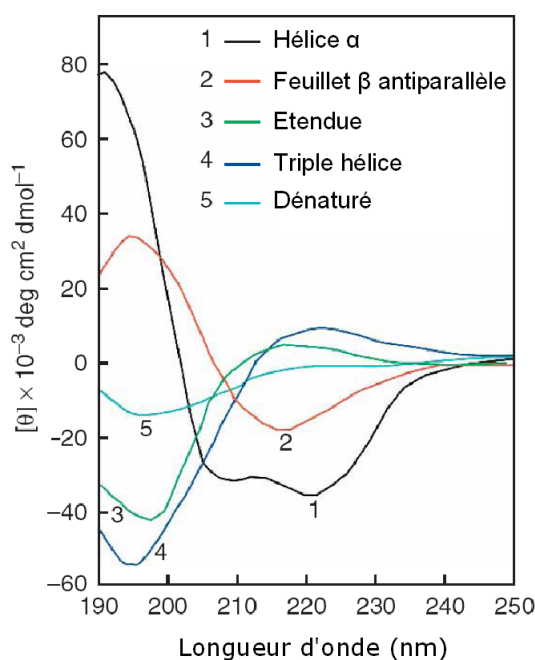


Figure III. 8 : Profil caractéristique des spectres obtenus en Dichroïsme circulaire pour un type de structure secondaire donné. Les protéines en hélice α présentent deux bandes négatives à 222 et 208 nm et une bande positive à 193 nm (noir, 1). Les protéines en feuillets β ont une bande négative à 218 nm et une positive à 195 nm (rouge, 2). Les protéines dénaturées présentent une bande négative vers 195 nm de faible intensité (bleu clair, 5). Figure d'après (Greenfield, 2006a).

VII. Outils bioinformatique

Analyse de séquence et de structure : les alignements de séquences ont été réalisés à l'aide des programmes mis à disposition sur Internet par l'EBI (European Biology Institute, Cambridge) pour Swissprot-TrEMBL (séquences protéiques) ou le programme Blast (recherche de séquences homologues) et Clustal W (alignement multiple) et par l'IGS (Institut de Génomique Structure, Marseille) pour Toffee (alignements multiples).

Les alignements et superpositions de structures et les calculs de déviations standard (ou RMSD), ont été réalisés avec le programme Lsqman (Uppsala Structure Factory, <http://xray.bmc.uu.se/usf/>). Les superpositions ont été effectuées avec un seuil de 3,5 Å. Les valeurs des RMSD sont basées sur les positions équivalentes des atomes de la chaîne principale (N, C α et C) et les superpositions concernent soit les protéines dans leur intégralité, soit les domaines individuels. Le programme Endsript combine 8 différents programmes d'alignement de séquence et d'analyse structurales pour tirer le maximum d'informations à partir d'un fichier PDB :

entre autres la recherche de séquences (BLAST) les alignements multiples (CLUSTAL W), les contacts intermoléculaires (CNS),

De nombreux outils bioinformatiques ont été utilisés de la mise en page d'alignements de séquences (Esript) à la réalisation d'images 3D ou le calcul des potentiels électrostatiques avec Pymol (Delano Inc) et le module APBS.

Les contacts intermoléculaires et intramoléculaires ont été calculés par le programme NCONT de l'interface Ccp4. Le seuil de distance est à 3,5 Å pour les contacts intramoléculaires et 4,5 Å pour les contacts cristallins. Le calcul de distances a également été effectué par le programme *Lsqkab* de l'interface Ccp4 par superposition des structures secondaires entière ou par domaines homologues. Les surfaces de la molécule et les surfaces de contacts ont été calculées par *areaimol* de l'interface Ccp4 (Accessible Surface Area Calculation).

VIII. Outils bureautiques

Les illustrations, figures ou les posters ont été réalisés en utilisant les logiciels libres du système d'exploitation Linux tels que le pack Open Office.org (suite bureautique), Gimp (retouche et création d'images numériques) appuyé de l'ouvrage d'utilisation Suse Linux (Suse Linux user guide, Haefner et al., 2002) et le guide des logiciels libres <http://www.infogiciel.info/article0235.html>. Le programme Illustrator pour Mac a également permis de confectionner certaines figures.

- ANNEXES -

ANNEXE I

La bactérie *Deinococcus radiodurans*

Une bactérie ultrarésistante

Deinococcus radiodurans est une bactérie extrémophile connue pour survivre à l'effet létal d'agents mutagènes puissants tels que l'exposition à de fortes radiations ionisantes et à la dessiccation. Parmi les effets provoqués par ces événements, les cassures double-brins sont considérées comme les plus redoutables des lésions faites à l'ADN. A titre de comparaison, la viabilité de *D. radiodurans* n'est pas affectée par des doses d'irradiations, qui, lorsqu'elles sont 250 fois inférieures, suffisent à stériliser la majorité des cultures bactériennes (Mattimore *et al.*, 1996), (http://www.exobio.cnrs.fr/article.php3?id_article=92). *D. radiodurans* y survit grâce à son génome multicopie et à sa capacité de réparation. Ainsi, lorsqu'il est soumis à des conditions de radiations extrêmes, le génome de *D. radiodurans* est divisé en une centaine de fragments d'ADN qui se réassemblent en un génome fonctionnel. Ce mécanisme, encore peu connu, nécessite deux copies du génome, des cassures aléatoires, l'ADN polymérase I, responsable de la synthèse des brins d'ADN complémentaires, et la protéine RecA impliquée dans la maturation des chromosomes circulaires (Cox et Battista, 2005; Zahradka *et al.*, 2006). La séquence complète de son génome a été publiée (White *et al.*, 1999).

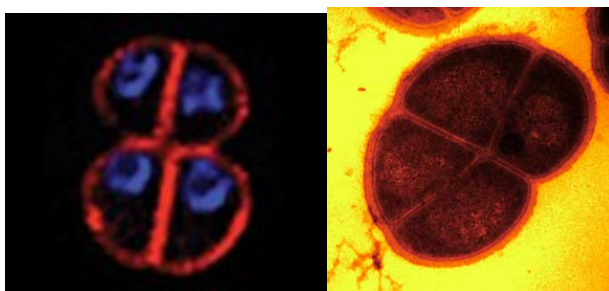


Figure : Une tétrade de *Deinococcus radiodurans*. L'anecdote raconte qu'elle a été découverte en tant que contaminant de viande "stérilisée" par exposition aux rayonnements γ , dans l'Oregon, en 1956 (Anderson *et al.*, 1956). Cette bactérie appartient au genre *Deinococcus* dont le nom découle de Deino, étrange et coccus, rond. Ces espèces sont rares mais trouvées dans des environnements très variés, souvent fortement exposés à la sécheresse. L'image de gauche a été obtenue par fluorescence, l'ADN étant coloré en bleu, les membranes en rouge (Cox et Battista, 2005).

Domaines insertions, une routine pour *Deinococcus*

La combinaison de différents domaines dans une protéine peut entraîner à le développement de fonctions différentes, activer la coopération entre des protéines existantes aux fonctions liées ou faciliter la régulation. *D. radiodurans* code pour 20 protéines multidomaines avec des combinaisons inhabituelles composées par exemple de domaine hydrolase et monooxygénase (gène DR_B0033) ou d'une protéine du cycle cellulaire avec un domaine déaminase (DR_1207) ou encore d'un domaine amidotransférase et un domaine HRD (ou Helicase et RNase D) impliqué dans la liaison aux acides nucléiques (DR_2444) (Makarova *et al.*, 2001; White *et al.*, 1999).

