



HAL
open science

Études structurales de la maturation des ARN de transfert : les méthyltransférases M1A22 et M1A58 des ARNt

Clément Dégut

► **To cite this version:**

Clément Dégut. Études structurales de la maturation des ARN de transfert : les méthyltransférases M1A22 et M1A58 des ARNt. Biologie structurale [q-bio.BM]. Université Sorbonne Paris Cité, 2015. Français. NNT : 2015USPCB226 . tel-03617253

HAL Id: tel-03617253

<https://theses.hal.science/tel-03617253v1>

Submitted on 23 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université Paris Descartes

Faculté de Pharmacie

Ecole Doctorale

Médicament - Toxicologie - Chimie - Imagerie

ETUDES STRUCTURALES DE LA MATURATION DES ARN DE TRANSFERT :

LES METHYLTRANSFERASES M^1A_{22} ET M^1A_{58} DES ARNt

Thèse de Doctorat

Présentée le 5 novembre 2015 par

Clément Dégut

Pour l'obtention du grade de

Docteur de l'Université Paris Descartes

Spécialité : Biologie structurale

DIRECTRICE DE THESE : **Carine Tisné**

COMPOSITION DU JURY :

Emmanuelle Schmitt (Rapporteur)

Jean-Pierre Simorre (Rapporteur)

Mélanie Éthève-Quelquejeu (Examineur)

Valérie Heurgué-Hamard (Examineur)

Préparée au Laboratoire de Cristallographie et RMN Biologiques

Université Paris Descartes, CNRS UMR 8015
4, avenue de l'Observatoire, 75006 Paris

Remerciements

Tout d'abord, je tiens à remercier les personnes qui ont accepté d'évaluer mon travail : Emmanuelle Schmitt et Jean- Pierre Simorre en tant que rapporteurs, Mélanie Ethève-Quelquejeu et Valérie Heurgué-Hamard en tant que membres du jury.

Je tiens également à remercier toutes les personnes qui m'ont aidé et soutenu pendant ces trois années de thèse.

J'adresse tout particulièrement mes remerciements à Carine Tisné, ma directrice de thèse, qui m'a laissé faire toutes les erreurs que je voulais faire sans jamais m'abandonner pour autant. Avec qui les discussions, qu'elles soient scientifiques ou non, sont toujours agréables et pleine de bonne humeur. Si l'expérience que je retire de mes trois années de thèse est extrêmement positive, c'est en grande partie de son fait. Merci aussi pour la relecture attentive de cette thèse.

Merci également à Pierre Barraud qui m'a formé à la RMN, et a toujours été disponible pour répondre à mes nombreuses questions et qui a également consacré un temps certain à la relecture de ce document.

Je tiens à remercier chaleureusement l'équipe RMN avec qui j'ai travaillé sur différents projets : Franck pour les nombreuses gouttes... (de robot), Marjorie (à qui je dois sûrement encore du chocolat), Irène, Assia, Luc, Valéry et Sylvie.

Je remercie également Thibaut Crépin qui m'a accueilli dans son laboratoire.

Merci à Carlos Kikkuti pour son aide et les discussions autour du Nanotemper.

Merci à Nelly Morellet pour ses conseils et son aide précieuse au TGIR-RMN.

Merci à toute l'équipe du laboratoire LCRB pour les bons moments que nous avons partagés et à toutes les machines à café qui ont succombées pendant cette période.

Merci enfin à ma famille et à mon entourage, qui m'a supporté et encouragé pendant ces trois ans. Un grand merci à Camille qui m'a également supporté... dans tous les sens du terme en particulier pendant la rédaction de cette thèse.

Résumé de la thèse

Les ARNt portent de nombreuses modifications post-transcriptionnelles des nucléosides. Ces réactions chimiques de maturation peuvent se produire sur les bases et le sucre de l'ARNt. La famille des ARNt est la famille d'ARN qui compte le plus grand nombre de nucléotides modifiés avec 92 modifications connues (voir la base de données des modifications des ARNt : <http://biochem.ncsu.edu/RNAmods>). Les fonctions de ces modifications commencent peu à peu à être élucidées. D'une manière générale, les modifications des nucléotides de la boucle anticodon sont nécessaires pour un appariement correct codon-anticodon et une reconnaissance par l'aminoacyl-ARNt-synthétase. Les modifications au niveau des boucles T et D facilitent le repliement et la stabilisation de la structure en L des ARNt. Pour les modifications localisées en dehors de ces boucles, elles peuvent affecter la stabilité de sa structure, l'identité de l'ARNt ou le mener vers des structures alternatives non fonctionnelles. Ces modifications sont donc essentielles pour le maintien de la fonction cellulaire normale, et des altérations de ces modifications conduisent inévitablement à des pathologies plus ou moins marquées chez l'Homme.

Parmi ces modifications, les méthylations sont les plus courantes. Elles sont assurées par des méthyltransférases qui utilisent le donneur universel de méthyle la SAM (S-adenosyl-L-méthionine) pour transférer son groupement méthyle au nucléoside cible.

Cette thèse porte sur l'étude structurale par RMN et cristallographie de deux de ces méthyltransférases des ARNt :

- TrmK la méthyltransférase m¹A de la position 22
- TrmI la méthyltransférase m¹A de la position 58 bactérienne.

Les modes de reconnaissance entre ces méthyltransférases et leurs substrats ARNt sont étudiés.

Mots clé : ARN de transfert, méthyltransférase, m¹A, TrmK, TrmI

Table des matières

CHAPITRE I INTRODUCTION	1
I. LES ARN DE TRANSFERT	1
<i>I.1 Structure des ARNt.....</i>	1
<i>I.1.1 Structure primaire.....</i>	<i>1</i>
<i>I.1.2 Structure secondaire.....</i>	<i>3</i>
<i>I.1.3 Structure tertiaire.....</i>	<i>5</i>
<i>I.1.1 Plasticité structurale.....</i>	<i>7</i>
<i>I.2 Fonctions des ARN de transfert.....</i>	12
<i>I.2.1 Fonction dans la traduction.....</i>	<i>12</i>
<i>Les ARNt dans la traduction.....</i>	<i>12</i>
<i>Le code génétique.....</i>	<i>14</i>
<i>Biais d'usage des codons</i>	<i>15</i>
<i>I.2.2 Activités catalytiques des ARNt</i>	<i>17</i>
<i>Rôle dans le transfert de peptide au niveau du ribosome</i>	<i>17</i>
<i>Rôle dans leur propre acylation/déacylation.....</i>	<i>18</i>
<i>I.2.3 Donneur d'acide aminé</i>	<i>19</i>
<i>Le peptidoglycane bactérien.....</i>	<i>19</i>
<i>Dégradation des protéines : N-end rule</i>	<i>21</i>
<i>Modification de la membrane bactérienne.....</i>	<i>22</i>
<i>Synthèse de composés organiques.....</i>	<i>23</i>
<i>I.2.4 Implications des ARNt dans les infections virales.....</i>	<i>24</i>
<i>I.2.5 Fragments d'ARNt.....</i>	<i>25</i>
<i>Demi-ARNt.....</i>	<i>25</i>
<i>Fragments 5' d'ARNt (5'tRF)</i>	<i>27</i>
<i>Fragments 3'CCA</i>	<i>27</i>
<i>Fragment 3'U.....</i>	<i>27</i>
<i>Implication dans des pathologies humaines</i>	<i>28</i>
<i>I.3 Maturations des ARN de transfert.....</i>	30
<i>I.3.1 Maturation de l'extrémité 5'</i>	<i>31</i>
<i>I.3.2 Maturation de l'extrémité 3'</i>	<i>32</i>

I.3.3	<i>Épissage</i>	33
I.3.4	<i>Aminoacylation</i>	33
I.4	<i>Modifications post-transcriptionnelles</i>	35
I.4.1	<i>Nature chimique des modifications</i>	35
I.4.2	<i>Diversité et répartition des nucléotides modifiés</i>	36
I.4.3	<i>Nucléotides modifiés des ARN de transfert</i>	39
	<i>La position 34 ou wobble</i>	40
	<i>La position purine 37</i>	41
	<i>Structuration et stabilisation des ARN de transfert</i>	42
I.4.4	<i>Fonctions biologiques des modifications post-transcriptionnelles</i>	45
I.4.5	<i>Dynamique des modifications</i>	46
	<i>Adaptation aux conditions environnementales</i>	47
	<i>Adaptation aux conditions cellulaires</i>	48
I.4.1	<i>Modifications post-transcriptionnelles des ARNt et pathologies humaines</i>	50
	<i>Désordre neurologique</i>	50
	<i>Cancer</i>	51
	<i>Maladies métaboliques</i>	51
II.	<i>LA M¹A DANS LES ARN DE TRANSFERT</i>	53
II.1	<i>Les méthyltransférases</i>	53
II.1.1	<i>Donneurs de méthyle</i>	53
II.1.2	<i>Les Méthyltransférases SAM-dépendantes</i>	56
	<i>Méthyltransférases de classe I</i>	57
	<i>Les autres méthyltransférases SAM-dépendantes</i>	59
	<i>Méthyltransférases à radical SAM</i>	61
II.2	<i>La méthylation des adénines en sur l'azote 1 : m¹A</i>	63
II.2.1	<i>Les méthylations dans les ARN de transfert</i>	63
II.2.2	<i>Importance de la modification m¹A</i>	64
	<i>Position 9</i>	65
	<i>Position 14</i>	66
	<i>Position 57</i>	66
	<i>Positions 22 et 58</i>	67
II.2.1	<i>Position 22 : TrmK</i>	67
	<i>Fonction</i>	68
	<i>Structure de TrmK</i>	69
	<i>Substrats</i>	70
II.2.2	<i>Position 58</i>	71
	<i>Fonctions</i>	73
	<i>Structure de TrmI</i>	73
	<i>Substrats</i>	76

CHAPITRE II ÉTUDE DE LA METHYLTRANSFERASE M¹A₅₈ DES

ARNT : TRMI	79
III. MUTANTS DU SITE ACTIF DE TRMI	81
III.1 Problématique et résumé de l'article 1.....	81
III.1 Article 1 : Dégut et al. (2015) <i>Biophys. Chem.</i>	82
IV. COMPLEXE COVALENT TRMI/BRAS T DE L'ARNT.....	90
IV.1 Introduction	90
IV.2 Mutations ponctuelles en cystéines.....	91
IV.2.1 Sélection des mutants cystéines de <i>TrmI</i>	91
IV.2.2 Production et purification des différents mutants de <i>TrmI</i>	92
IV.3 Formation du complexe covalent <i>TrmI</i> /BrasT	94
IV.3.1 Tests de formation des complexes covalents en microvolume.....	94
IV.3.2 Essais de cristallogenèse du complexe covalent <i>TrmI</i> Q200C/ARN G57G*	96
IV.4 Perspectives.....	100

CHAPITRE III ÉTUDE DE LA METHYLTRANSFERASE M¹A₂₂ DES

ARNT : TRMK.....	101
V. PRODUCTION D'UN SUBSTRAT DE TRMK : L'ARN ^{T^{SER}} DE <i>B. SUBTILIS</i>	103
V.1 Problématique et résumé de l'article 2.....	103
V.2 Article 2 : Dégut et al. (2016)	105
V.3 Cristaux de l'ARN ^{T^{SER}} produit <i>in vivo</i>	127
VI. ATTRIBUTION DU SQUELETTE PEPTIDIQUE DE TRMK	128
VI.1 Etude de <i>TrmK</i> par RMN.....	128
VI.2 Principe de l'attribution séquentielle.....	129
VI.3 Article 3 : Dégut et al. (2015)	135
VII. ETUDES DE L'INTERACTION DE LA METHYLTRANSFERASE TRMK AVEC SON SUBSTRAT ARNT 144	
VII.1 Problématique et résumé de l'article 4.....	144
VII.2 Article 4 Dégut et al (2015).....	145
VII.3 Etude en cours et perspectives sur <i>TrmK</i>	175

CONCLUSION.....	178
------------------------	------------

BIBLIOGRAPHIE	184
----------------------------	------------

Chapitre I Introduction

I. Les ARN de transfert

Les acides ribonucléiques de transfert (ARNt) sont les adaptateurs du code génétique. Ils font le lien physique entre la séquence d'acide nucléique d'un ARN messenger (ARNm) et la séquence d'acides aminés d'une protéine. Ils s'apparient grâce à leur anticodon correspondant au codon du code génétique, à l'ARN messenger apportant l'acide aminé adéquat au ribosome pour la synthèse peptidique. Ils ont été mis en évidence pour la première fois en 1958 par Mahlon Hoagland qui a montré l'existence dans la cellule d'ARN solubles portant des acides aminés, et a prouvé leur implication dans la synthèse peptidique [1].

I.1 Structure des ARNt

I.1.1 Structure primaire

Les ARN sont des polymères organiques formés par la succession de briques élémentaires : les nucléotides composés d'une base azotée, d'un sucre (le ribose) et d'un groupement phosphate (**Figure I-1 A**). Les nucléotides sont reliés entre eux par une liaison phosphodiester impliquant l'oxygène 3' du ribose d'un nucléotide et le phosphate en 5' du nucléotide précédant dans la séquence. La polarité des ARN est définie par la nomenclature des sucres, on parlera ainsi d'extrémité 3' ou 5' de l'ARN. Le carbone anomérique (en 1') porte la base azotée qui donne au nucléotide ses caractéristiques. Il existe pour les ARN quatre bases canoniques : deux purines (la guanine et l'adénine) et deux pyrimidines (l'uracile, et la cytosine) (**Figure I-1**

B). Nous verrons au cours de cette introduction que ces nucléotides peuvent être très modifiés chimiquement par des enzymes dites de modification.

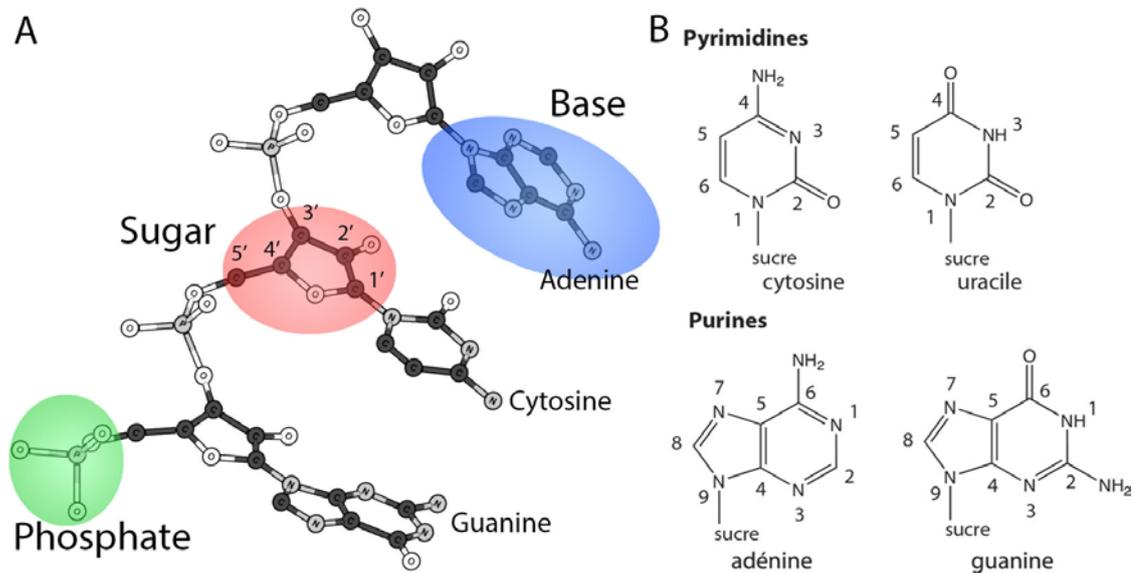


Figure I-1 Organisation des ARN A) Structure générale, le groupement phosphate est sur fond vert, le sucre sur fond rouge et la base sur fond bleu. B) Les quatre bases canoniques de l'ARN.