Système traductionnel et aminoacylation espèce spécifique

D. radiodurans ne possède pas de voie standard de biosynthèse de la méthionine et de la cystéine mais est capable de les produire par d'autres voies de biosynthèses. Toutes les enzymes clés nécessaires à la biosynthèse de la lysine sont absentes. Cette bactérie est composée des éléments de la machinerie traditionnelle typique d'une bactérie mais présente quelques propriétés atypiques pour son espèce. Elle possède un répertoire unique de gènes et de réactions de formation de Gln-ARNt^{Gln} et Asn-ARNt^{Asn}. Le génome de *D. radiodurans* code pour des acteurs des voies directes et indirectes de formation des 2 couples : une GlnRS et le trimère GatCAB. GatCAB est capable de catalyser la formation de la Gln-ARNt^{Gln} et l'Asn-ARNt^{Asn} *in vitro* et seulement l'Asn-ARNt^{Asn} *in vivo*, sa GluRS étant discriminante. La présence de GatCAB est donc uniquement destinée à la voie indirecte de synthèse de l'Asn-ARNt^{Asn} (Curnow *et al.*, 1998).

ANNEXES II

Étude structurale de la TyrRS mitochondriale humaine

Durant mon travail de thèse, j'ai également mis à profit mes compétences cristallographiques pour participer à la résolution de la structure de la première synthétase à localisation exclusivement mitochondriale, la TyrRS mitochondriale humaine (mt-TyrRS) dont les résultats sont présentés et discutés dans deux articles (Article 2, Article 3).

Bien qu'elle soit codée par le génome nucléaire, la mt-TyrRS présente plus de similarité de séquence avec les TyrRS bactériennes qu'avec son homologue cytosolique. La mt-TyrRS est, en effet, un système hybride entre bactérie et eucaryote appuyé par le fait qu'elle présente des séquences caractéristiques des TyrRS bactériennes mais aussi d'archées, en particulier dans la région de reconnaissance de l'extrémité 3' de l'ARNt^{Tyr}. Ses particularités fonctionnelles se définissent par sa capacité à fixer efficacement la tyrosine sur l'ARNt^{Tyr} bactérien, d'archée et d'eucaryote, ce qui fait d'elle la première TyrRS qui s'affranchisse de la barrière espèce pour la spécificité de la réaction d'aminoacylation. Enfin, une mutation dans le gène de la TyrRS cytoplasmique humaine est impliquée dans une neuropathie de Charcot-Marie-Tooth, une maladie héréditaire qui affecte le système nerveux périphérique (Jordanova *et al.*, 2006).

L'architecture des TyrRS diverge selon les différents règnes. Si elles sont toutes homodimériques et présentent un domaine catalytique conservé, le module C-terminal de reconnaissance de l'anticodon varie en séquence, en longueur et en organisation : chez les bactéries il est riche en hélice α et renferme un domaine supplémentaire dit "S4-like", homologue en structure et en séquence à la protéine ribosomique S4. Chez les archées et les eucaryotes, le domaine est homologue à celui des TrpRS eucaryotiques et présente un module additionnel pour les eucaryotes vascularisés [revue dans Bonnefond *et al.*, 2005].

Les TyrRS sont probablement les synthétases qui ont été les plus étudiées avec des structures résolues pour quatre bactéries (Brick et Blow, 1987; Kobayashi *et al.*, 2005; Qiu *et al.*, 2001; Yaremchuk *et al.*, 2002), quatre archées (Kobayashi *et al.*, 2003; Kuratani *et al.*, 2006), deux eucaryotes (Tsunoda *et al.*, 2007; Yang *et al.*, 2002), deux mitochondries (Bonnefond *et al.*, 2007; Paukstelis *et al.*, 2005) et un virus (Abergel *et al.*, 2005). La structure de complexes ARNt^{Tyr}/TyrRS est aussi connue dans les trois règnes du vivant (Kobayashi *et al.*, 2003; Tsunoda *et al.*, 2007; Yaremchuk *et al.*, 2002), (Voir l'*Introduction* au § II. C). Ces informations ont permis de comprendre la différence dans le mode de reconnaissance de l'ARNt^{Tyr}. Depuis la résolution de la structure cristallographique de la TyrRS de *B. stearothermophilus* en 1987 le nombre de structures de TyrRS résolues a considérablement augmenté. Néanmoins, de nombreuses questions concernant leur structure, fonction et évolution restent à éclaircir.

Pour ce travail, la TyrRS mitochondriale humaine a été cristallisée sous sa forme entière et dans une version minimale active (incluant les domaines catalytique et de reconnaissance de l'ARNt). Les cristaux obtenus ont permis d'atteindre une résolution de 2.7 Å puis de 2.2 Å en présence d'un analogue de l'intermédiaire réactionnel Tyr-AMP. Tandis que les cristaux de TyrRS de l'enzyme native contiennent un seul monomère dans l'unité asymétrique, ceux de la TyrRS en présence de l'analogue présentent l'homodimère entier (**Article 2**). L'analyse des données a permis de résoudre la première structure d'une aaRS humaine à localisation exclusivement mitochondriale (**Article 3**).

De manière intéressante, les 2 monomères ne sont pas identiques dans le cristal présentant un rmsd global de 0,5 Å qui peut atteindre 3,5 Å pour certains domaines isolés. Au niveau structural, cette TyrRS est très proche des TyrRS des bactéries, telles que *E. coli* et *B. stearothermophilus*. Par contre la TyrRS mitochondriale humaine présente deux insertions dans sa structure, absentes chez les homologues bactériennes. La première insertion, d'une dizaine d'acides aminés, se situe à proximité du site de reconnaissance de l'extrémité acceptrice de l'ARNt^{Tyr} mitochondrial et pourrait être impliqué dans la stabilisation de la branche acceptrice. De plus, cette insertion semble caractéristique des TyrRS mitochondriales de vertébrés. La seconde insertion est longue d'une vingtaine d'acides aminés, et sa séquence est retrouvée chez toutes les TyrRS mitochondriales. Elle pourrait être impliquée dans la fixation de cofacteurs de la réaction d'aminacylation ou dans l'ancrage à la membrane interne de la mitochondrie. La résolution de cette structure a également permis d'analyser le mode de fixation de l'acide aminé ce qui nous a montré que le mécanisme de liaison de l'ATP et de la tyrosine est globalement conservé chez toutes les TyrRS.

J'ai contribué à la partie cristallographique de ce travail plus précisément à l'enregistrement des données, leur traitement et leur affinement, ainsi que la reconstruction de la structure. Les jeux de données conduisant aux meilleurs scores ont fait l'objet d'une analyse plus poussée. Les structures ont été résolues par remplacement moléculaire puis affinées. Je me suis, lors de ce travail, impliquée dans la formation de Luc Bonnefond par l'initiation aux suites de programmes spécialisées (Suite HKL, Ccp4, Cns, Coot; Voir aussi [Matériel et Méthodes](#)). Luc Bonnefond est maintenant apte à résoudre une structure de façon autonome.

Tyrosyl-tRNA synthetase: the first crystallization of a human mitochondrial aminoacyl-tRNA synthetase

L. BONNEFOND, M. FRUGIER, E. TOUZÉ, B. LORBER, C. FLORENTZ, R. GIEGÉ, J. RUDINGER-THIRION and C. SAUTER

Structural biology and crystallization communications, 2007, Vol. 63, Pages 338-341

Pages 135-138 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1107/S1744309107012481>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://www-sicd.u-strasbg.fr/services/peb/>

Crystal Structure of Human Mitochondrial Tyrosyl-tRNA Synthetase Reveals Common and Idiosyncratic Features

Luc BONNEFOND, Magali FRUGIER, Elodie TOUZÉ, Bernard LORBER, Catherine FLORENTZ, Richard GIEGÉ, Claude SAUTER and Joëlle RUDINGER-THIRION

Structure, 2007, Vol. 15, N° 11, Pages 1505-1516

Pages 139-150 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

<http://dx.doi.org/10.1016/j.str.2007.09.018>

La version imprimée de cette thèse peut être consultée à la bibliothèque ou dans un autre établissement via une demande de prêt entre bibliothèques (PEB) auprès de nos services :

<http://www-sicd.u-strasbg.fr/services/peb/>

- RÉFÉRENCES BIBLIOGRAPHIQUES -

RÉFÉRENCES BIBLIOGRAPHIQUES

A

- Abergel, C., Chenivresse, S., Byrne, D., Suhre, K., Arondel, V. and Claverie, J.-M. (2005) Mimivirus TyrRS: preliminary structural and functional characterization of the first amino-acyl tRNA synthetase found in a virus. *Acta Cryst.*, **F61**, 212-215.
- Agris, P.F. (1996) The importance of being modified: Roles of modified nucleosides and Mg²⁺ in RNA structure and function. *Prog. Nucleic Acid Res. Mol. Biol.*, **53**, 79-129.
- Ambrogelly, A., Korencic, D. and Ibba, M. (2002) Functional annotation of class I lysyl-tRNA synthetase phylogeny indicates a limited role for gene transfer. *J Bacteriol*, **184**, 4594-4600.
- An, S. and Musier-Forsyth, K. (2004) Trans-editing of Cys-tRNA^{Pro} by *Haemophilus influenzae* YbaK protein. *J. Biol. Chem.*, **279**, 42359-42362.
- Anderson, A.W., Nordon, H.C., Cain, R.F., Parrish, G. and Duggan, G. (1956) Isolation, morphology, cultural characteristics and resistance to gamma radiation. *Food Technology*, **10**, 575-578.
- Arnez, J., Harris, D.C., Mitschler, A., Rees, B., Francklyn, C. and Moras, D. (1995) Crystal structure of histidyl-tRNA synthetase from *E. coli* complexed with histidyl-adenylate. *EMBO J.*, **14**, 4143-4155.
- Arnez, J.G. and Moras, D. (1997) Structural and functional considerations of the aminoacylation reaction. *Trends Biochem. Sci.*, **22**, 211-216.
- Ataka, M. and Wakayama, N.I. (2002) Effects of a magnetic field and magnetization force on protein crystal growth. Why does a magnet improve the quality of some crystals? *Acta Crystallogr*, **D58**, 1708-1710.

B

- Bahadur, R.P., Chakrabarti, P., Rodier, F. and Janin, J. (2004) A dissection of specific and non-specific protein-protein interfaces. *J Mol Biol*, **336**, 943-955.
- Banerjee, R., Dubois, D., Gauthier, J., Lin, S., Roy, S. and Lapointe, J. (2004) The zinc-binding site of a class I aminoacyl-tRNA synthetase is a SWIM domain that modulates amino acid binding via the tRNA acceptor arm. *Eur. J. Biochem.*, **271**, 724-733.
- Barat C, Lullien V, Schatz O, Keith G, Nugeyre MT, Grüniger-Leitch F, Barré-Sinoussi F, LeGrice SF, Darlix JL. (1989) HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J.*, **8** (11), 3279-3285.
- Basavappa, R. and Sigler, P.B. (1991) The 3Å crystal structure of yeast initiator tRNA: Functional implications in initiator/elongator discrimination. *EMBO J.*, **10**, 3105-3111.
- Becker, H.D., Min, B., Jacobi, C., Racznik, G., Pelaschier, J., Roy, H., Klein, S., Kern, D. and Söll, D. (2000a) The heterotrimeric *Thermus thermophilus* Asp-tRNA^{Asn} amidotransferase can also generate Gln-tRNA^{Gln}. *FEBS Lett.*, **238**, 1-5.
- Becker, H.D., Roy, H., Moulinier, L., Mazauric, M.-H., Keith, G. and Kern, D. (2000b) *Thermus thermophilus* contains an eubacterial and an archaeobacterial aspartyl-tRNA synthetase. *Biochemistry*, **39**, 3216-3230.
- Bedouelle H, Carter P, Waye MM, Winter G, Lowe DM, Wilkinson AJ, Fersht AR., (1985) Engineering of tyrosyl tRNA synthetase, *Biochimie*, **67** (7-8), 737-43.
- Biertümpfel C., Basquin J., Suck D., Sauter C. (2002) Crystallization of biological macromolecules using agarose gel. *Acta Crystallogr D Biol Crystallogr*. **Pt 10 Pt 1**, 1657-1659.
- Blaise, M., Becker, H.D., Lapointe, J., Cambillau, C., Giegé, R. and Kern, D. (2005) Glu-Q-tRNA(Asp) synthetase coded by the yadB gene, a new paralog of aminoacyl-tRNA synthetase that glutamylates tRNA(Asp) anticodon. *Biochimie*, **87**, 847-861.
- Bonnefond, L., Frugier, M., Touze, E., Lorber, B., Florentz, C., Giegé, R., Rudinger-Thirion, J. and Sauter, C. (2007) Tyrosyl-tRNA synthetase: the first crystallization of a human mitochondrial aminoacyl-tRNA synthetase. *Acta Crystallogr*, **F63**, 338-341.
- Bonnefond, L., Giegé, R. and Rudinger-Thirion, J. (2005) Evolution of the tRNA^{Tyr}/TyrRS aminoacylation systems. *Biochimie*, **87**, 873-883.

- Brick, P., Bhat, T.N. and Blow, D.M. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with tyrosyl adenylate intermediate. *J. Mol. Biol.*, **208**, 83-98.
- Brick, P. and Blow, D.M. (1987) Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine. *J. Mol. Biol.*, **194**, 287-297.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.*, **D54**, 905-921.
- Bullard, J., Cai, Y.-C., Demeler, B. and Spemulli, L. (1999) Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. *J. Mol. Biol.*, **288**, 567-577.
- Bullock, T.L., Uter, N., Nissan, T.A. and Perona, J.J. (2003) Amino acid discrimination by a class I aminoacyl-tRNA synthetase specified by negative determinants. *J Mol Biol*, **328**, 395-408.
- Bult, C.J., Xxx, X., Woese, C.R. and Venter, J.C. (1996) Complete genome sequence of the methanogenic archeon, *Methanococcus jannaschii*. *Science*, **273**, 1058-1073.

C

- Cahuzac, B., Berthonneau, E., Birlirakis, N., Guittet, E., Mirande, M., (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases, *EMBO J.*, **19** (3), 445-52.
- Carugo, O. and Argos, P. (1997) Protein-protein crystal-packing contacts. *Protein Sci.*, **6**, 2261-2263.
- Cavarelli, J., Rees, B., Eriani, G., Ruff, M., Boeglin, M., Gangloff, J., Thierry, J.-C. and Moras, D. (1994) The active site of yeast aspartyl-tRNA synthetase: Structural and functional aspects of the aminoacylation reaction. *EMBO J.*, **13**, 327-337.
- Charron, C., Giegé, R. and Lorber, B. (2004) Structure and packing contacts of thaumatin in a hexagonal lattice. *Acta Cryst.*, **D60**, 83-89.
- Chayen, N.E. (1997) The role of oil in macromolecular crystallization. *Current Biol.*, **5**, 1269-1274.
- Chayen, N.E. (1998) Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Cryst.*, **D54**, 8-15.
- Collaborative Computational Project, N., 1994. (1994) The CCP4 Suite: programs for protein crystallography. *Acta Cryst.*, **D50**, 760-763.
- Cox, M.M. and Battista, J.R. (2005) *Deinococcus radiodurans* - the consummate survivor. *Nat Rev Microbiol*, **3**, 882-892.
- Crosio, M.-P., Janin, J. and Jullien, M. (1992) Crystal packing in six crystal forms of pancreatic ribonuclease. *J. Mol. Biol.*, **228**, 243-251.
- Curnow, A.W., Hong, K.W., Yuan, R., Kim, S.I., Martins, O., Winkler, W., Henkin, T.M. and Söll, D. (1997) Glu-tRNA^{Gln} amidotransferase: A novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 11819-11826.
- Curnow, A.W., Ibba, M. and Söll, D. (1996) tRNA-dependent asparagine formation. *Nature*, **382**, 589-590.
- Curnow, A.W., Tumbula, D., Pelaschier, J., Min, B. and Söll, D. (1998) Glutamyl-tRNA^{Gln} amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 12838-12843.
- Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. and Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature*, **347**, 249-255.
- Cusack, S., Hartlein, M. and Leberman, R. (1991) Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res*, **19**, 3489-3498.
- Cusack, S., Yaremchuk, A. and Tukalo, M. (2000) The 2Å crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO J.*, **19**, 2351-2361.