En 1968, le Prix Nobel de Médecine ou Physiologie a été accordé à Robert Holley (conjointement avec Har Gobind Khorana et Marshall Nirenberg) « pour leur interprétation du code génétique et de ses fonctions dans la synthèse protéique. » C'est en effet dans son équipe qu'a été séquencé le premier ARNt [2]. Les ARNt ont des tailles variables de 74 à 98 nucléotides avec une moyenne de 76. Leurs séquences primaires sont codées par des gènes spécifiques, et ils sont synthétisés par l'ARN polymérase III chez les eucaryotes (il n'y a qu'une polymérase chez les procaryotes). Un important travail d'alignement de séquence d'ARNt a été entrepris dans les années 80 par M. Sprinzl et continué jusqu'à aujourd'hui [3-5], donnant lieu à une nomenclature systématique. En effet, de nombreuses positions invariantes (U₈, A₁₄, G₁₈, G₁₉, A₂₁, U₃₃, G₅₃, T₅₄, Ψ₅₅, C₅₆, A₅₈, C₆₁, C₇₄, C₇₅ et A₇₆) et semi-invariantes (Y₁₁,

R₁₅, R₂₄, Y₃₂, R₃₇, Y₄₈, R₅₇ et Y₆₀)* ont été décrites [6]. Dans ce contexte, les ARNt sont conventionnellement numérotés du nucléotide 1 pour l'extrémité 5' (il existe parfois un nucléotide -1), au nucléotide 76 en 3', et ce quel que soit le nombre réel de nucléotides. Les trois nucléotides de l'anticodon sont ainsi toujours les nucléotides 34, 35 et 36. L'extrémité 3' est caractérisée par la présence du triplet de nucléotides CCA, l'adénine terminale étant la base qui sera aminoacylée sur le 2'OH ou 3'-OH selon la classe de l'aminoacyl-ARNt synthétase (aaRS).

I.1.2 Structure secondaire

Certains groupes de nucléotides de la chaîne de l'ARNt sont capables de s'apparier entre eux, *via* des liaisons hydrogènes. Ces appariements sont de type Watson-Crick : paire de base G=C et A-U, et de type wobble le plus fréquemment la paire G○U† (**Figure I-2 B**). Ces appariements provoquent la formation de quatre bras qui donnent aux ARNt leur structure secondaire dite en feuille de trèfle (**Figure I-2 B**). Ces bras sont :

- Le bras accepteur qui relie les extrémités 5' et 3', laissant libre une à deux bases suivies de l'extrémité CCA
- Le bras D, qui doit son nom à la présence très fréquente d'une ou plusieurs dihydrouridine (noté D)
- Le bras anticodon portant au niveau de la boucle les trois nucléotides de l'anticodon aux positions 34, 35 et 36

* Les pyrimidines sont notées Y et les purines R.

† Dans ce manuscrit les appariements seront notés de façon simplifiée : les paires de bases Watson-Crick seront notées avec deux traits pour G=C et un seul pour A-U, les appariements de type wobble avec un rond (par exemple G○U) et les autres appariements non canoniques avec un rond plein (par exemple m⁷G₄₆●C₂).

- Le bras T, caractérisé par la présence fréquente de ribothymidine (m^5U)*. Il est aussi parfois appelé « T Ψ C », du fait de la présence de la séquence conservée thymine, pseudouridine (noté Ψ), cytosine au niveau de la boucle aux positions 54, 55 et 56.

Entre le bras anticodon et le bras T se trouve une région variable qui porte le plus souvent trois à quatre nucléotides, mais peut en comporter jusqu'à une vingtaine et former une tige-boucle supplémentaire. On parle d'ARNt de classe I pour les ARNt contenant une région variable courte, et de classe II pour ceux ayant une région variable longue.

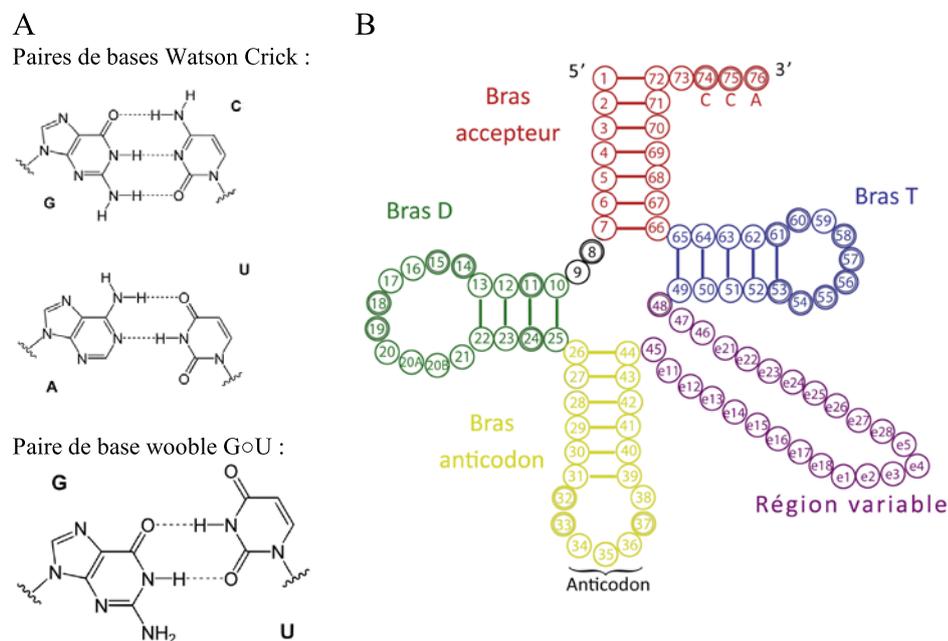


Figure 1-2 Structure secondaire des ARNt. A) Les deux paires de bases Watson-Crick des ARNt, et la paire de bases B) Schéma général de la structure des ARNt. Les bases sont représentées par des cercles contenant leur numérotation conventionnelle, les traits entre bases indiquent les appariements, les positions invariantes sont marquées

* La notation des bases modifiées est la suivante : n^iX_km avec : **n** étant l'indication de la modification ; **i** sa position sur la base modifiée ; **X** le nucléotide en question ; **k** le numéro du nucléotide dans la séquence; et **m** la modification sur le 2'OH du ribose.

d'un double cercle. Les ARNt ayant une région variable longue (présence de nucléotide "e") sont dit de classe II, les régions variables courtes définissent la classe I.

I.1.3 Structure tertiaire

La première structure cristallographique d'un ARNt (ARNt^{Phe} de levure) a été résolue en 1974 par Sung-Hou Kim [7, 8], la même équipe ayant publié précédemment des données à basse résolution, qui permettait déjà d'avoir une idée de l'organisation générale de l'ARNt [9, 10]. De nombreuses structures d'ARNt seuls ou en complexe ont depuis été résolues. Ces structures ont montré que le repliement tridimensionnel des ARNt est extrêmement bien conservé. Ils adoptent une structure tridimensionnelle caractéristique dite en forme de « L » (**Figure I-3**). Cette structure résulte de l'empilement d'une part du bras accepteur et du bras T, et d'autre part du bras D et du bras anticodon, qui forment deux hélices perpendiculaires. Cette structure est maintenue par des interactions de type liaisons hydrogène entre les nucléotides des boucles T et D. On trouve en particulier la paire de bases G₁₉=C₅₆ qui forme « l'angle » de l'ARNt et est présente dans la majorité des séquences connues [11].

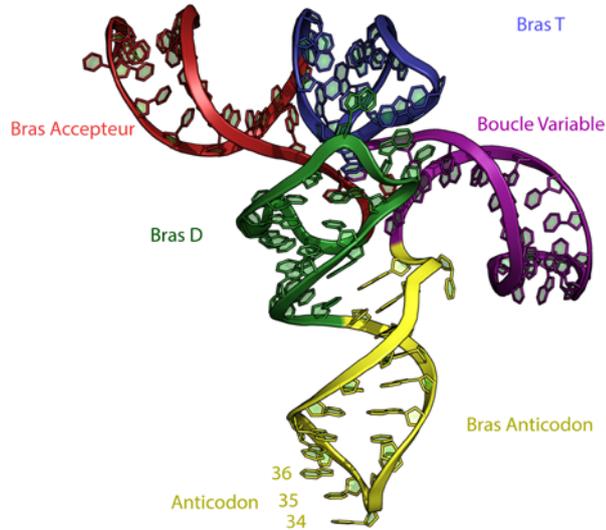


Figure I-3 Structure cristallographique d'un ARNt de classe II : l'ARNt^{Sec} humain [12] (code d'accèsion PDB : 3HL2), l'empilement des bras T et accepteur ainsi que des bras D et anticodon formes deux hélices perpendiculaires. La boucle variable forme une hélice supplémentaire sans interaction avec les autres bras de l'ARNt.

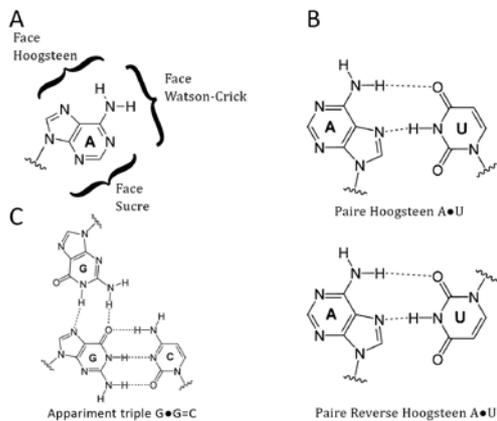


Figure I-4 Exemples d'appariement non WC. A) Trois faces d'une base peut former des appariements avec d'autres bases. B) Appariement Hoogsteen et Reverse-Hoogsteen pour la paire A•U, la paire de bases Reverse Hoogsteen est caractérisée par la position en trans des liaisons glycosidiques. C) Exemple d'appariement triple ici G•G=C.

Ces interactions ne se limitent pas aux appariements Watson-Crick classiques, de nombreuses autres types d'appariements dit non canoniques sont possibles (**Figure I-4**). On pourra citer la paire G₁₅•C₄₈; la paire G₁₈•Ψ₅₅ qui ne comporte qu'une liaisons hydrogène ; ou encore deux paires Reverse-Hoogsteen très conservées :

T₅₄•m¹A₅₈ et U₈•A₁₄. On trouve également les appariements triples U₈•A₁₄•A₂₁ (présent dans 82% des ARNt) G₄₅•m²G₁₀=C₂₅, m⁷G₄₆•G₂₂=C₁₃ et A₉•A₂₃•U₁₂.

Une structure plus récente de l'historique ARNt^{Phe} de levure [13] résolue à 1.93 Å a montré la présence de 11 sites de fixations spécifiques pour des métaux, et notamment pour des métaux divalents. Il est connu, de longue date, que les ARN structurés ont besoin de cation et en particulier de sodium et de magnésium pour stabiliser leur structure. Les cations permettent d'écranter les charges négatives du squelette phosphodiester, et ainsi stabilisent la compaction de la structure globale de l'ARN.

I.1.1 Plasticité structurale

La conservation de cette structure globale en "L" s'explique par le fait que les différents ARNt de la cellule interagissent avec les mêmes partenaires, comme le ribosome ou les facteurs de traduction. Les petites variations locales, de structure ou de séquence, expliquent les interactions spécifiques que chaque ARNt peut entretenir avec certains partenaires cellulaires, notamment son aaRS. Certains ARNt vont avoir des flexibilités plus importantes ou pourront atteindre des conformations locales différentes des autres qui permettent à leurs partenaires de les discriminer. Cette plasticité structurale joue un rôle important dans les fonctions biologiques des ARNt. Quelques exemples sont donnés dans les paragraphes suivants.

La **Figure I-5** illustre les importantes déformations que peuvent subir les ARNt au cours de leurs interactions avec leurs partenaires cellulaires.

La plasticité structurale est révélée notamment par le changement dans l'orientation relative des deux hélices de l'ARNt, l'angle variant de 65° (**Figure I-5 d**) à 110° (**Figure I-5 b**). Quelques distances caractéristiques entre les phosphates des bases sont indiquées (en ångström). La distance D5 correspondant à l'écart entre l'anticodon et la dernière base avant le CCA. C'est celle qui varie le plus : 38%

entre l'ARNt^{Asp} libre et en complexe avec son aaRS [14] (**Figure I-5 b et d**). De la même façon dans les différents sites du ribosome, les ARNt adoptent des conformations sensiblement différentes pour s'adapter aux différents sites du ribosome (**Figure I-5 e et f**).

Enfin, un phénomène remarquable est celui de l'ajustement induit à longue distance que l'on peut observer sur ces ARNt. En effet, dans la structure de l'ARNt^{Phe} lié au facteur d'élongation EF-Tu (**Figure I-5 c**), la boucle anticodon est dans une conformation très différente de celle de la structure libre (**Figure I-5 a**), alors que EF-Tu se lie sur la partie opposée de l'ARNt.