D

- D'Arcy, A., Elmore, C., Stihle, M. and Johnston, J.E. (1996) A novel approach to crystallizing proteins under oil. *J. Crystal Growth*, **168**, 175-180.
- Danbolt, N.C. (2001) Glutamate uptake. *Prog Neurobiol*, **65**, 1-105.
- Dardel, F., Panvert, M. and Fayat, G. (1990) Transcription and regulation of expression of the *Escherichia coli* methionyl-tRNA synthetase gene. *Mol. Gen. Genet.*, **223**, 121-133.
- Delagoutte, B., Moras, D. and Cavarelli, J. (2000) tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding. *EMBO J.*, **19**, 5599-5610.
- Deniziak, M., Sauter, C., Becker, H.D., Paulus, C.A., Giegé, R. and Kern, D. (2007) *Deinococcus* glutaminyl-tRNA synthetase is a chimera between proteins from an ancient and the modern pathways of aminoacyl-tRNA formation. *Nucleic Acids Res.* 1421-1431.
- Deniziak, M.A., Sauter, C., Becker, H.D., Giegé, R. and Kern, D. (2004) Crystallization and preliminary X-ray characterization of the atypical glutaminyl-tRNA synthetase from *Deinococcus radiodurans*. *Acta Cryst. D*, **60**, 2361-2363.
- Diaz-Lazcoz, Y., Aude, J.C., Nitschké, P., Chiapello, H., Landès-Devauchelle, C. and Risler, J.L. (1998) Evolution of genes, evolution of species: the case of aminoacyl-tRNA synthetases. *Mol. Biol. Evol.*, **15**, 1548-1561.
- Doublié, S., Bricogne, G., Gilmore, C. and Carter, C.W. (1995) Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to tyrosyl-tRNA synthetase. *Structure*, **3**, 17-31.
- Dubois, D.Y., Blaise, M., Becker, H.D., Campanacci, V., Keith, G., Giegé, R., Cambillau, C., Lapointe, J. and Kern, D. (2004) An aminoacyl-tRNA synthetase-like protein encoded by the *Escherichia coli* yadB gene glutamylates specifically tRNA^{Asp}. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 7530-7535.
- Ducruix, A. and Giegé, R. (eds.). (1999a) *Crystallization of Nucleic Acids and Proteins: A Practical Approach (second edition)*. IRL Press, Oxford.
- Ducruix, A. and Giegé, R. (1999b) Methods of crystallization. In Ducruix, A. and Giegé, R. (eds.), *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. IRL Press, Oxford, pp. 121-147.

E

- Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst.*, **D60**, 2126-2132.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature*, **347**, 203-206.

F

- Fechter, P., Rudinger, J., Giegé, R. and Théobald-Dietrich, A. (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. *FEBS Lett.*, **436**, 99-103.
- Feng, L., Sheppard, K., Tumbula-Hansen, D. and Soll, D. (2005) Gln-tRNA^{Gln} formation from Glu-tRNA^{Gln} requires cooperation of an asparaginase and a Glu-tRNA^{Gln} kinase. *J. Biol. Chem.*, **280**, 8150-8155.
- Förster, C., Limmer, S., Zeidler, W. and Sprinzl, M. (1994) Effector region of the translation elongation factor EF-Tu-GTP complex stabilizes an orthoester acid intermediate structure of aminoacyl-tRNA in a ternary complex. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 4254-4257.
- Francin, M., Kaminska, M., Kerjan, P. and Mirande, M. (2002) The N-domain of mammalian lysyl-tRNA synthetase is a functional tRNA binding domain. *J. Biol. Chem.*, **277**, 1762-1769.
- Frugier, M., Söll, D., Giegé, R., Florentz, C. (1994) Identity switches between tRNAs aminoacylated by class I glutaminyl- and class II aspartyl-tRNA synthetases, *Biochemistry*, **33**, 9912-21.
- Frugier, M., Moulinier, L. and Giegé, R. (2000) A domain in the N-terminal extension of class IIb eukaryotic aminoacyl-tRNA synthetases is important for tRNA binding. *EMBO J.*, **19**, 2371-2380.

- Frugier, M., Ryckelynck, M. and Giegé, R. (2005) tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. *EMBO Rep.*, **6**, 860-865.
- Fukunaga, R., Yokoyama, S., (2007) Structure of the AlaX-M trans-editing enzyme from *Pyrococcus horikoshii*. *Acta Crystallogr D Biol Crystallogr.* **63**, 390-400.

G

- Giegé R, Lorber B, Ebel JP, Moras D, Thierry JC., (1980) Crystallization of the complex formed between yeast aspartyl tRNA and its specific aminoacyl tRNA synthetase, *C R Seances Acad Sci D.*, **29**, 291(4), 393-396.
- Giegé, R., Moras, D. and Thierry, J.-C. (1977) Yeast transfer RNA^{Asp}: A new high resolution X-ray diffracting crystal form of a transfer RNA. *J. Mol. Biol.*, **115**, 91-96.
- Giegé, R., Sissler, M. and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res.*, **26**, 5017-5035.
- Greenfield, N.J. (2006a) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc*, **1**, 2527-2535.
- Greenfield, N.J. (2006b) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*, **1**, 2876-2890.
- Gustilo, E.M., Dubois, D.Y., Lapointe, J. and Agris, P.F. (2007) E. coli glutamyl-tRNA Synthetase is inhibited by anticodon stem-loop domains and a Minihelix. *RNA Biol*, **4**.

H

- Hampel A, Labanauskas M, Connors PG, Kirkegard L, RajBhandary UL, Sigler PB, Bock RM. (1968) Single crystals of transfer RNA from formylmethionine and phenylalanine transfer RNA's. *Science* , **162**(860),1384-1387.
- Han, J.M., Kim, J.Y. and Kim, S. (2003) Molecular network and functional implications of macromolecular tRNA synthetase complex. *Biochem Biophys Res Commun*, **303**, 985-993.
- Hayase, Y., Jahn, M., Rogers, M.J., Sylvers, L.A., Koizumi, M., Inoue, H., Ohtsuka, E. and Söll, D. (1992) Recognition of bases in *E. coli* tRNA^{Gln} by glutamyl-tRNA synthetase: a complete identity set. *EMBO J.*, **11**, 4159-4165.

I

- Ibba, M., Morgan, S., Curnow, A.W., Pridmore, D.R., Vothknecht, U.C., Gardner, W., Lin, W., Woese, C.R. and Söll, D. (1997) A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases. *Science*, **278**, 1119-1122.
- Ibba, M. and Söll, D. (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.*, **69**, 617-650.
- Ilyin, V.A., Temple, B., Hu, M., Li, G., Yin, Y., Vachette, P. and Carter, C.W., Jr. (2000) 2.9 Å crystal structure of ligand-free tryptophanyl-tRNA synthetase: domain movements fragment the adenine nucleotide binding site. *Protein Sci.*, **9**, 218-231.
- Isel C, Ehresmann C, Keith G, Ehresmann B, Marquet R. (1995) Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA₃^{Lys} (template/primer). *J. Mol. Biol.*, **247**(2), 236-50.

J

- Jakubowski, H. (1999) Misacylation of tRNA^{Lys} with noncognate amino acids by lysyl-tRNA synthetase. *Biochemistry*, **38**, 8088-8093.
- Janin, J. (1997) Specific versus non-specific contacts in protein crystals. *Nature Struct. Biol.*, **4**, 973-974.
- Jordanova, A., Irobi, J., Thomas, F.P., Dijck, P.V., Meerschaert, K., Dewil, M., Dierick, I., Jacobs, A., Vriendt, E.D., Guergueltcheva, V., Rao, C.V., Tournev, I., Gondim, F.A.A., D'Hooghe, M., Gerwen, V.V., Callaerts, P., Bosch, L.V.D., Timmermans, J.-P., Robberecht, W.,

Gettemans, J., Thevelein, J.M., Jonghe, P.D., Kremensky, I. and Timmerman, V. (2006) Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. *Nat Genet.*, **38**, 197-202.

K

- Kavran JM, Gundllapalli S, O'Donoghue P, Englert M, Söll D, Steitz TA.(2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl. Acad. Sci.*, **104** (27),11268-11273
- Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, Wang AH, Seeman NC, Rich A. (1974) Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science*, **185**, 435-40.
- Kim, Y., Shin, J., Li, R., Cheong, C., Kim, K. and Kim, S. (2000) A novel anti-tumor cytokine contains an RNA binding motif present in aminoacyl-tRNA synthetases. *J.Biol.Chem.*, **275**, 27062--27068.
- Kisselev, L.L. and Wolfson, A.D. (1994) Aminoacyl-tRNA synthetases from higher eukaryotes. *Prog. Nucl. Acids Res. Mol. Biol.*, **48**, 83-142.
- Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K. and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. *Nat. Struct. Biol.*, **10**, 425-432.
- Kobayashi, T., Takimura, T., Sekine, R., Vincent, K., Kamata, K., Sakamoto, K., Nishimura, S. and Yokoyama, S. (2005) Structural Snapshots of the KMSKS Loop Rearrangement for Amino Acid Activation by Bacterial Tyrosyl-tRNA Synthetase. *J. Mol. Biol.*, **346**, 105-117.
- Kundrot, C.E. (2004) Which strategy for a protein crystallization project? *Cell Mol Life Sci*, **61**, 525-536.
- Kuratani, M., Sakai, H., Takahashi, M., Yanagisawa, T., Kobayashi, T., Murayama, K., Chen, L., Liu, Z.-J., Wang, B.-C., Kuroishi, C., Kuramitsu, S., Terada, T., Bessho, Y., Shirouzu, M., Sekine, S.-I. and Yokoyama, S. (2006) Crystal Structures of Tyrosyl-tRNA Synthetases from Archaea. *J. Mol. Biol.*, **355**, 395-408.

L

- Lamour, V., Quevillon, S., Diriong, S., N'guyen, V.C., Lipinski, M. and Mirande, M. (1994) Evolution of the Glx-tRNA synthetase family: the glutamyl enzyme as a case of horizontal gene transfer. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 8670-8674.
- Lee, S., Cho, B., Park, S. and Kim, S. (2004) Aminoacyl-tRNA synthetase complexes: beyond translation. *J. Cell Science*, **117**, 3725-3734.
- Levengood JD, Roy H, Ishitani R, Söll D, Nureki O, Ibba M. (2007) Anticodon recognition and discrimination by the alpha-helix cage domain of class I lysyl-tRNA synthetase. *Biochemistry*, **46**(39), 11033-11038.
- Lin, S.X., Baltzinger, M. and Remy, P. (1983) Fast kinetic study of yeast phenylalanyl-tRNA synthetase: an efficient discrimination between tyrosine and phenylalanine at the level of the aminoacyladenylate-enzyme complex. *Biochemistry*, **22**, 681-689.
- Lin, S.X., Zhou, M., Azzi, A., Xu, G.J., Wakayama, N.I. and Ataka, M. (2000) Magnet used for protein crystallization: novel attempts to improve the crystal quality. *Biochem Biophys Res Commun*, **275**, 274-278.
- Lincecum, T.L., Jr., Tukalo, M., Yaremchuk, A., Mursinna, R.S., Williams, A.M., Sproat, B.S., Van Den Eynde, W., Link, A., Van Calenbergh, S., Grotli, M., Martinis, S.A. and Cusack, S. (2003) Structural and mechanistic basis of pre- and posttransfer editing by leucyl-tRNA synthetase. *Mol. Cell*, **11**, 951-963.
- Ling, J., Roy, H. and Ibba, M. (2007) Mechanism of tRNA-dependent editing in translational quality control. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 72-77.
- Lipman, R.S., Chen, J., Evilia, C., Vitseva, O. and Hou, Y.M. (2003) Association of an aminoacyl-

- tRNA synthetase with a putative metabolic protein in archaea. *Biochemistry*, **42**, 7487-7496.
- Liu, J., Gagnon, Y., Gauthier, J., Furenlid, L., L'Heureux, P.J., Auger, M., Nureki, O., Yokoyama, S. and Lapointe, J. (1995) The zinc-binding site of Escherichia coli glutamyl-tRNA synthetase is located in the acceptor-binding domain. Studies by extended x-ray absorption fine structure, molecular modeling, and site-directed mutagenesis. *J.Biol.Chem.*, **270**, 15162-15169.
- Liu, Y., Liao, J., Zhu, B., Wang, E.D. and Ding, J. (2006) Crystal structures of the editing domain of Escherichia coli leucyl-tRNA synthetase and its complexes with Met and Ile reveal a lock-and-key mechanism for amino acid discrimination. *Biochem J*, **394**, 399-407.
- Lorber, B., Giegé, R., Ebel, J.P., Berthet, C., Thierry, J.C., Moras, D. (1983) Crystallization of a tRNA aminoacyl-tRNA synthetase complex. Characterization and first crystallographic data. *J.Biol.Chem.*, **258** (13), 8429-8435.
- Lorber B, Kern D, Mejdoub H, Boulanger Y, Reinbolt J, Giege R. (1987) Comparative efficacy and safety of ceftizoxime, cefotaxime and latamoxef in the treatment of bacterial pneumonia in high risk patients. *J Antimicrob Chemother.*, **19** (2), 239-248.
- Lorber, B., Skouri, M., Munch, J.-P. and Giegé, R. (1993) The influence of impurities on protein crystallization. The case of lysozyme. *J. Crystal Growth*, **128**, 1203-1211.
- Lorber, B., Sauter, C., Robert, M.-C., Capelle, B. and Giegé, R. (1999) Crystallization within agarose gel in microgravity improves the quality of thaumatin crystals. *Acta Cryst.*, **D55**, 1491-1494.

M

- Makarova, K.S., Aravind, L., Wolf, Y.I., Tatusov, R.L., Minton, K.W., Koonin, E.V. and Daly, M.J. (2001) Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed from the perspective of comparative genomics. *Microbiol Mol Biol Rev*, **65**, 44-79.
- Marzi S, Myasnikov AG, Serganov A, Ehresmann C, Romby P, Yusupov M, Klaholz BP. (2007) Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell*. **130** (6), 1019-1031.
- Mattimore V, Battista JR. (1996) Radioresistance of Deinococcus radiodurans: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *JBacteriol*. **178**(3), 633-637.
- McPherson, A. and Cudney, B. (2006) Searching for silver bullets: an alternative strategy for crystallizing macromolecules. *J. Struct. Biol.*, **156**, 387-406.
- Mechulam Y, Schmitt E, Maveyraud L, Zelwer C, Nureki O, Yokoyama S, Konno M, Blanquet S. (1999) Crystal structure of Escherichia coli methionyl-tRNA synthetase highlights species-specific features. *J. Mol. Biol.*, **294**(5), 1287-1297.
- Mirande, M. (1991) Aminoacyl-tRNA synthetase family from prokaryotes and eukaryotes: Structural domains and their implications. *Prog. Nucleic Acid Res. Mol. Biol.*, **40**, 95-142.
- Moras, D., Comarmond, M.B., Fischer, J., Weiss, R., Thierry, J.-C., Ebel, J.-P. and Giegé, R. (1980) Crystal structure of yeast tRNA^{Asp}. *Nature*, **288**, 669-674.
- Mosyak, L., Reshetnikova, L., Goldgur, Y., Delarue, M. and Safo, M.G. (1995) Structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus*. *Nature Struct. Biol.*, **2**, 537-547.
- Murzin AG, Brenner SE, Hubbard T, Chothia C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.*, **247**, 536-40.