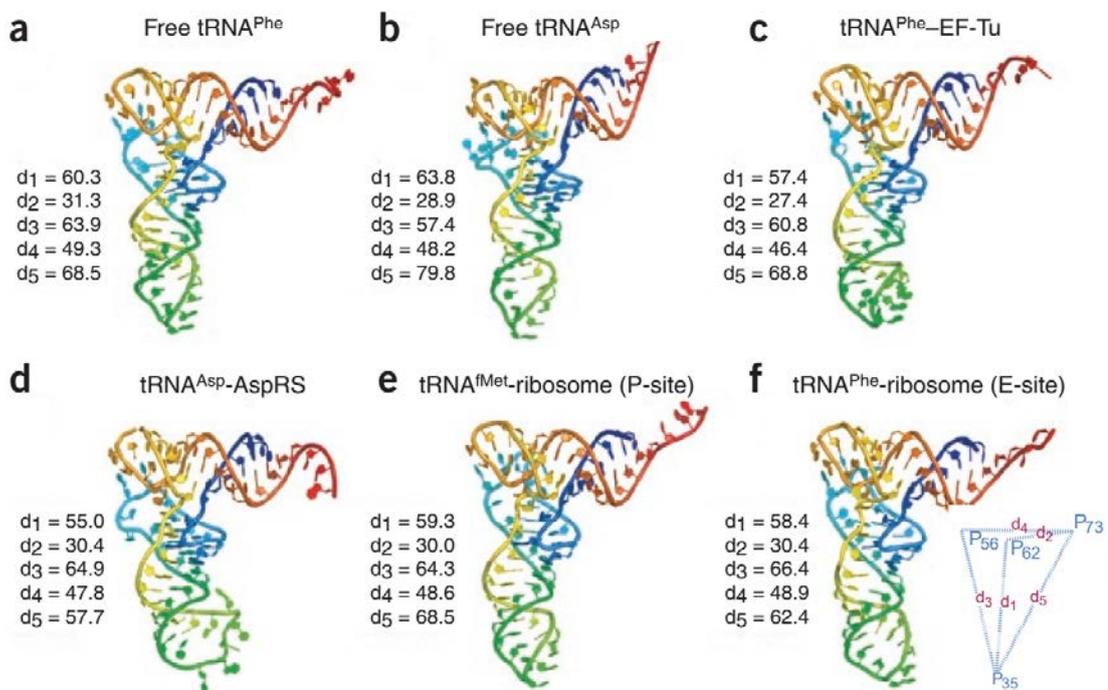


Figure I-5 Plasticité structurale des ARNt [14]. Six structures d'ARNt sont montrées sous la même orientation, représentant différents états conformationnels. (a) ARNt^{Phe} de levure libre. (b) ARNt^{Asp} de levure mimant l'interaction avec un ARNm, l'anticodon GUC étant en effet relié à un autre anticodon par les contacts cristallins. (c) ARNt^{Phe} de levure en complexe avec le facteur d'élongation EF-Tu. (d) ARNt^{Asp} de levure en complexe avec son aminoacyl-ARNt synthétase. (e) ARNt^{Met} de levure dans le site P du

ribosome. (f) ARNt^{Phe} de levure dans le site E du ribosome.) (Codes d'accèsion PDB dans l'ordre 4TRA, 2TRA, 2TTT, 1ASY, 2J00, 2J00).

La structure de l'ARNt^{Met} de *E. coli* [15] illustre bien comment une conformation différente de l'anticodon que cet ARNt est probablement le seul à pouvoir atteindre peut jouer un rôle biologique. La région de l'anticodon montre une structure unique impliquant deux des trois paires de bases G=C de la tige anticodon caractéristiques des ARNt initiateurs (**Figure I-6 a**). Cette structure unique peut être décrite par trois points :

- Une paire de bases Cm₃₂•A₃₈ de type wobble qui étend la tige anticodon d'une paire supplémentaire (**Figure I-6 c**).
- un triplet de bases entre la base A₃₇ située après le triplet de l'anticodon et la paire G₂₉=C₄₁ du côté grand sillon (**Figure I-6 d**).
- Un empilement particulier des bases de la boucle anticodon, où sur la base A₃₈ est directement empilée sur la dernière base de l'anticodon U₃₆, alors qu'elle est empilée sur la base 37 dans toutes les structures d'ARNt élongateurs (**Figure I-6 b**). Cette structure permet probablement sa sélection au site P du ribosome en tant qu'ARNt initiateur et permet de le distinguer des ARNt élongateurs.

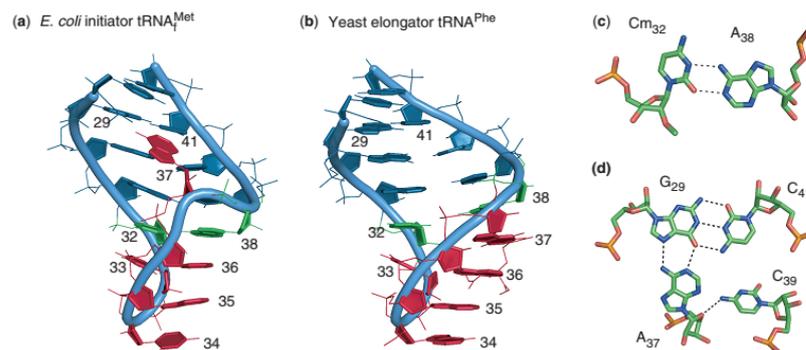


Figure I-6 Conformation unique du bras anticodon de l'ARNt^{Met} d'*E. coli* [15]. (a) Structure de la boucle anticodon de l'ARNt^{Met}, la base A₃₈ est empilée sur l'anticodon, la base A₃₇ interagit avec plusieurs bases du bras anticodon (b) la

structure de l'ARNt^{Phe} élongateur est donnée pour comparaison. (c et d) La structure de cette boucle est expliquée notamment par la paire de bases Cm₃₂○A₃₈ (en vert sur les structures a et b) et l'appariement triple A₃₇●G₂₉=C₄₁.

Les deux exemples suivant montrent que des ARNt peuvent aussi avoir des structures tridimensionnelles différentes de cette forme en "L".

Les ARNt mitochondriaux peuvent présenter une diversité structurale très forte avec les ARNt cellulaires. En effet, on retrouve plusieurs ARNt auxquels il manque les bras T ou D, voire les deux [16, 17]. Ces ARNt sont pourtant aminoacylés normalement et peuvent prendre part à la synthèse protéique dans le ribosome mitochondrial [18].

La structure la plus extrême de ces déformations induites par les interactions avec des partenaires est très certainement la conformation "λ" des ARNt en interaction avec certaines enzymes de modification. Cette conformation a été découverte avec la structure de l'enzyme de modifications ARNt-guanine transglycosylase (arcTGT) en complexe avec son substrat ARNt [19]. La modification a lieu sur la guanine 15 qui est enfouie au cœur de la structure tertiaire de l'ARNt sous forme de L. Pour pouvoir atteindre cette position, l'enzyme doit déformer considérablement la structure de l'ARNt, qui adopte alors une structure dite "λ". Dans cette structure, la région D ne forme plus de tige-boucle, mais est complètement ouverte. Toutes les interactions tertiaires entre boucles D et T sont détruites, la région variable forme une hélice avec une partie du bras D (hélice DV) et est empilé sur la tige anticodon (**Figure I-7**).

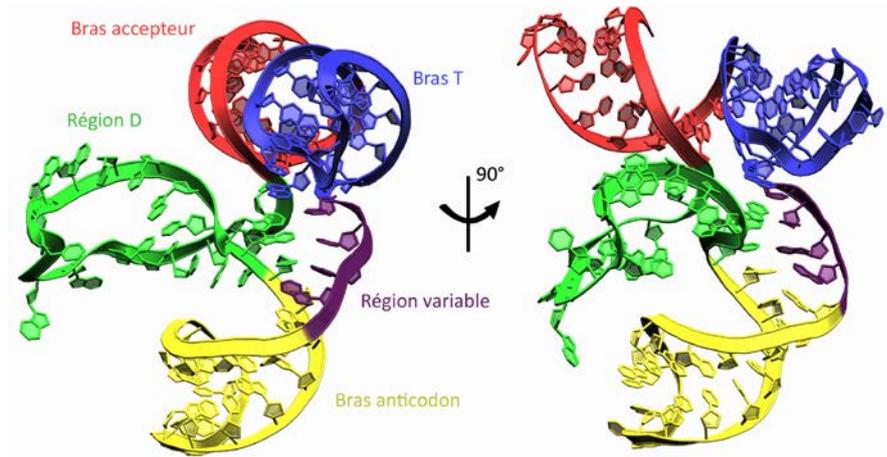


Figure I-7 Structure de l'ARNt sous forme λ en complexe avec l'enzyme de modification arcTGT (non montrée). Le code couleur pour les bras est le même que précédemment. Dans cette forme, le bras D est déformé en une large boucle et ne fait plus d'interaction avec la boucle T (code d'accèsion PDB : 1J2B) [19].

Les ARNt ont tous la même conformation globale, mais leur structure est très dynamique, ils sont constamment déformés par les différentes interactions avec leurs partenaires cellulaires. Ils portent, en outre, des particularités structurales qui leur permettent de s'adapter à leur différentes fonctions.

I.2 Fonctions des ARN de transfert

La fonction la plus connue des ARNt est leur rôle dans la traduction, mais ils ne se restreignent pas à cette fonction et ils jouent de nombreux rôles dans les activités cellulaires. Les différentes fonctions des ARNt sont décrites dans les paragraphes suivants.

I.2.1 Fonction dans la traduction

Les ARNt dans la traduction

La traduction est le processus par lequel un ARN messager (ARNm) est décodé par le ribosome, un polypeptide est alors synthétisé d'après la séquence de l'ARNm. Les ARNt sont centraux dans ce processus, ils font le lien entre le code génétique et la protéine en devenir. Le ribosome possède trois sites de liaison aux ARNt, les sites A (Aminoacyl), P (Peptidyl) et E (Exit) (**Figure I-8**).

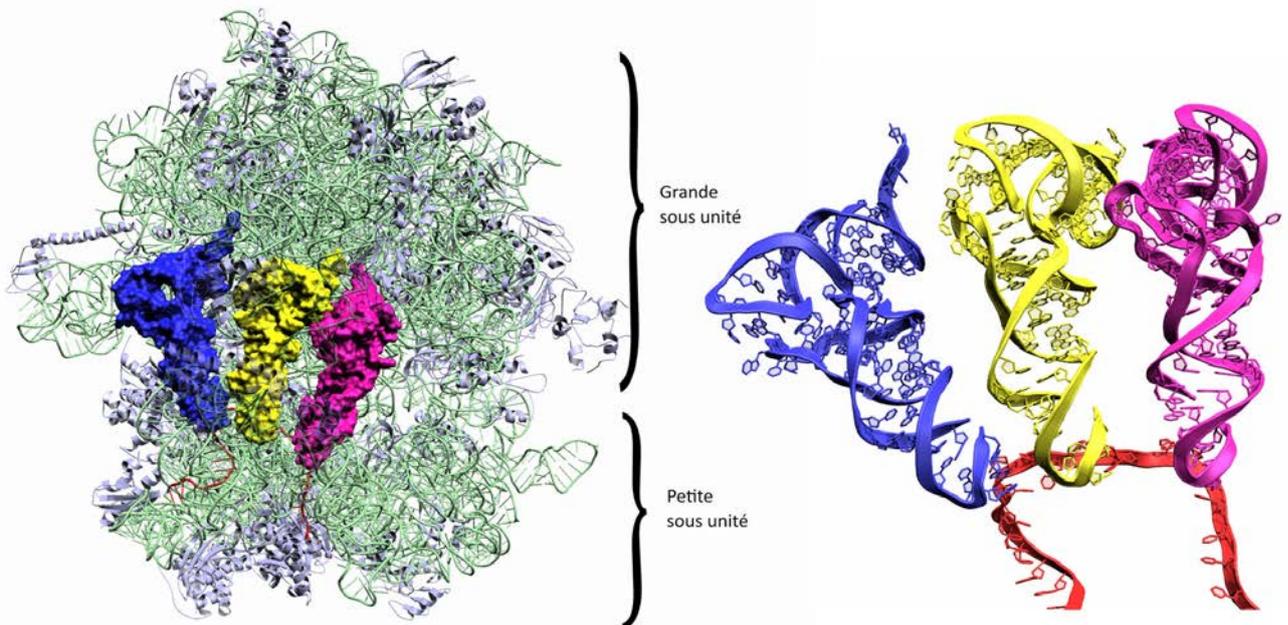


Figure I-8 Structure du Ribosome 70S de *T. thermophilus* les protéines sont en bleu clair et les ARN ribosomiques (ARNr) en vert. Un ARNt^{phe} est présent dans chaque site

P (en jaune), E (en bleu) et A (en violet) l'ARN messenger est en rouge. (code d'accèsion PDB : 4V51) [20].

L'initiation de la traduction est la première étape de la traduction, et elle se fait sur un codon spécifique, dit codon d'initiation, dans le site P du ribosome. AUG est de loin le codon d'initiation le plus fréquent (83% des gènes chez *E. coli* commencent par ce codon), puis GUG (14%), et UUG (3%). Ces codons conduisent tous à l'incorporation du premier résidu N-terminal (la formyl-méthionine chez les procaryotes et la méthionine chez les eucaryotes) par l'action d'un ARNt initiateur (ARNt_i) spécifique.

L'élongation implique les ARNt élongateurs porteurs des acides aminés qui sont intégrés à la protéine. Les ARNt matures et acylés, donc prêts pour rentrer dans la traduction sont liés par le facteur d'élongation Tu (EF-Tu) qui permet la liaison au ribosome dans le site A (**Figure I-9 A - B**). La liaison peptidique est alors formée entre le polypeptide jusqu'alors attaché à l'ARNt du site P, et l'acide aminé porté par l'ARNt nouvellement arrivé au site A. La protéine en formation est donc transférée du site P vers le site A en gagnant un acide aminé. L'ARNt du site E sort du ribosome en même temps (**Figure I-9 C - D**).

La translocation a alors lieu sous l'action de la translocase EF-G chez les procaryotes ou eEF-2 chez les eucaryotes. L'ARNt porteur du polypeptide est déplacé du site A vers le site P et celui du site A vers le site E : le ribosome se déplace alors d'un codon (**Figure I-9 E - F**).

Il est également de plus en plus proposé que les ARNt jouent un rôle important dans la catalyse de la formation de la liaison peptidique, ce point sera discuté dans le paragraphe I.2.2 en page 17.

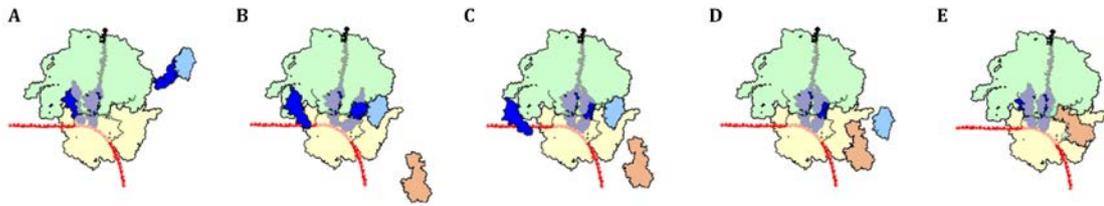


Figure I-9 Etapes de l'élongation. Les petites et grandes sous-unité du ribosomes sont représentées respectivement en jaune et vert, les ARNt sont bleu foncé, EF-Tu est en bleu clair et EF-G est orange.