N

- Nagel, G.M. and Doolittle, R.F. (1991) Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 8121-8125.
- Nakamura, A., Yao, M., Chimnarong, S., Sakai, N. and Tanaka, I. (2006) Ammonia channel couples glutaminase with transamidase reactions in GatCAB. *Science*, **312**, 1954-1958.
- Ng, J.D., Kuznetsov, Y.G., Malkin, A.J., Keith, G., Giegé, R. and McPherson, A. (1997a) Visualization of RNA crystal growth by Atomic Force Microscopy. *Nucleic Acids Res.*, **25**, 2582-2588.

- Ng, J.D., Lorber, B., Giegé, R., Koszelak, S., Day, J., Greenwood, A. and McPherson, A. (1997b) Comparative analysis of thaumatin crystals grown on earth and in microgravity. *Acta Cryst.*, **D53**, 724-733.
- Nureki, O., Vassylyev, D.G., Tateno, M., Shimida, A., Nakama, T., Fukai, S., Konno, M., Hendrickson, T., Schimmel, P. and Yokoyama, S. (1998) Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science*, **280**, 578-582.

O

- O'Donoghue, P. and Luthey-Schulten, Z. (2003) On the evolution of structure in aminoacyl-tRNA synthetases. *Microbiol. Mol. Biol. Rev.*, **67**, 550-573.
- Oshikane H, Sheppard K, Fukai S, Nakamura Y, Ishitani R, Numata T, Sherrer RL, Feng L, Schmitt E, Panvert M, Blanquet S, Mechulam Y, Söll D, Nureki O. (2006) Structural basis of RNA-dependent recruitment of glutamine to the genetic code. *Science*, **312**, 1950-1954.
- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307-326.

P

- Park, S.G., Jung, K.H., Lee, J.S., Jo, Y.J., Motegi, H., Kim, S., Shiba, K. (1999), Precursor of pro-apoptotic cytokine modulates aminoacylation activity of tRNA synthetase. *J.Biol.Chem...*, **274**(24), 16673-16676.
- Park, S.G., Ewalt, K.L. and Kim, S. (2005) Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends Biochem. Sci.*, **30**, 569-574.
- Paukstelis, P.J., Coon, R., Madabusi, L., Nowakowski, J., Monzingo, A., Robertus, J. and Lambowitz, A.M. (2005) A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Molecular Cell*, **17**, 417-428.
- Perona, J.J., Rould, M.A. and Steitz, T.A. (1993) Structural basis for transfer RNA aminoacylation by glutaminyl-tRNA synthetase. *Biochemistry*, **32**, 8758-8771.
- Perret V, Florentz C, Puglisi JD, Giegé R. (1992) Effect of conformational features on the aminoacylation of tRNAs and consequences on the permutation of tRNA specificities. *J Mol Biol.*, **226**, 323-333.
- Perutz, M.F., (1960) Structure of hemoglobin. *Brookhaven Symp Biol.*, **13**, 165-183.

Q

- Qiu, X., Janson, C.A., Smith, W.W., Green, S.M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C.S.V. and Jarvest, R.L. (2001) Crystal structure of *Staphylococcus aureus* tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. *Protein Sci.*, **10**, 2008-2016.
- Quevillon, S., Robinson, J.-C., Berthonneau, E., Siatecka, M. and Mirande, M. (1999) Macromolecular assemblage of aminoacyl-tRNA synthetases: Identification of protein-protein interactions and characterization of a core protein. *J. Mol. Biol.*, **285**, 183-195.

R

- Ramachandran, N. and Leslie, F.W. (2005) Using magnetic field to control convection during protein crystallization -- analysis and validation studies. *J. of Crystal Growth* **274**, 297-306.
- Rauert, W., Eddine, A.N., Kaufmann, S.H., Weiss, M.S. and Janowski, R. (2007) Reductive methylation to improve crystallization of the putative oxidoreductase Rv0765c from *Mycobacterium tuberculosis*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **63**, 507-511.
- Retailleau, P., Huang, X., Yin, Y., Hu, M., Weinreb, V., Vachette, P., Vonnrhein, C., Bricogne, G., Roversi, P., Ilyin, V. and Carter, C.W. (2003) Interconversion of ATP binding and conformational free energies by tryptophanyl-tRNA synthetase: Structures of ATP bound to open and closed, pre-transition-state conformations. *J. Mol. Biol.*, **325**, 39-63.
- Ribas de Pouplana, L. and Schimmel, P. (2001a) Aminoacyl-tRNA synthetases: potential markers of genetic code development. *Trends Biochem. Sci.*, **26**, 591-596.
- Ribas de Pouplana, L. and Schimmel, P. (2001b) Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell*, **104**, 191-193.
- Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BF, Klug A. (1974) Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature*. **250**, 546-551.
- Robinson, J.C., Kerjan, P. and Mirande, M. (2000) Macromolecular assemblage of aminoacyl-tRNA synthetases: quantitative analysis of protein-protein interactions and mechanism of complex assembly. *J Mol Biol*, **304**, 983-994.
- Rogers KC, Crescenzo AT, Söll D. (1995) Aminoacylation of transfer RNAs with 2-thiouridine derivatives in the wobble position of the anticodon. *Biochimie*. **77**, 66-74.
- Romby, P., Caillet, J., Ebel, C., Sacerdot, C., Graffe, M., Eyermann, F., Brunel, C., Moine, H., Ehresmann, C., Ehressman, B. and Springer, M. (1996) The expression of *E. coli* threonyl-tRNA synthetase is regulated at the translational level by symmetrical operator-repressor interactions. *EMBO J.*, **15**, 5976-5987.
- Rossmann, M.G., Moras, D. and Olsen, K.W. (1974) Chemical and biological evolution of a nucleotide-binding protein. *Nature*, **250**, 194-199.
- Rould, M.A., Perona, J.J., Söll, D. and Steitz, T.A. (1989) Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA^{Gln} and ATP at 2.8 Å resolution. *Science*, **246**, 1135-1142.
- Roy, H., Ling, J., Irnov, M. and Ibba, M. (2004) Post-transfer editing in vitro and in vivo by the beta subunit of phenylalanyl-tRNA synthetase. *EMBO J*, **23**, 4639-4648.
- Ruff, M., Cavarelli, J., Mikol, V., Lorber, B., Mitschler, A., Giegé, R., Thierry, J.-C. and Moras, D. (1988) A high resolution diffracting crystal form of the complex between yeast tRNA^{Asp} and aspartyl-tRNA synthetase. *J. Mol. Biol.*, **201**, 235-236.
- Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C. and Moras, D. (1991) Class II aminoacyl transfer RNA synthetases: Crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^{Asp}. *Science*, **252**, 1682-1689.
- Ryckelynck, M., Giegé, R. and Frugier, M. (2005a) tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. *Biochimie*, **87**, 835-845.
- Ryckelynck, M., Masquida, B., Giegé, R. and Frugier, M. (2005b) An Intricate RNA Structure with two tRNA-derived motifs directs complex formation between yeast Aspartyl-tRNA Synthetase and its mRNA. *J. Mol. Biol.* **354**(3), 614-29.
- Rymo, L. and Lagerkvist, U. (1970) Crystallization of lysyl-tRNA synthetase from yeast. *Nature*, **226**, 77.