Le code génétique

Le code génétique utilise des codons de trois bases pour encoder le type d'acides aminés, il existe donc 64 (4^3) codons possibles pour seulement 20 acides aminés – 61 en réalité puisqu'il existe 3 codons-stops (**Tableau I-1**). Le code génétique est donc dit dégénéré, plusieurs codons synonymes codent pour un même acide aminé. Chaque organisme est capable de lire tous les codons du code génétique universel. Ainsi, plusieurs ARNt avec des anticodons différents peuvent porter le même acide aminé. Il n'existe néanmoins pas un ARNt par anticodon la plupart des organismes en possédant une quarantaine seulement [21, 22]. Certains ARNt peuvent lire plusieurs codons différents. L'appariement codon/anticodon repose sur trois paires de bases. Les deux premières sont des appariements Watson-Crick stricts et ne tolèrent aucun mésappariement. La troisième paire de bases en revanche repose sur un appariement moins strict, dit appariement wobble [23–25]. En position wobble plusieurs paires de base différentes peuvent être possibles, une uridine wobble est par exemple capable de faire un appariement avec n'importe laquelle des quatre bases A,G,C,T. Les codons d'un même acide aminé ayant souvent les deux premières bases identiques, l'appariement wobble sur la dernière paire permet d'économiser le nombre d'ARNt. Par exemple tous les codons glycine commencent par les deux bases G_1G_2 , ces deux base pourront être suivies par n'importe laquelle

des quatre autres bases (**Tableau I-1**), et pourront tous être lu par le même anticodon UCC.

		2 ^e Base									
		U		C		A		G			
1 ^{re} Base	U	UUU	F Phe	UCU	S Ser	UAU	Y Tyr	UGU	C Cys	U	
		UUC		UCC		UAC		UGC		C	
		UUA	L Leu	UCA		UAA	STOP Ocre	UGA	STOP Opale / U Sec / W Trp	A	
	UUG	L Leu / START	UCG	UAG		STOP Ambr e / O Pyl	UGG	W Trp	G		
	C	CUU	L Leu	CCU		P Pro	CAU	H His	CGU	R Arg	U
		CUC		CCC			CAC		CGC		C
		CUA		CCA	CAA		CGA	A			
		CUG		CCG	CAG		CGG	G			
	A	AUU	I Ile	ACU	T Thr	AAU	N Asn	AGU	S Ser	U	
		AUC		ACC		AAC		AGC		C	
		AUA		ACA		AAA	AGA	A			
		AUG	M Met / START	ACG		AAG	K Lys	AGG	R Arg	G	
	G	GUU	V Val	GCU	A Ala	GAU	D Asp	GGU	G Gly	U	
		GUC		GCC		GAC		GGC		C	
		GUA		GCA		GAA	GGA	A			
		GUG	V Val / S TART	GCG		GAG	E Glu	GGG		G	

Tableau I-1 Code génétique universel. Les codons alternatifs principaux sont notés après une barre oblique (Sec = sélénocystéine, Pyl = pyrolysine). Les codons d'initiation sont marqués START. Les codons synonymes sont dans la majorité des cas regroupés en "cases" (entourées d'un double trait) pour lesquelles les deux premières bases sont identiques.

Biais d'usage des codons

Avec la dégénérescence du code génétique vient la notion essentielle du biais d'usage des codons. Les organismes utilisent la totalité de ces codons, mais les séquençages des génomes montrent également une forte préférence de certains organismes pour des codons particuliers [26, 27]. Et chez un organisme donné, chaque gène est composé préférentiellement certains codons [28] (**Figure I-10**). Par ailleurs, la disponibilité intracellulaire des ARNt isoaccepteurs (portant le même acide aminé) varie également très fortement selon leur anticodon. L'utilisation par l'organisme des codons correspondant aux ARNt les plus disponibles permet donc d'optimiser la synthèse protéique. À l'inverse, l'utilisation de codons peu courants

permet de diminuer le taux de traduction de certaines protéines [28, 29]. Il est également proposé que des codons rares placés à des points stratégiques du gène permettent de stopper temporairement la traduction, laissant ainsi le temps à la protéine de se replier [30]. Le biais d'usage des codons permet donc un niveau supplémentaire de régulation de la traduction qui dépend directement de la biodisponibilité des ARNt dans l'organisme.

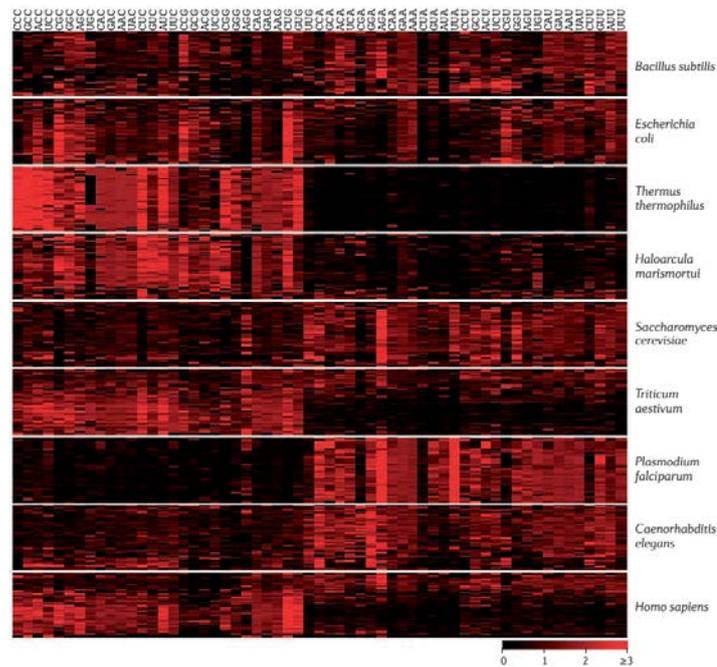


Figure I-10 Biais d'usage des codons dans différents organismes (Adapté de [26]). L'utilisation relative des codons est montrée pour 50 gènes pris au hasard dans différents organismes. Les gènes sont les lignes, les codons les colonnes. On remarque que certains organismes (*T. thermophilus*, *P. falciparum* en particulier) n'utilisent presque jamais certains codons. Chez l'Homme, le biais d'usage dépend très fortement du gène sélectionné ce qui laisse imaginer une importante régulation par ce moyen.

I.2.2 Activités catalytiques des ARNt

Rôle dans le transfert de peptide au niveau du ribosome

L'exemple principal de l'activité d'aide à la catalyse des ARNt est celui du centre peptidyl-transférase du ribosome. L'une des étapes clés de cette catalyse est l'activation par déprotonation du groupement α -NH₂ de l'acide aminé porté par l'ARNt (dans le site A du ribosome). Il n'y a pas de groupement catalytique protéique au niveau du site peptidyltransférase, la catalyse est donc forcément ribonucléique [31]. Si le résidu A₂₄₅₁ de l'ARNr 23S a tout d'abord été proposé comme le groupement catalytique principal [32], des mutations ont montré que ce résidu ne jouait pas de rôle essentiel dans la catalyse [33, 34]. Les structures cristallographiques ont montré que le seul groupe capable d'agir comme une base était alors le 2'OH de l'ARNt dans le site P. Des expériences avec des ARNt modifiés, soit par un 2'OF, soit un déoxy-2', ont montré une baisse drastique de l'efficacité catalytique du centre peptidyl-transférase [35]. Il est donc très probable que le groupement 2'OH du nucléotide A₇₆ agisse comme une base générale pour la formation de la liaison peptidique (**Figure I-11**).

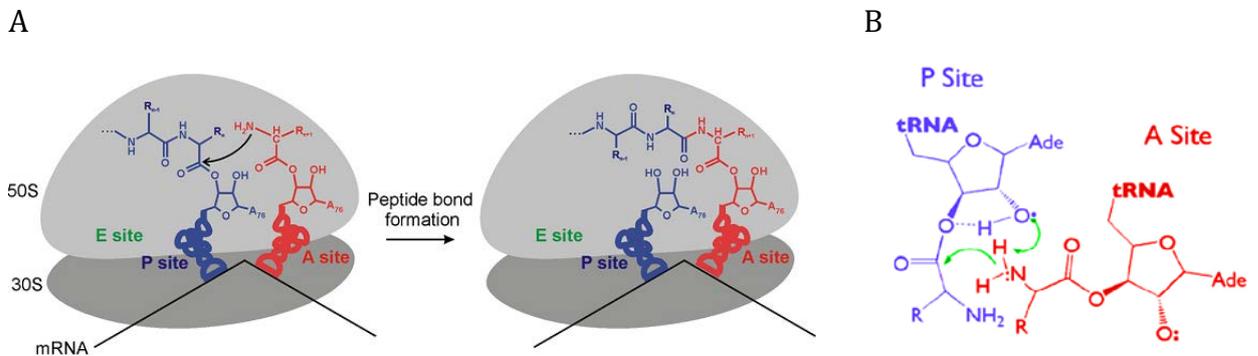


Figure I-11 *Activité catalytique de l'ARNt dans le site P du ribosome. A position relative des ARNt et des différents groupements terminaux dans le ribosome [36]. B Détail de la réaction : Le groupe α -NH₂ dans le site A (en rouge) fait une attaque nucléophile sur le carbone du carbonyle du peptide en cours de synthèse dans le site P*

(en bleu). Le groupement 2' OH de l'ARNt dans le site P agirait comme une base générale dans cette réaction [37].

Rôle dans leur propre acylation/déacylation

Il semblerait d'après plusieurs études [38–40] que les ARNt participent directement à la catalyse de leur propre acylation et déacylation. Il a été montré pour l'aaRS thréonine, que seule l'histidine 309 de cette protéine est primordiale pour la catalyse, or cette histidine lie directement le 2'OH terminal de l'ARNt^{Thr} [40, 41]. Ici encore le remplacement du 2'OH par un groupement fluor ou sa suppression conduisent à une diminution très forte de la catalyse indiquant l'implication de ce groupe dans la catalyse [40]. Les auteurs ont proposé que l'histidine 309 déprotone le groupement 2'OH, qui active à son tour le 3'-OH qui recevra l'acide aminé (**Figure I-12 A**). Enfin, un mécanisme du même type a été proposé dans le cadre de la déacylation des ARNt misacylés par l'aaRS Phe, qui est capable de détecter la misacylation et de déacyler l'ARNt. Dans ce cadre, la réaction étant une hydrolyse, elle est médiée par une molécule d'eau [38, 39] (**Figure I-12 B**).

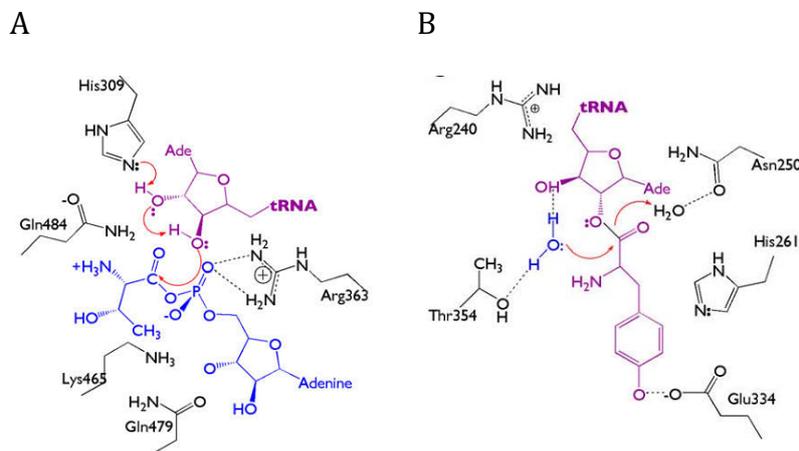


Figure I-12 Mécanisme d'acylation/déacylation d'ARNt. (A) Mécanisme catalytique proposé pour l'acylation de l'ARNt^{Thr} [40]. L'histidine 309 arrache le proton du groupe 2'OH du nucléotide 76 de l'ARNt, qui déprotone à son tour groupement en 3' qui va

alors attaquer le carbone du carbonyle et être ainsi acylé. (B) Mécanisme proposé de la réaction de déacylation par la synthétase de l'ARNt^{Phe} misacylé [37].

Une dernière activité dans la catalyse sera abordée dans la section suivante.

I.2.3 Donneur d'acide aminé

Certains acides aminés sont incorporés dans des peptides dont la synthèse est indépendante du ribosome, ou sont utilisés comme précurseurs dans la production de petits métabolites. Les ARNt aminoacylés sont alors souvent utilisés comme donneurs d'acides aminés.*

Le peptidoglycane bactérien

La muréine ou peptidoglycane bactérien est le composant principal de la paroi bactérienne. Cette paroi permet aux bactéries de ne pas exploser, malgré la très forte pression osmotique entre leur milieu intérieur et extérieur. La muréine est un polymère complexe formé par quatre éléments (**Figure I-13**) :

- Une structure principale composée par l'alternance de deux sucres : N-acétylglucosamine (NAG) et acide N-acétylmuramique (NAM) liés en $\beta(1,4)$
- Lié au NAM, on trouve une chaîne peptidique de trois à cinq acides aminés. Dans cette chaîne on peut notamment trouver des acides aminés D (D-Ala et D-Glu en particulier) ce qui est unique aux parois bactériennes.
- Des ponts interpeptidiques qui peuvent comporter jusqu'à cinq acides aminés (cinq glycines chez *S. aureus*) reliant entre eux tous les peptides deux à deux, formant une structure maillée complexe.
- Des acides lipotéichoïques qui fixent le peptidoglycane à la membrane bactérienne (**Figure I-13 C**).

* Le terme de synthèse peptidique non-ribosomale ne doit alors pas être utilisé car ce processus est défini comme également indépendante des ARNt.

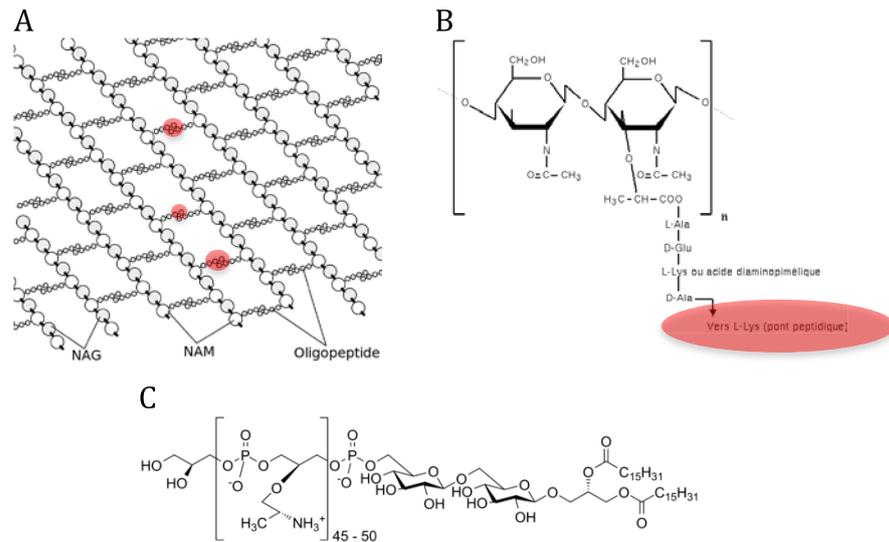


Figure I-13 Structure du peptidoglycane. A) Structure générale. B) Structure chimique du NAG (tout à gauche) et NAM. Une chaîne d'acide aminé est liée au NAM par une liaison peptidique. C) acide lipoteichoïque *. La position du pont peptidique est montrée sur fond rouge.