S

- Salazar JC, Ambrogelly A, Crain PF, McCloskey JA, Söll D. (2004) A truncated aminoacyl-tRNA synthetase modifies RNA. *Proc Natl Acad*, **101**, 7536-7541.
- Sareen, D., Steffek, M., Newton, G.L. and Fahey, R.C. (2002) ATP-dependent L-cysteine:1D-myo-inositol 2-amino-2-deoxy-alpha-D-glucopyranoside ligase, mycothiol biosynthesis enzyme MshC, is related to class I cysteinyl-tRNA synthetases. *Biochemistry*, **41**, 6885-6890.

- Sampath P, Mazumder B, Seshadri V, Gerber CA, Chavatte L, Kinter M, Ting SM, Dignam JD, Kim S, Driscoll DM, Fox PL. (2004) Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell*. **119**, 195-208.
- Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., Ibb, M. and Söll, D. (2005) RNA-dependent cysteine biosynthesis in Archaea. *Science*, **307**, 1969-1972.
- Sauter, C., Lorber, B., Cavarelli, J., Moras, D. and Giegé, R. (2000) The free yeast aspartyl-tRNA synthetase differs from the tRNA^{Asp}-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon binding domain. *J. Mol. Biol.*, **299**, 1313-1324.
- Schmitt, E., Panvert, M., Blanquet, S. and Mechulam, Y. (1995) Transition state stabilization by the 'high' motif of class I aminoacyl-tRNA synthetases: the case of Escherichia coli methionyl-tRNA synthetase. *Nucleic Acids Res*, **23**, 4793-4798.
- Schmitt, E., Panvert, M., Blanquet, S. and Mechulam, Y. (2005) Structural Basis for tRNA-Dependent Amidotransferase Function. *Structure*, **13**, 1421-1433.
- Schray B, Knippers R. (1991) Binding of human glutamyl-tRNA synthetase to a specific site of its mRNA. *Nucleic Acids Res*. **19**, 5307-5312.
- Schulze JO, Masoumi A, Nickel D, Jahn M, Jahn D, Schubert WD, Heinz DW., (2006) Evolutionary relationship between initial enzymes of tetrapyrrole biosynthesis. *J. Mol. Biol.*, **358**, 1212-1220.
- Sekine, S., Nureki, O., DuBois, D.Y., Bernier, S., Chênevert, R., Lapointe, J., Vassilyev, D. and Yokoyama, S. (2003) ATP binding by glutamyl-tRNA synthetase is switched to the productive mode by tRNA binding. *EMBO J.*, **22**, 676-688.
- Sekine S, Shichiri M, Bernier S, Chênevert R, Lapointe J, Yokoyama S. (2006) Structural bases of transfer RNA-dependent amino acid recognition and activation by glutamyl-tRNA synthetase. *Structure*. **14**, 1791-1799.
- Sherlin, L. and Perona, J.J. (2003) tRNA-dependent active site assembly in a class I aminoacyl-tRNA synthetase. *Structure*, **11**, 591-603.
- Siatecka, M., Rozek, M., Barciszewski, J. and Mirande, M. (1998) Modular evolution of the Glx-tRNA synthetase family--rooting of the evolutionary tree between the bacteria and archaea/eukarya branches. *Eur J Biochem*, **256**, 80-87.
- Silvian, L.F., Wang, J. and Steitz, T.A. (1999) Insights into editing from an ile-tRNA synthetase structure with tRNA^{ile} and mupirocin. *Science*, **285**, 1074-1077.
- Simader, H., Hothorn, M., Kähler, C., Basquin, J., Simos, G. and Suck, D. (2006a) Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. *Nucleic Acids Res*. **62**(Pt 12), 1510-1519
- Simader, H., Hothorn, M. and Suck, D. (2006b) Structures of the interacting domains from yeast glutamyl-tRNA synthetase and tRNA-aminoacylation and nuclear-export cofactor Arc1p reveal a novel function for an old fold. *Acta Crystallogr*, **D62**, 1510--1519.
- Simos, G., Sauer, A., Fasiolo, F. and Hurt, E.C. (1998) A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol. Cell*, **1**, 235-242.
- Sissler, M., Delorme, C., Bond, J., Ehrlich, S.D., Renault, P. and Francklyn, C. (1999) An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 8985-8990.
- Stathopoulos, C., Kim, W., Li, T., Anderson, I., Deutsch, B., Palioura, S., Whitman, W. and Söll, D. (2001) Cysteiny-tRNA synthetase is not essential for viability of the archaeon *Methanococcus maripaludis*. *Proc. Natl. Acad. Sci.*, **98**, 14292-14297.
- Stout CD, Mizuno H, Rubin J, Brennan T, Rao ST, Sundaralingam M. (1976) Atomic coordinates and molecular conformation of yeast phenylalanyl tRNA. An independent investigation. *Nucleic Acids Res*. **3**, 1111-1123.

T

- Tang, S.-N. and Huang, J.-F. (2005) Evolution of different oligomeric glycyl-tRNA synthetases. *FEBS Lett.*, **579**, 1441-1445.
- Terada T, Nureki O, Ishitani R, Ambrogelly A, Ibba M, Söll D, Yokoyama S. (2002) Functional convergence of two lysyl-tRNA synthetases with unrelated topologies. *Nat Struct Biol*, **9**, 257-262.
- Tsunoda, M., Kusakabe, Y., Tanaka, N., Ohno, S., Nakamura, M., Senda, T., Moriguchi, T., Asai, N., Sekine, M., Yokogawa, T., Nishikawa, K. and Nakamura, K.T. (2007) Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms. *Nucleic Acids Res*, **35**, 4289-4300.
- Tumbula, D.L., Becker, H.D., Chang, W.Z. and Söll, D. (2000) Domain-specific recruitment of amide amino acids for protein synthesis. *Nature*, **407**, 106-110.

U

- Uter NT, Perona JJ. (2006) Active-site assembly in glutaminyl-tRNA synthetase by tRNA-mediated induced fit. *Biochemistry*, **45**, 6858-6865.

V

- Vergara, A., Lorber, B., Zagari, A. and Giegé, R. (2003) Physical aspects of protein crystal growth investigated with the Advanced Protein Crystallization Facility in reduced-gravity environments. *Acta Cryst.*, **D59**, 2-15.

W

- Waller, J.-P., Risler, J.-L., Monteilhet, C. and Zelwer, C. (1971) Crystallization of trypsin-modified methionyl-tRNA synthetase from *Escherichia coli*. *FEBS Lett.*, **16**, 186-188.
- Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I. and Grimes, J.M. (2006) Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*, **14**, 1617-1622.
- Wang CC, Morales AJ, Schimmel P. (2000) Functional redundancy in the nonspecific RNA binding domain of a class I tRNA synthetase. *J.Biol.Chem.*, **275**, 17180-17186.
- Waye, M.M.Y., Winter, G., Wilkinson, A.J. and Fersht, A.R. (1983) Deletion mutagenesis using an 'M13 splint': the N-terminal structural domain of tyrosyl-tRNA synthetase (*B. stearothermophilus*) catalyses the formation of tyrosyl adenylate. *EMBO J.*, **2**, 1827-1829.
- Weik, M. (2003) Low-temperature behavior of water confined by biological macromolecules and its relation to protein dynamics. *Eur. Phys. J. E. Soft Matter*, **12**, 153-158.
- Westhof E, Dumas P, Moras D. (1983) Loop stereochemistry and dynamics in transfer RNA. *J Biomol Struct Dyn.*, **1**, 337-355.
- White, O., Eisen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D., Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L., Moffat, K.S., Qin, H., Jiang, L., Pamphile, W., Crosby, M., Shen, M., Vamathevan, J.J., Lam, P., McDonald, L., Utterback, T., Zalewski, C., Makarova, K.S., Aravind, L., Daly, M.J., Minton, K.W., Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S., Smith, H.O., Venter, J.C. and Fraser, C.M. (1999) Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science*, **286**, 1571-1577.
- Whitmore, L. and Wallace, B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res*, **32**, W668-673.