Les chaînes peptidiques sont synthétisées de façon indépendante des ribosomes ou des ARNt dans le cytoplasme [42], les ponts inter-peptidiques ont par contre besoin pour leur synthèse d'ARNt aminoacylé (aa-ARNt) donneurs d'acides aminés [43, 44]. Chez *S. aureus*, il existe un ARNt^{Gly} qui n'a pas d'activité dans la traduction ribosomale et qui est dédié uniquement à la production du pont inter-peptidique [45]. La synthèse de ce pont est catalysée par la famille de protéines transférases Fem qui sont toutes des peptidyl-transférases non ribosomales de la superfamille des GNAT acétyl-transférases [46].

La structure récente de FemX_{Wv} en complexe avec une molécule de synthèse ARN-peptide-UDP-MurNAc (**Figure I-14**) a permis de montrer que l'ARNt est impliqué dans la catalyse enzymatique (**Figure I-15**) [47].

* Source des illustrations : Peptidoglycane.svg (public) et Mureine.svg (CC-BY-SA-3.0) - Lipoteichoic acid.svg (CC-BY-SA-3.0) -@ Wikimedia commons

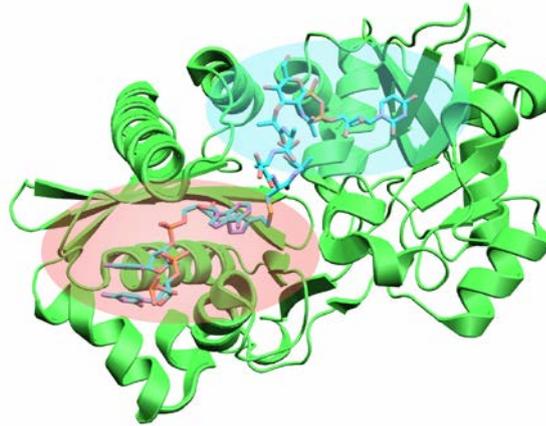


Figure I-14 Structure de FemX_{Ww} en complexe avec un ARN-peptide-UDP-MurNac [47]. La partie ARN est sur fond rouge et le pentapeptide-UDP-MurNac en bleu. (Code d'accèsion PDB : 4II9)

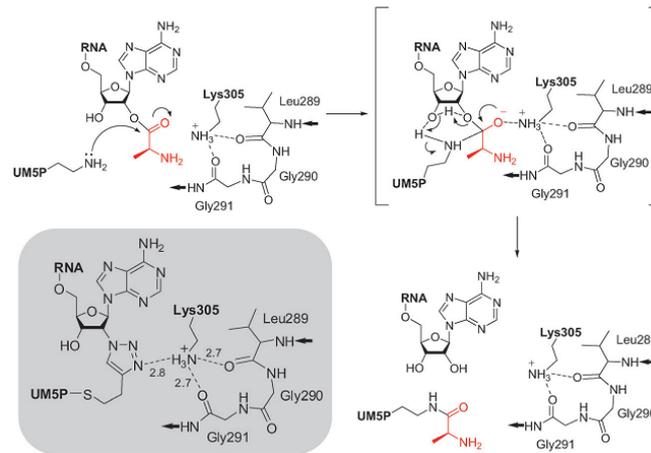


Figure I-15 Modèle proposé pour la réaction d'aminoacylation (adapté de [47]). UM5P : peptide-UDP-MurNac ; RNA : ARnt. Le groupement 2'OH du ribose de A₇₆ joue un rôle actif dans les transferts de protons impliqués dans la catalyse.

Dégradation des protéines : N-end rule

La "N-end rule" (règle du N-terminal) est une voie de signalisation entraînant la dégradation des protéines qui est conservée chez tous les organismes, eucaryotes comme procaryotes [48]. Ce système détermine la demi-vie des protéines en

fonction de leur résidu N_{ter}. Ce résidu terminal aura soit une action stabilisante, soit déstabilisante selon son type. La dégradation des protéines passe par le système ubiquitine-protéasome chez les organismes qui le possède (eucaryotes et certains procaryotes), ou par des protéases spécifiques chez les autres organismes. Une modulation de ce système existe *via* les protéines L, F, et R-transférases. Ce sont des peptidyl-transférases capables d'ajouter un résidu en N-terminal d'une protéine, en utilisant un ARNt comme donneur, leur site catalytique étant homologue à celui des protéines Fem X/A/B [49]. Les transférases L et F (leucine et phénylalanine) sont bactériennes, et la transférase R (arginine) est eucaryote. Les résidus F, L et R sont tous les trois des résidus fortement déstabilisant [50]. L'ajout en N_{ter} de ces acides aminés provoque la dégradation des protéines cibles.

Residue X	<i>E. coli</i>	<i>S. cerevisiae</i>
Arg	2 min	2 min
Lys	2 min	3 min
Phe	2 min	3 min
Leu	2 min	3 min
Trp	2 min	3 min
Tyr	2 min	10 min
His	> 10 h	3 min
Ile	> 10 h	30 min
Asp	> 10 h	3 min
Glu	> 10 h	30 min
Asn	> 10 h	3 min
Gln	> 10 h	10 min
Cys	> 10 h	> 30 h
Ala	> 10 h	> 30 h
Ser	> 10 h	> 30 h
Thr	> 10 h	> 30 h
Gly	> 10 h	> 30 h
Val	> 10 h	> 30 h
Pro	?	> 5 h
Met	> 10 h	> 30 h

Tableau I-2 Règle du N-terminal [48]. Demi-vie approximative de la protéine \underline{X} - β gal ou \underline{X} est le résidu N_{ter}, chez *E. coli* à 37 °C et *S. cerevisiae* à 30°C.

Modification de la membrane bactérienne

Les aminoacyl phosphatidyl glycérol synthases (aa-PG synthase) sont un groupe d'enzymes qui attache différents acides aminés à des phosphatidyl-glycérol (PG). Ces lipides modifiés neutralisent les charges globalement négatives des membranes

(groupement phosphate des phospholipides). Ils rendent ainsi les membranes plus imperméables aux peptides antibactériens et aux antibiotiques cationiques (classe des aminoglycosides par exemple) [51]. Toutes les aa-PG synthases sont dépendantes d'ARNt comme donneur d'acide aminé [37].

Synthèse de composés organiques

Les acides aminés portés par les ARNt peuvent être des précurseurs dans des synthèses de composés organiques. On peut citer par exemple la synthèse de l'antibiotique valinomycine (**Figure I-16 A**) qui utilise deux acides aminés comme précurseurs, la valine et la sérine ; la sérine étant apportée par un ARNt [52]. L'ensemble des protéines responsables de la production de cet antibiotique est regroupé dans un cluster de quatorze gènes [53]. L'ARNt sérine utilisée dans cette réaction est aminoacylé par une aa-ARNt synthétase spécifique –et essentielle faisant partie du cluster de gènes : VlmL [54]. La raison de la présence d'une synthétase spécifique n'est pas claire, puisque l'ARNt^{Ser} utilisé est un ARNt élongateur normal de la synthèse peptidique. Les cyclodipeptides synthases, responsables de la synthèse de dipeptide cyclique (**Figure I-16 B**) bactérien (cyclo L-Phe-L-Pro entre autres) sont également dépendantes des ARNt comme donneurs d'acides aminés dans ces synthèses [55]. Ces peptides très variés ont des activités notamment antibiotique ou antifongique. Enfin, on peut citer le cas de la synthèse des tétrapyrroles (**Figure I-16 C**) (hèmes, chlorophylles, bilines –pigments biliaires-...) chez les plantes, les algues et certaines bactéries. Ce processus utilise comme précurseur le glutamate chargé sur son ARNt^{Glu} [56, 57].

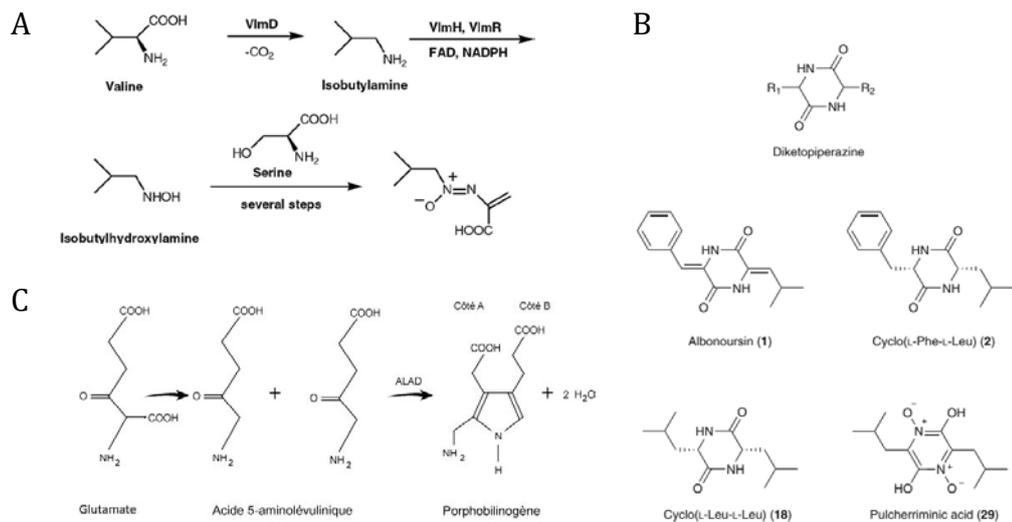


Figure I-16 Les acides aminés portés par les ARNt sont précurseurs de nombreuses réactions de synthèse biochimique. (A) Le séryl-ARNt^{Ser} est utilisée pour la biosynthèse de la valinomycine [52]. (B) Différents cyclodipeptides, leurs précurseurs sont des aa-ARNt [55]. (C) le glutamyl-ARNt^{Glu} est le précurseur des tétrapyrroles par la voie du porphobilinogène [56, 57].

I.2.4 Implications des ARNt dans les infections virales

L'un des rôles les plus connus des ARNt en dehors de la traduction, est leur utilisation par les rétrovirus et certains rétrotransposons, comme initiateur de la transcription inverse. Les transcriptases inverses sont des polymérases ADN, ARN dépendantes, qui convertissent donc des ARN en ADN. Elles ont besoin pour cela d'une amorce qui s'apparie sur la matrice ARN, les ARNt sont détournés de leur fonction première pour jouer ce rôle d'amorce. Le génome du VIH comporte une région appelé PBS ('primer binding site') complémentaire sur 18 bases en 3' de l'ARNt₃^{Lys} cellulaire [58]. L'hybride entre le PBS et l'ARNt sert d'amorce pour la transcriptase inverse lors de la synthèse du brin d'ADN (-). Lors de la synthèse du brin (+) la séquence de l'ARNt ne doit pas être répliquée au-delà des 18 nucléotides complémentaires du génome viral. C'est la méthylation de l'adénine 58 (m¹A₅₈) de

l'ARNt₃^{Lys} qui permet d'arrêter la polymérisation du brin, la polymérase étant incapable de répliquer cette base [59–61].

I.2.5 Fragments d'ARNt

Les ARN non codants pour des protéines (ARNnc) sont omniprésents dans la vie cellulaire. Il a longtemps été cru qu'une vaste partie du génome eucaryote et procaryote était inutile, car ne codant pas pour des protéines (junk DNA [62]). On sait aujourd'hui que la majorité des génomes sont transcrits en ARN, mais qu'une petite partie seulement sera traduite en protéine. Le reste restant sous forme d'ARNnc qui peuvent avoir une variété de rôles allant du *silencing* de gènes par interaction avec les ARNm (leur première fonction connue) à la maintenance de la chromatine (Pour une introduction [63] et pour des revues [64, 65]). Le sujet étant vaste, nous n'aborderons ici que les activités d'ARNnc liées aux fragments d'ARNt.

La dégradation régulée des ARNt est connue depuis longtemps. Dans les années 90, il a été montré qu'*E. coli* produit en cas d'infection par le phage T4 l'endonucléase PrrC qui digère les ARNt^{Lys} dans la boucle anticodon. La déplétion en ARNt^{Lys} ainsi induite ralentit la traduction des protéines du phage, et ralentit sa propagation dans la colonie [66]. Les fragments de dégradation des ARNt ont été longuement observés dans les cellules, mais ce n'est que récemment que l'on s'est rendu compte qu'ils pouvaient également avoir une activité fonctionnelle comme ARNnc [67]. Toutes les fonctions abordées ci-dessous sont résumées dans la **Figure I-18** en page 28.

Demi-ARNt

Les demi-ARNt sont produits par le clivage dans la boucle anticodon d'ARNt mature, portant la plupart du temps toutes leurs modifications post-transcriptionnelles [68, 69]. Le déclencheur de ce clivage est en général un stress cellulaire spécifique (nutritionnel, physicochimique...). Les protéines les plus connues pour réaliser ce clivage étant Rny1 chez la levure, et l'angiogénine chez les eucaryotes évolués.

L'angiogénine est une protéine de la famille des RNase A dont la fonction la plus anciennement documentée est l'angiogénèse est de ce fait directement impliquée dans la croissance des tumeurs cancéreuses [70, 71]. C'est également une protéine cytotoxique activée par le stress cellulaire et dégradant spécifiquement certains ARNt. Il a été montré que la moitié 5' de ces ARNt clivés est capable de diminuer fortement la vitesse de traduction, mais le mécanisme est encore débattu [67]. Il semblerait que l'association de ces ARNnc avec des facteurs d'initiation de la traduction soit une piste. Ces demi-ARNt par leur interaction avec le complexe Dicer-2 pourraient aussi être des régulateurs directs de la production de petits ARN interférents (siARN). En l'inhibant, ils provoquent la traduction d'ARNm spécifiques au stress cellulaire [72]. Une forte activité pro-prolifération cellulaire des demi-ARNt est depuis quelque temps mise en avant, et leur implication dans certains processus cancéreux est de plus en plus claire [73, 74] (**Figure I-17**).

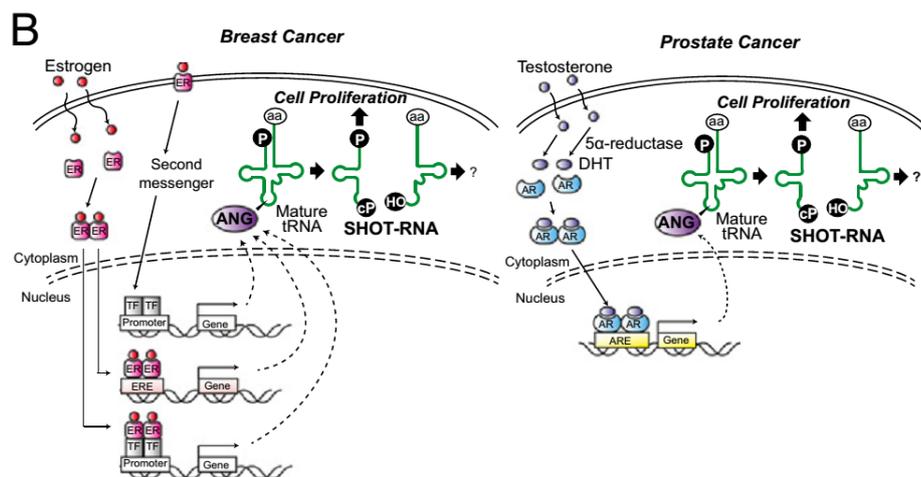


Figure I-17 Mécanisme proposé d'implication des demi-ARNt dans la prolifération cellulaire induite par les hormones sexuelles dans les cancers du sein et de la prostate [73]. Dans les deux cas, la production d'angiogénine est augmentée par les hormones sexuelles provoquant le clivage d'ARNt en demi-ARNt.

Fragments 5' d'ARNt (5'tRF)

Les fragments 5' sont produits par la coupure dans ou autour de la boucle D, ce sont donc des fragments de plus petites tailles. Comme pour les demi-ARNt, ces fragments proviennent d'ARNt mature. Mais il ne peut pas être exclu que ces fragments proviennent plutôt de demi-ARNt-5', et soient ainsi le produit d'une seconde étape de dégradation. Souvent, la majeure partie des 5'tRF retrouvés provient d'un seul ARNt particulier ce qui laisse plutôt penser qu'ils possèdent une voie de biosynthèse spécifique [67]. Leur activité intracellulaire n'est pas parfaitement établie, mais il semblerait qu'ils soient capables de lier la machinerie de production d'ARN interférent. Par leur grande quantité, ils entreraient en compétition avec les substrats normaux de cette machinerie, et provoqueraient ainsi une inhibition de l'ARN interférence [75]. Il a également été montré que ces ARN sont capables de se lier au ribosome (*in vivo* et *in vitro*) en inhibant ainsi la traduction. [76, 77]

Fragments 3'CCA

Ces fragments sont produits par clivage dans la boucle T, et portent l'extrémité 3'CCA. Ils proviennent également d'un ARNt mature. L'enzyme responsable de cette production est très probablement Dicer, mais peut-être pas uniquement [78, 79]. Ces ARN s'associent à la machinerie ARN interférence et semblent avoir une action d'ARN interférents. Il semblerait également que ces ARN aient une activité antivirale [78]. Un mécanisme proposé étant que ces ARN lient des ARN viraux qui utilisent des ARNt comme amorces, et les dirige vers DICER pour la dégradation. Un système qui rappellerait le système immunitaire CRISPR-CAS9 bactérien [80].

Fragment 3'U

Ces derniers types de fragments sont produits à partir de précurseurs d'ARNt (pré-ARNt), et se terminent par une longue série de U. Il semblerait, de façon surprenante, que ces ARNnc soient produits dans le cytoplasme à partir de pré-

ARNt alors que ceux-ci ne devraient pas être présents hors du noyau [81]. Ils ont également un rôle dans l'ARN interférence mais sont encore mal connus [73].

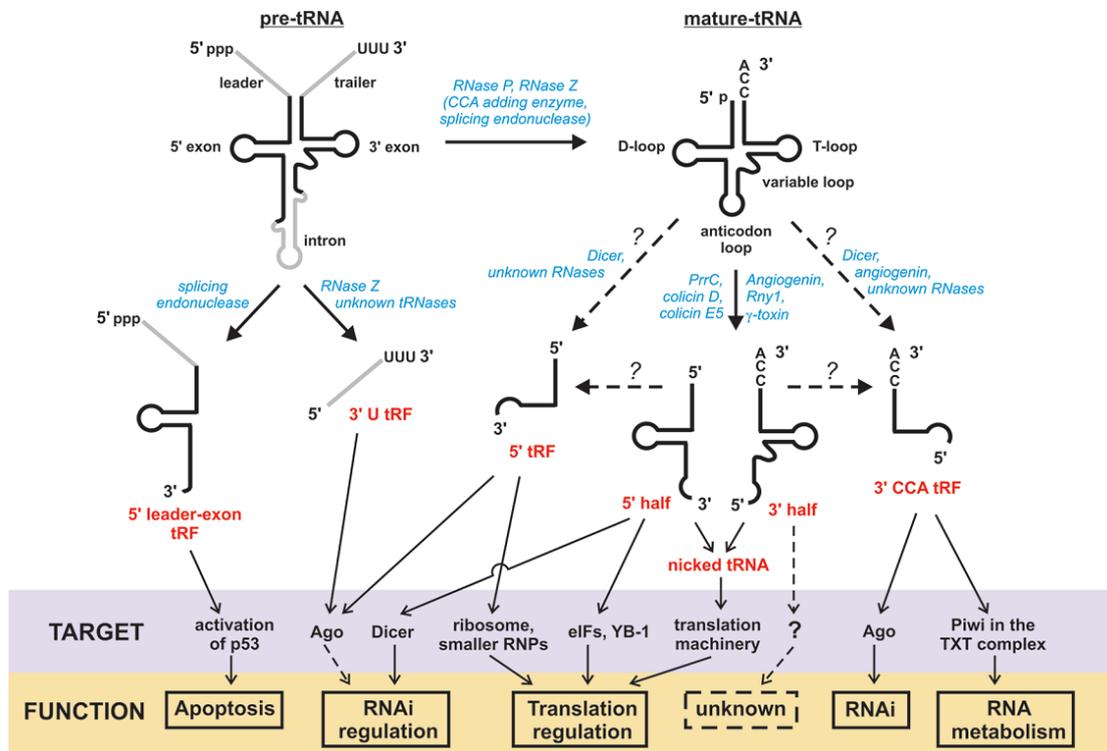


Figure I-18 Les fragments de dégradation d'ARNt ont de nombreux rôles dans la signalisation intracellulaire (adapté [67]). Différentes voies métaboliques connues (flèches pleines) ou supposées (pointillés) sont représentées. Il manque pour compléter ce schéma, les actions des 5'-demi-ARNt sur la prolifération cellulaire [73].

Implication dans des pathologies humaines

Outre l'implication dans les processus cancéreux déjà évoquée, le métabolisme des fragments d'ARN semble impliqué dans certains processus pathologiques (pour une revue voir [82]).

Des mutations de l'angiogénine réduisant son activité nucléasique ont été directement impliquées dans la pathogénogénèse de la Sclérose Latérale Amyotrophique (ALS ou maladie de Charcot) [83], mais également dans la maladie de Parkinson [84]. Par ailleurs, la majorité des mutations causales de ces maladies

sont des pertes de fonctions dans des RNases [85], et l'administration d'angiogénine sur les cultures cellulaires ou des modèles murins montrent des effets protecteurs de cette enzyme [86, 87]. Ces résultats laissent entrevoir l'implication forte des fragments d'ARNt dans ces processus pathologiques, peut-être par leur capacité à inhiber l'apoptose [88], ou à promouvoir les granules de stress [68].

La méthyltransférase 2 de la famille NOP2/Sun (NSUN2) produit le m^5C_{34} de l'ARNt^{Leu}_{CCA} [89] mais aussi aux positions 48, 49 et 50 d'autres ARNt [90, 91]. La mutation de ce gène est responsable de plusieurs pathologies, dont certaines formes de handicap mental [92, 93], ou le syndrome de Dubowitz [94]. En effet, l'absence de méthylation sur ces ARNt, provoque leur dégradation et l'accumulation de fragments 5' qui agissent comme ARN interférent et limite la croissance cellulaire, expliquant les retards de développement dans ces différents syndromes [95].

Ce cas souligne l'importance des modifications post-transcriptionnelles dans la régulation des activités des fragments d'ARNt. Les fragments provenant majoritairement d'ARNt mature, ils portent toutes les modifications toutes les modifications post-transcriptionnelles des ARNt dont ils sont issus. Les modifications post-transcriptionnelles font en effet parties de la maturation des ARNt, qui fait l'objet du chapitre suivant.

I.3 Maturations des ARN de transfert

Les précurseurs des ARNt sont synthétisés à partir de gènes spécifiques. Chez les bactéries et les archées, il n'existe qu'une seule ARN polymérase, et c'est donc celle-ci qui prend en charge la synthèse des ARNt. Les eucaryotes possèdent plusieurs ARN polymérases (ARN pol) qui sont responsables de la synthèse spécifique des ARN ribosomiques pour l'ARN pol I, des ARN messagers pour l'ARN pol II, des ARN de transfert ainsi que l'ARN ribosomique 5S et d'autres petits ARN pour l'ARN pol III. Chez les plantes, une ARN polymérase supplémentaire l'ARN pol IV est responsable de la synthèse des ARN interférents. Les ARNt sont synthétisés par l'ARN polymérase sous forme de précurseurs avec des extrémités 5' et 3' devant être maturées et peuvent aussi contenir des introns. Par ailleurs, les ARN pol n'incorporant que les nucléotides canoniques, les précurseurs ne portent aucune modification. On peut distinguer plusieurs étapes majeures qui seront détaillées dans les prochains paragraphes (pour des revues complètes voir [96–99]) (**Figure I-19**):

- L'excision de l'extrémité 5' par la RNase P
- L'excision de l'extrémité 3' par plusieurs endo/exonucléases
- L'addition du CCA terminal chez les eucaryotes (souvent déjà présent chez les bactéries, et parfois chez les archées)
- L'épissage des introns
- L'incorporation de modifications post-transcriptionnelles (qui interviennent tout au long de la maturation)
- L'aminoacylation

Il est important de noter que l'ordre de ces étapes, ainsi que leur localisation cellulaire peut varier selon l'organisme et/ou l'ARNt considéré.

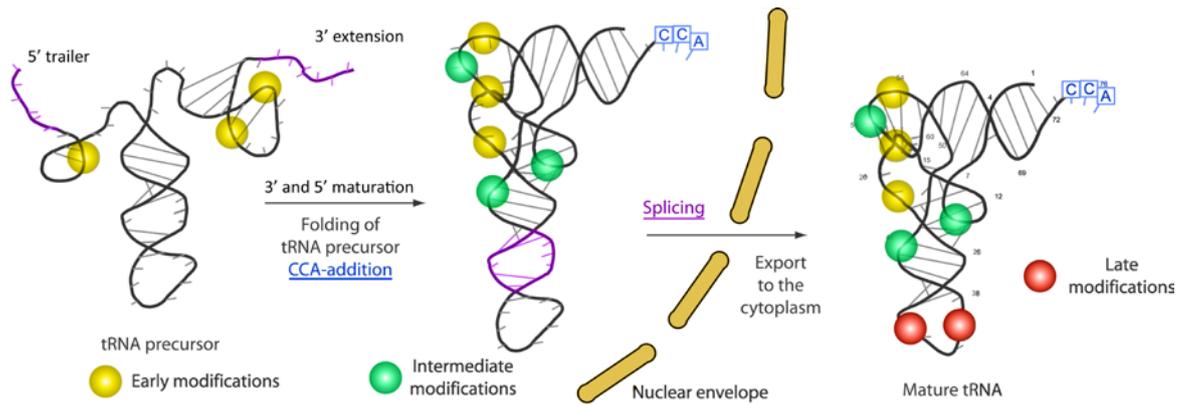


Figure I-19 Biogenèse des ARNt (adapté de [97]). Les ARNt sont synthétisés sous forme de précurseurs pour lesquels les extrémités 5' et 3' sont maturées, l'intron épissé (si présent), et le CCA terminal ajouté. Ils sont ensuite exportés hors du noyau cellulaire. Tout au long de cette maturation, des modifications post-transcriptionnelles peuvent être incorporées. L'ordre et la localisation (cytoplasme/noyau) de ces étapes varient d'un organisme à l'autre.

I.3.1 Maturation de l'extrémité 5'

La maturation de l'extrémité 5' se fait sous l'action d'une endonucléase retrouvée chez tous les organismes connus* : la ribonucléase-P (RNase P) (**Figure I-20**). Ce complexe nucléoprotéique est composé d'un ARN et d'une protéine chez les procaryotes [100]. En l'absence de protéine, l'ARN a toujours une activité catalytique à forte force ionique [101]. C'est donc l'ARN du complexe qui catalyse le clivage, il s'agit d'un ribozyme. L'ARN possède, en plus de son domaine catalytique, un domaine de reconnaissance des ARNt. Chez les eucaryotes, le complexe contient souvent plus d'une protéine (10 sous-unités protéiques chez la levure [102]). Il a récemment été montré que la RNase P eucaryote serait également un ribozyme

* À l'exception notable de l'archée *Nanoarchaeum equitans* qui synthétise ses ARNt directement sous forme mature en 5' -l'extrémité 5' étant alors tri-phosphate [11].

[103], mais l'ARN seul a une activité catalytique très faible, les protéines associées étant indispensables au moins à sa structuration [104].

Les RNases P des mitochondries et des chloroplastes, bien que ces organites dérivent de bactéries, ne contiennent pas d'ARN catalytique, la fonction étant portée par des protéines sans aucune relation avec la RNase P cellulaire [105].

I.3.2 Maturation de l'extrémité 3'

La maturation de l'extrémité 3' fait intervenir tout d'abord des endonucléases (RNase E et III chez les bactéries) qui clivent l'extension 3' en son milieu [106], puis des exonucléases (RNase II, BN' et PH toujours chez les bactéries) qui finissent de digérer l'extension. Chez les eucaryotes, la digestion peut également être catalysée par la RNase Z qui coupe en 5' du CCA [107] (**Figure I-20**).

Cette étape de digestion peut laisser un ARN mature avec le CCA, mais souvent l'addition du CCA à l'ARNt sera effectuée *a posteriori* par une enzyme spécifique ("CCA-adding" enzyme). Cette enzyme a la particularité d'être une ARN polymérase matrice indépendante, elle est capable de reconnaître les ARNt et d'ajouter une séquence spécifique (CCA) sans avoir pour autant besoin d'une matrice ARN ou ADN pour la guider [108, 109]. Cette enzyme possède une seule poche catalytique capable de fixer le CTP comme l'ATP [110].