- Wilcox, M. (1969) Gamma-glutamyl phosphate attached to glutamine-specific tRNA. A precursor of glutamyl-tRNA in *Bacillus subtilis*. *Eur. J. Biochem.*, **11**, 405-412.
- Wilcox, M. and Nirenberg, M. (1968) Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proc. Natl. Acad. Sci. U S A*, **61**, 229-236.
- Wilkinson, A.J., Fersht, A.R., Blow, D.M. and Winter, G. (1983) Site-directed mutagenesis as a probe of enzyme structure and catalysis: Tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. *Biochemistry*, **22**, 3581-3586.
- Winter, G., Fersht, A.R., Wilkinson, A.J., Zoller, M. and Smith, M. (1982) Redesigning enzyme using structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding. *Nature*, **299**, 756-758.
- Woese, C.R., Olsen, G.J., Ibba, M. and Söll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. and Mol. Biol. Reviews*, **64**, 202-236.
- Wolf, Y.I., Aravind, L., Grishin, N.V. and Koonin, E.V. (1999) Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.*, **9**, 689-710.
- Wolfe, C.L., Warrington, J.A., Davis, S., Green, S. and Norcum, M.T. (2003) Isolation and characterization of human nuclear and cytosolic multisynthetase complexes and the intracellular distribution of p43/EMAPII. *Protein Sci.*, **12**, 2282-2290.
- Wong, J.T., (1975) A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci.*, **72**, 1909-12.
- Woo, N.H., Roe, B.A. and Rich, A. (1980) Three-dimensional structure of *Escherichia coli* initiator tRNA^{fMet}. *Nature*, **286**, 346-351.
- Wukovitz, S.W. and Yeates, T.O. (1995) Why protein crystals favour some space-groups over others. *Nature Struct. Biol.*, **2**, 1062-1067.

X

- Xu, M.G., Chen, J.F., Martin, F., Zhao, M.W., Eriani, G. and Wang, E.D. (2002) Leucyl-tRNA synthetase consisting of two subunits from hyperthermophilic bacteria *Aquifex aeolicus*. *J.Biol.Chem.*, **277**, 41590-41596.

Y

- Yang, D. (1996) Mammalian aminoacyl-tRNA synthetases. *Curr.Top.Cell.Regul.*, **34**, 101-136
- Yang, X.L., Skene, R.J., McRee, D.E. and Schimmel, P. (2002) Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15369-15374.
- Yaremchuk, A., Cusack, S. and Tukalo, M. (2000) Crystal structure of a eukaryote/archaeon-like prolyl-tRNA synthetase and its complex with tRNA(Pro)(CGG). *EMBO J.*, **19**, 4745-4758.
- Yaremchuk, A., Krikiviyi, I., Tukalo, M. and Cusack, S. (2002) Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J.*, **21**, 3829-3840.

Z

- Zahradka, K., Slade, D., Bailone, A., Sommer, S., Averbek, D., Petranovic, M., Lindner, A.B. and Radman, M. (2006) Reassembly of shattered chromosomes in *Deinococcus radiodurans*. *Nature*, **443**, 569-573.
- Zhang, C.-M. and Hou, Y.-M. (2005) Domain-Domain Communication for tRNA Aminoacylation: the importance of covalent connectivity. *Biochemistry*, **44**, 7240-7249.

Cristallogénèse et études structurales appliquées aux aminoacyl-ARNt synthétases

Les aminoacyl-ARNt synthétases sont des enzymes modulaires responsables de la fixation spécifique des acides aminés sur l'ARNt. De manière générale, chaque cellule comporte 20 aminoacyl-ARNt synthétases, chacune étant spécifique d'un acide aminé. Ces dernières années, plusieurs exceptions ont été observées, notamment l'absence de la glutaminyl-ARNt synthétase (GlnRS) chez de nombreuses bactéries ou archées. Il existe chez ces organismes une voie indirecte d'aminocyclation impliquant une glutamyl-ARNt synthétase (GluRS) non discriminante couplée à une amidotransférase (AdT).

Mon travail de thèse concerne principalement les aspects de cristallogénèse et l'étude structurale de la glutaminyl-ARNt synthétase de la bactérie *Deinococcus radiodurans* (GlnRS-Dr). Si le cœur de cette enzyme est globalement conservée, certaines de ses insertions et extensions la distinguent de ses homologues bactériennes. La plus notable s'étend sur 220 résidus dans le région C-terminale et présente en partie une homologie de séquence avec la sous unité GatB de l'AdT. Des expériences de biochimie ont montré que la suppression cette région entraîne une perte d'affinité pour l'ARNt.

La structure cristallographique de la GlnRS-Dr résolue au laboratoire a montré de très faibles variations structurales avec les domaines catalytiques et de liaison à l'anticodon de la GlnRS de *E. coli*. Néanmoins, la plupart des insertions et surtout l'extrémité C-terminale ne sont pas visibles dans cette structure qui compte 35% de régions mobiles dans le cristal.

Afin de caractériser ces régions non visibles, un premier volet de mon travail doctoral a consisté à trouver de nouvelles conditions de cristallisation pour générer un empilement cristallin différent ou à les stabiliser en présence de ces différents substrats. Plusieurs conditions ont permis l'obtention de cristaux de la forme native ou complexée (glutamine, ou Gln, analogues de glutaminyl-adénylate, ARNt^{Gln} sous forme de transcrit). Les structures de la GlnRS co-cristallisée avec la glutamine, deux analogues d'adénylate (glutaminol et glutaminyl-sulfamoyl adénylates), et de la GlnRS tronquée de son domaine additionnel C-terminal (GlnRS-ΔC) ont été résolues entre 1,9 à 2,5 Å. Ces résultats ont permis de caractériser les variations structurales du site actif induites par la présence des substrats et de comparer leur mode de fixation avec les données disponibles. L'architecture globale de l'enzyme en présence d'adénylates est proche de l'enzyme libre. La résolution de la structure 3D de la GlnRS cocristallisée avec la Gln et la GlnRS-ΔC a permis de résoudre une longue boucle qui verrouille le site actif. Celle-ci n'existe que dans 2 espèces du genre *Deinococcus*. En revanche, aucune des nouvelles structures ne permet de visualiser l'extrémité C-terminale.

Une autre approche de mon travail concerne l'étude de la géométrie de l'empilement cristallin de la GlnRS-Dr. En présence de petits substrats, on observe, pour un même groupe d'espace, une augmentation jusqu'à 14% du volume de la maille, révélant la plasticité de cet empilement. Il en résulte une diminution notable des contacts intermoléculaires ainsi qu'un élargissement des canaux de solvant. Ces derniers sont suffisamment importants pour accommoder une extrémité C-terminale structurée mais mobile, que ce soit en présence ou en absence de substrat.

Pour compléter cette étude, une analyse structurale en solution par résonance magnétique nucléaire (RMN) a été entreprise sur l'extrémité C-terminale isolée (YqeY-λ) et suggère la présence d'une région majoritairement structurée. Les expériences complémentaires de RMN en deux dimensions sur la partie YqeY-λ marquée en ¹⁵N l'ont confirmé. Des mesures en dichroïsme circulaire nous ont également informés sur la composition en structures secondaires de ce module.

Enfin, en parallèle de mon projet doctoral principal, j'ai mis à profit les compétences acquises en cristallographie en participant à l'étude structurale d'une autre synthétase. Ce travail a conduit à la détermination de la première structure tridimensionnelle d'une synthétase humaine d'origine mitochondriale : la Tyrosyl-ARNt synthétase (mt-TyrRS). La structure correspond à un complexe d'analogue de substrat qui ressemble globalement à celle des autres TyrRS mais montre plusieurs caractéristiques mitochondriales qui les en distingue.