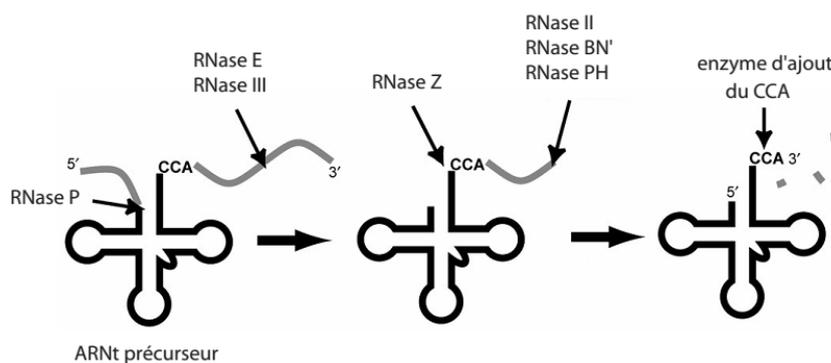


Figure I-20 Maturation des extrémités 5' et 3' (Adapté de Nakanishi & Nureki [96]). La RNase P clive l'extrémité 5' en une seule étape. L'extrémité 3' est tout d'abord clivée

en son milieu par des endonucléases, puis la digestion est terminée par des exonucléases. Enfin, une nucléotidyl-transférase vient réparer l'extrémité CCA si besoin.

I.3.3 Épissage

Les ARNt peuvent contenir un intron qui est épissé. Il est situé entre les bases 37 et 38. Chez les bactéries, les introns sont des introns autocatalytiques (de groupe I) [111]. Les eucaryotes et les archées ont en revanche un mécanisme d'épissage spécifique aux ARNt, qui contrairement au spliceosome des ARNm, ne fait pas intervenir d'ARN. L'épissage est catalysé par trois enzymes capables de reconnaître tous les substrats. L'ARNt est tout d'abord clivé par une endonucléase membranaire [112] (membrane interne du noyau cellulaire) au niveau des deux sites d'épissage. Une phosphotransférase NAD-dépendante supprime le groupe phosphate resté en 5' [113]. Les deux demi-ARNt sont ensuite reliés entre eux par l'ARNt ligase [114].

I.3.4 Aminoacylation

L'aminocyclation des ARNt est prise en charge par une enzyme spécifique à chaque acide aminé. Les organismes ont donc en général une aaRS pour chacun des vingt acides aminés, ces enzymes étant capables de reconnaître différents ARNt isoaccepteurs. L'aminocyclation des ARNt se fait en deux étapes, la première étant l'adénylation de l'acide aminé pour donner un acide aminé-AMP, cette réaction pouvant se produire en l'absence d'ARNt. L'acide aminé est ensuite transféré sur l'ARNt cible et l'AMP relargué. Les aaRS sont les garantes du code génétique dans le sens où ce sont les seules molécules qui reconnaissent à la fois l'acide aminé, et l'anticodon et/ou les déterminants structuraux définissant l'ARNt. Elles sont séparées en deux grandes classes selon leur architecture générale, la classe I adoptant une architecture en pli de Rossmann qui lie l'ATP, ce repliement est aussi retrouvé dans les méthyltransférase de classe I (*voir page 57*). De nombreuses aaRS

possèdent en plus une capacité d'édition permettant d'hydrolyser la liaison entre l'acide aminé et l'ARNt en cas de mauvais appariement (pour des revues voir les références [115, 116]).

Les aaRS reconnaissent la structure complète des ARNt, et ont de ce fait un rôle important dans le contrôle qualité de la structure des ARNt. L'acylation est en effet souvent une condition *sine qua non* pour l'export nucléo-cytoplasmique de l'ARNt [117].

Il est intéressant de noter que chez un grand nombre de bactéries les aaRS pour la glutamine et l'asparagine sont absentes [118], les ARNt dont les codons correspondent à ces acides aminés sont chargés par les aa-synthétases glutamate (pour la glutamine) et aspartate (pour l'asparagine). Ces ARNt sont donc au départ misacylés. Le complexe entre l'ARNt misacylé et la synthétase sont reconnus et pris en charge par des amidotransférases spécifiques (Glu-AdT et Asp-AdT) qui vont catalyser la production des bons acides aminés [119]. Ces enzymes reconnaissent spécifiquement la structure des ARNt misacylés [120]. Un mécanisme du même type existe chez certaines archées pour l'ARNt^{Cys}, qui est chargé par la synthétase Serine.

I.4 Modifications post-transcriptionnelles

Les modifications post-transcriptionnelles font partie intégrante de la maturation des ARNt, mais le sujet étant vaste, et central dans ma thèse, une partie complète leur est consacrée.

I.4.1 Nature chimique des modifications

Nous avons déjà très rapidement abordé l'existence de bases modifiées dans les ARNt. Ces bases modifiées diffèrent par leur nature chimique des quatre bases canoniques A, C, G, U. Les modifications sont post-transcriptionnelles parce qu'elles sont incorporées sur l'ARN après l'étape de transcription, l'ARN polymérase n'incorporant que des bases canoniques dans la molécule produite. Elles ne sont donc pas inscrites dans le gène codant pour l'ARN, mais sont dues à des enzymes qui catalysent les réactions de modification. La nature des modifications peut être très variée (**Figure I-21**) : on trouvera fréquemment la simple addition d'un groupement méthyle (m^5C par exemple), mais aussi des additions de groupement beaucoup plus complexes (on pourra citer la mannosyl-queuosine – manQ des ARNt eucaryotes), voire plusieurs groupements ($gcmnm^5s^{2*}$ pour citer un exemple découvert récemment [121]) (**Figure I-21 B**). D'une façon générale, on regroupera les modifications dans deux grandes catégories selon la complexité des voies métaboliques qu'elles engagent, on discernera ainsi :

Les modifications simples, pour lesquelles une seule réaction enzymatique est nécessaire, comme les méthylations, thiolations, isomérisation (la pseudouridine ψ est obtenue par isomérisation de l'uridine), réduction (dihydrouridine D)...

Les modifications complexes, sont généralement des suites de modifications simples sur le même nucléotide (à des positions différentes ou non), et résultent toujours de voies métaboliques impliquant de multiples enzymes. Ainsi on trouve

* géranyl-5-carboxyméthylaminométhyl-2-thiouridine

des modifications aussi complexes que la $gcmnm^5s^2U$ ou la mannosyl-queuosine citées précédemment.*

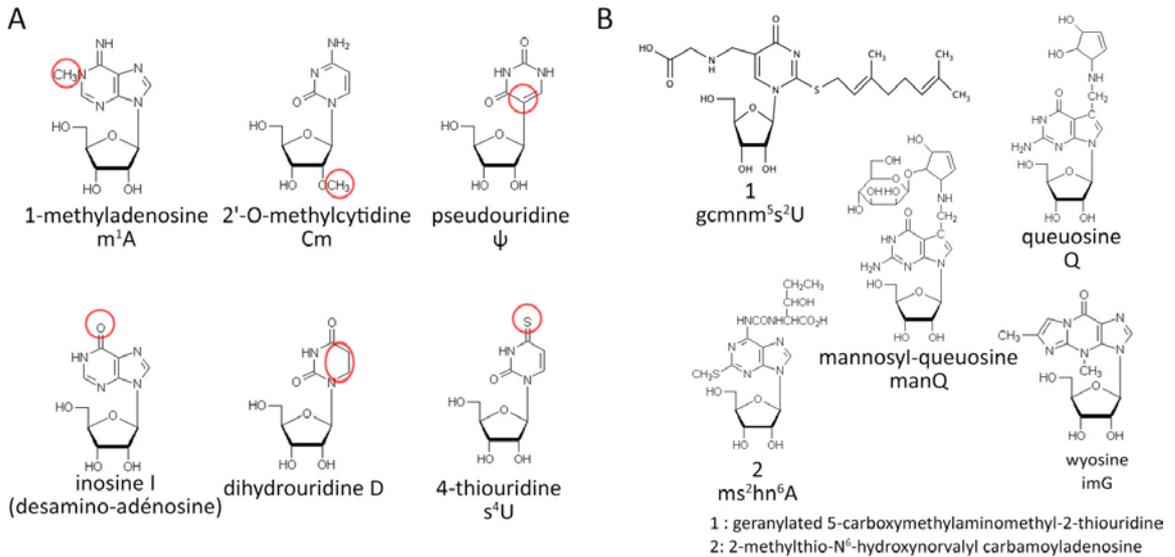


Figure I-21 Exemples de bases modifiées des ARN, (A) modifications simples, la position de la modification est indiquée par un cercle rouge (B) modifications complexes. †

La majorité des atomes des bases, carbone ou azote, endo ou exocyclique peuvent porter des modifications. Des modifications sur le ribose sont aussi possibles. Dans ce cas la modification sera toujours en 2'OH.

I.4.2 Diversité et répartition des nucléotides modifiés

Parmi les ARN, ce sont les ARNt qui portent le plus grand nombre et qui montre la plus grande diversité de modifications post-transcriptionnelle (**Figure I-22**). Ainsi, 109 nucléotides modifiés distincts ont été identifiés dans les différents ARN, dont 93

* Différents exemples de voies métaboliques de formation des bases modifiées sont consultables à l'adresse : <http://modomics.genesilico.pl/pathways>

† Toutes les structures proviennent de « The RNA Modification Database », <http://mods.rna.albany.edu/>

dans les ARNt [122]. Le nombre de modifications sur un ARNt donné varie entre trois et une vingtaine (**Tableau I-3**). Si l'on considère, non plus le type d'ARN, mais les règnes du vivant dans lesquels on retrouve ces modifications, on remarquera qu'une grande partie des modifications est spécifique d'un règne seul un faible nombre d'entre elles se retrouve dans les trois règnes (**Figure I-23**).

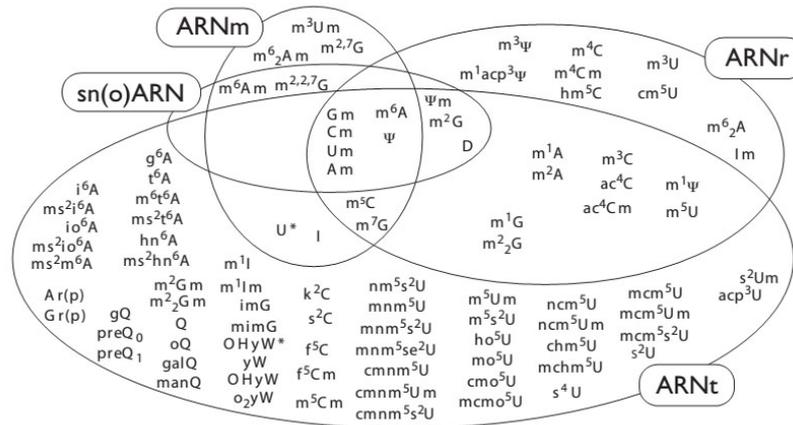


Figure I-22 Répartition des diverses modifications parmi les ARN (adapté de [123]). Seul un petit groupe de modifications sont retrouvé dans tous les types d'ARN. Les ARNt regroupent le plus grand nombre de modifications et les plus complexes.

Type d'ARN	ARNt 74-98 nt	sn(o)ARN 70-1200 nt	ARNr 1600-4000 nt	ARNm 100-3000 nt
Nombre de nt modifiés par ARN	3-20 nt (3-26 %)	0-13 nt (0-10 %)	65-200 nt (1-3 %)	0-5 nt (< 1 %)
Diversité	93	11	31	13

Tableau I-3 Diversité des modifications parmi différents types d'ARN (adapté de [122]). Les ARNt sont les plus modifiés, et leurs modifications les plus diverses. (nt : nucléotide)

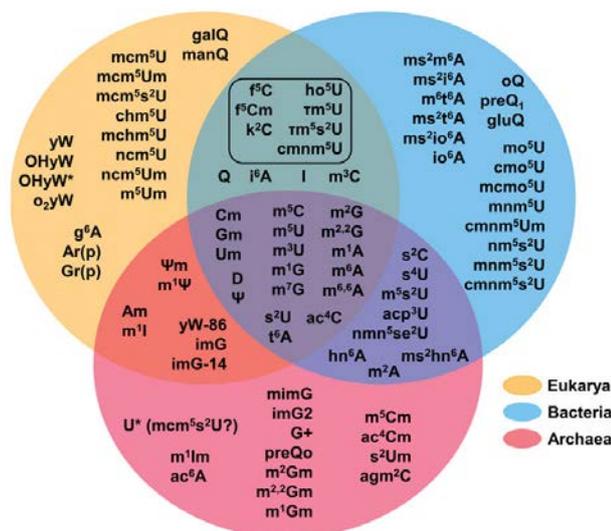


Figure I-23 Répartitions des modifications au sein des trois règnes du vivant (adapté de [124]). La majorité des modifications sont spécifiques d'un règne, seule une petite fraction est retrouvée dans les trois règnes. Ce sont majoritairement des modifications simples. Les modifications encadrées sont spécifiques des organites (chloroplastes et mitochondries).

Les modifications simples sont les plus répandues en nombre. On notera également que les seules modifications communes à tous les types d'ARN sont des méthylation et la pseudouridine (**Figure I-22**). Les modifications communes à tous les règnes du vivant sont en grande majorité simples également. Ces modifications communes sont très probablement les plus anciennes d'un point de vue évolutif (apparition avant la différenciation des trois règnes), et certains auteurs évoquent la possibilité que la méthylation m^7G_{37} des ARNt commune à toutes les espèces soit, sinon la première, une des premières modifications apparues du fait de son importance dans la stabilisation du cadre de lecture [125].

La base modifiée la plus communément trouvée dans tous les règnes du vivant et tous les types d'ARN est la pseudouridine [126]. On la retrouve notamment sur plusieurs sites conservés des ARN ribosomiaux où elle est indispensable [127]. La présence de cette base modifiée dans les ARN nucléaires est également

indispensable à l'épissage des ARN messagers [128, 129]. Sa récente détection dans les zones codantes des ARN messagers laisse également imaginer un rôle régulateur important dans la synthèse protéique [130–132]. Il est en particulier proposé que la pseudouridylation des uridines des anticodons soit à la base de codages alternatifs des codons stops [133].

I.4.3 Nucléotides modifiés des ARN de transfert

Environ la moitié des nucléotides des ARNt peuvent porter des modifications, mais certaines positions sont modifiées avec une fréquence beaucoup plus élevée que d'autres (**Figure I-24**). La tige acceptrice est plutôt peu modifiée, et ce sont les boucles T, D et anticodon qui reçoivent le plus de modifications. Il est par ailleurs notable, que certaines de ces positions hypermodifiées sont également des positions extrêmement conservées (position 34 et 37 dans la boucle anticodon, 54, 55 et 58 du bras T ou encore 47 de la région variable). Enfin, les deux nucléotides qui portent les modifications les plus variées et souvent les plus complexes sont la position wobble (34) de l'anticodon, et le nucléotide qui suit juste l'anticodon à la position 37 [134].

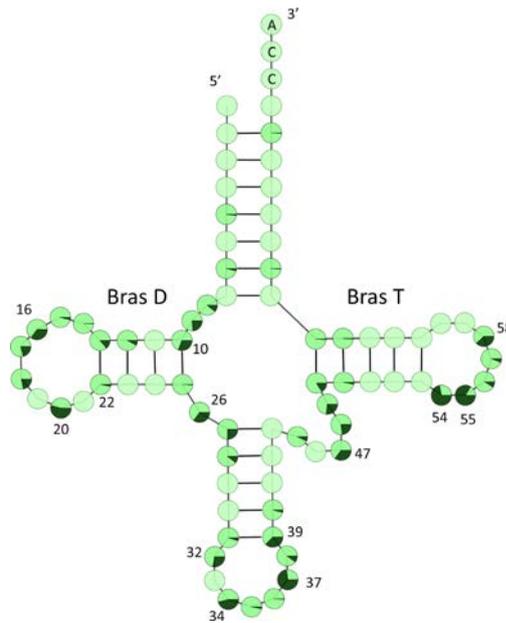


Figure I-24 Fréquence de modification des nucléotides pour chaque position des ARNt. Fréquence calculée sur les séquences des ARNt disponibles dans la base de données Modomics [135–137] (<http://genesilico.pl/trnamodviz>). La portion en noir des cercles dessinés à chaque position correspond à la proportion modifiée.

La position 34 ou wobble

Comme expliqué précédemment (Fonction des ARNt p. 12), le code génétique est dégénéré. Plusieurs codons codent pour le même acide aminé, et il n'existe pas autant d'ARNt que de codons. Des ARNt sont capables de reconnaître plusieurs codons différents, ceci grâce à un appariement moins exigeant de type wobble entre le troisième nucléotide du codon et le premier nucléotide de l'anticodon de l'ARNt (position 34) (**Tableau I-1 p.15**). L'appariement wobble se fait par la face Watson-Crick des bases, les contraintes stériques à l'intérieur du ribosome ne permettant ni à l'ARNt ni à l'ARNm de présenter les faces Hoogsteen ou sucre de leurs bases. Par ailleurs, les deux premières paires de bases codon/anticodon sont toujours de type Watson Crick. La microhélice résultant de l'appariement est reconnue par le ribosome de façon très spécifique par huit liaisons hydrogène, un mésappariement

codon/anticodon empêchant la formation de ces liaisons hydrogène provoque l'éjection de l'ARNt du ribosome [24, 138, 139].

L'appariement wobble repose en premier lieu sur quatre types de paires de bases : la guanine avec l'uracile G \circ U et l'inosine avec soit la cytosine, l'uracile, ou l'adénine I \circ C/U/A (**Figure I-25**). L'inosine n'étant pas une base canonique, sa formation par désamination de l'adénine est indispensable. Il est notable que cette modification n'est pas une exception, mais la règle : une adénine en position 34 non désaminée en inosine n'a été trouvée que dans un seul ARNt jusqu'à présent [140]. Par ailleurs, il a été montré qu'un ARNt dans le site A du ribosome et portant une adénosine en position 34 rend plus difficile le recrutement d'un autre ARNt dans le site P [141]. D'autres modifications permettent comme l'inosine d'étendre les possibilités de reconnaissance de la base 34, ou de les diminuer. En effet, l'uridine non modifiée en position 34 a la possibilité de reconnaître l'adénine, mais aussi n'importe laquelle des autres bases [21]. C'est pourquoi des modifications spécifiques servant à contraindre la position wobble sont également nécessaires. L'uridine est donc presque toujours modifiée (s²U₃₄; mnm⁵U₃₄; cmo⁵U₃₄...) même lorsqu'elle doit reconnaître une adénine (Pour une revue détaillée de la position wobble voir [24]).

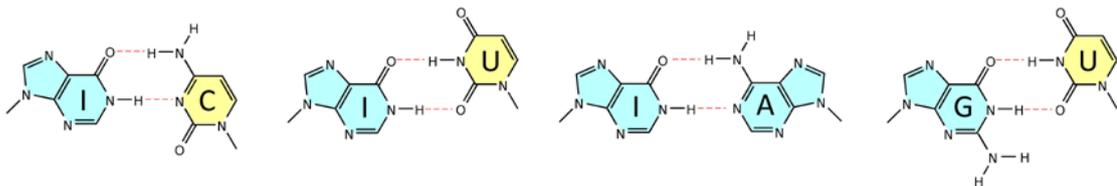


Figure I-25 Principales paires de bases de l'appariement wobble de gauche à droite I \circ C/U/A et G \circ U.

La position purine 37

La position 37 de l'ARNt est l'autre base hypermodifiée de la boucle anticodon. Elle suit directement l'anticodon et joue plusieurs rôles importants dans la traduction.

Tout d'abord, elle a un rôle dans la stabilisation des appariements wobble. Lors d'un appariement entre deux pyrimidines en position wobble, la structure de la microhélice codon/anticodon est déformée par la forte proximité des deux bases [142, 143], ce qui provoque le rejet de l'ARNt. Ce désassemblage n'a pas lieu si la position 37 porte des modifications adéquates pour stabiliser la structure de l'hélice codon/anticodon en forçant la structuration correcte de l'anticodon [24].

Les modifications de la purine 37 ont également un rôle primordial dans la prévention du glissement du cadre de lecture lors de la traduction, en stabilisant l'appariement codon/anticodon, et en rendant celui-ci moins sensible au contexte de séquence du codon. Par exemple, la modification $ms^2i^6A_{37}$ (2-méthylthio-N6-isopentényl-adénosine) présente dans la majorité des ARNt dont l'anticodon fini par un A, empêche le glissement du cadre de lecture. Le soufre vient faire une interaction d'empilement avec le cycle de l'uridine 1 du codon, cette interaction restant stable dans les trois sites du ribosome [144] (**Figure I-26**).

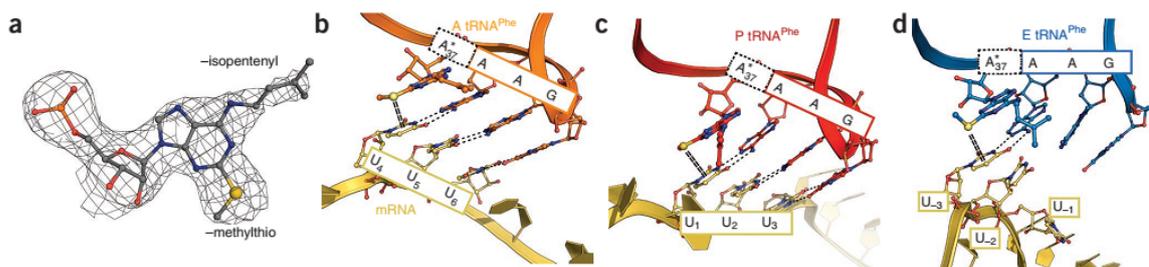


Figure I-26 Structure du ribosome 70S à 3.1 Å [144] (code d'accèsion PDB 4V6F), (a) densité électronique autour de la base modifiée $ms^2i^6A_{37}$; (b), (c) et (d) interaction de la base modifiée avec le premier nucléotide du codon dans les sites A, P et E du ribosome. L'atome de soufre de la modification méthyle-thiol de l'adénine A_{37} fait une interaction d'empilement avec l'uridine U_1 du codon dans les trois sites.

Structuration et stabilisation des ARN de transfert

Le rôle des bases modifiées sur la structure des ARNt est étudié de longue date. De nombreuses études de dénaturation thermique ont tout d'abord été conduites

[145–148]. Puis, d'autres techniques ont été utilisées comme le marquage chimique et enzymatique [149–151], la cristallographie aux rayons X [152], ou encore la résonance magnétique nucléaire (RMN) [153, 154], [155]. Ces différents travaux s'accordent sur le fait que la plupart des ARNt non modifiés (produit par transcription *in vitro*) possèdent le repliement canonique des ARNt matures, à l'exception de certains ARNt mitochondriaux qui, s'ils ne possèdent pas la modification m¹A₉, n'adoptent pas la structure secondaire en feuille de trèfle [156].

Des effets locaux sur la structure de l'ARNt sont notables aux environs des modifications. Ces effets structuraux peuvent être dus à des différences d'interaction avec les molécules d'eau ou avec les cations en solution, mais aussi à des modifications d'interactions tertiaires au sein de l'ARNt. Il a été montré, par exemple, que les nucléotides modifiés participaient à la fixation d'ion magnésium. Ainsi, du fait de sa moins bonne chélation du magnésium, un transcrit *in vitro* a besoin d'une concentration de Mg²⁺ plus élevée pour se replier qu'un ARNt extrait de bactéries [157]. Il a également été proposé que l'effet de rigidification locale dû à la présence de pseudouridine soit dû à la fixation particulière des molécules d'eau autour de cette base [152].

Certaines modifications changent de façon profonde la structure, le potentiel électrostatique, ou encore la propension à faire des interactions d'empilement des bases modifiées. La conformation C3'-endo du sucre est favorisée dans l'ARN par rapport à la conformation C2'-endo qui est plutôt présente dans l'ADN. Ceci étant dû à la présence du 2'OH du ribose. La **2'-O-méthylation** (méthylation du ribose en 2'OH) favorise plus encore la conformation C3'-endo en augmentant l'encombrement du groupement en position 2'. Cette modification a pour effet général de rigidifier la structure et d'augmenter la stabilité thermodynamique de l'ARN. Les uridines thiolées en position 2 (**s²U**) stabilisent également la structure des ARNt en favorisant la conformation C3'-endo. Les méthylation **m⁵C** et **m⁵U** stabilisent les ARN en favorisant l'empilement des bases. Au contraire, le cycle de la **dihydrouridine** n'étant pas plan, les interactions d'empilement avec cette base sont

fortement défavorisées. De plus, cette modification favorisant la conformation C2'-endo du sucre, elle apporte de la flexibilité à l'ARN. Enfin, les méthylations **m¹A**, **m³C** et **m⁷G** ont la particularité de conférer une charge nette positive au niveau de la base. Cette charge peut avoir un rôle important dans la compaction et la stabilité de l'ARN en écrantant les charges négatives du squelette phosphate, au même titre que les cations du solvant évoqués précédemment. La **pseudouridine** comporte un donneur de liaisons hydrogène supplémentaire par rapport à l'uridine ce qui a aussi pour effet de stabiliser les ARN.

Les bras T et D et la région variable portent un grand nombre de modifications, elles ont pour rôle d'augmenter la stabilité générale de l'ARNt en le rigidifiant ou l'assouplissant localement. On retrouvera également des modifications qui, comme pour les positions 37 et 34, modulent les capacités d'appariement de la base modifiée, afin d'empêcher des repliements non souhaitables [158], et de favoriser des appariements à longue distance entre les boucles T et D. Par exemple, l'enzyme qui catalyse la formation de m¹A₅₈ chez les eucaryotes est essentielle, bien que cette modification soit loin de l'anticodon [159]. Elle est indispensable à l'initiation par l'ARNt^{Met} [159, 160]. La structure de l'ARNt^{Met} est en effet très particulière [161] au niveau des boucles T et D. Trois adénines (A₂₀, A₅₄, et A₆₀) conservées dans les initiateurs eucaryotes se structurent autour de la m¹A₅₈ avec notamment une paire de bases *trans*-Hoogsten A₅₄•m¹A₅₈ (**Figure I-27**). L'absence de la modification provoque un repliement aberrant de cet ARNt qui est alors polyadénylé pour être ensuite dégradé [162].

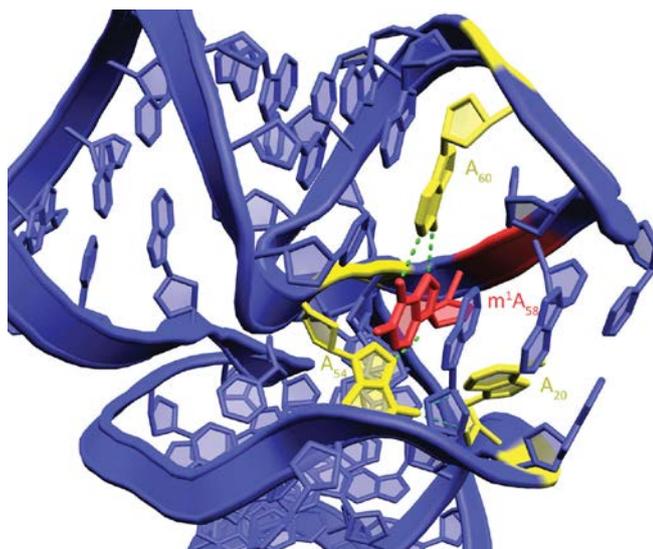


Figure I-27 Structure de l'ARN_t_i de levure (code d'accèsion PDB : 1YFG) [11]. Le nucléotide *m*¹A₅₈ (en rouge) est au cœur d'un réseau d'interactions entre plusieurs adénines (en jaune) conservées de l'ARN_t_i^{Met}. Les liaisons hydrogène sont indiquées en pointillés verts.

I.4.4 Fonctions biologiques des modifications post-transcriptionnelles

Les études de mutations des gènes pour les enzymes de modifications ont montré que si certaines modifications sont essentielles pour la croissance cellulaire, la plupart des mutations n'ont pas d'effet phénotypique. Pourtant, leurs conservations extrêmes dans les organismes, malgré un coût de production énergétique non négligeable, laissent imaginer une pression de sélection très forte pour cette persistance. Certains phénotypes sont néanmoins décrits de longue date. Par exemple en positions 34 et 37, il n'est pas rare d'observer des ralentissements de croissance en cas de mutations, du fait du rôle important de ces modifications dans la traduction [163, 164]. Chez les organismes extrémophiles, on repère assez facilement la présence de modifications spécifiques aux conditions de développement de ces organismes : augmentation de la stabilité par la présence de nombreuses pseudouridines chez les thermophiles, ou au contraire augmentation

de la flexibilité par des dihydrouridines chez les psychrophiles, chacune de ces modifications provoquant des variations de la stabilité aux températures extrêmes.

Pour beaucoup d'autres modifications, les mutants ne présentent pas de phénotype, mais le développement récent de tests plus sensibles commencent à démasquer le rôle de certaines modifications. La pseudouridine ψ_{55} est une modification ultra conservée, mais la mutation du gène *truB* chez *E. coli*, responsable de sa production ne produit aucun phénotype identifiable [165]. L'utilisation de culture en compétition entre cellules mutées et cellules sauvages a montré après plusieurs cycles de cultures que les cellules mutées sont peu à peu supplantées par les sauvages, la compétition leur étant légèrement défavorable [165]. La fonction précise de la ψ_{55} et les mécanismes de cette perte de compétitivité restent néanmoins inconnus.

Une autre approche est l'utilisation de test –dit test synthétique– mettant en avant les interactions génétiques. De nombreuses mutations pour des modifications qui sont sans phénotype apparent peuvent devenir létales chez des doubles mutants. Cette approche a été utilisée notamment avec le gène *Trm2*, qui encode pour la méthyltransférase responsable de la méthylation m^5U_{54} donc de la présence très conservée de la ribothymidine en position 54 des ARNt. La mutation ou la délétion de cette méthyltransférase chez *S. cerevisiae* ne donne aucun phénotype [166, 167], mais dans des souches où le gène *sup61+* de l'ARNt_{GGA}^{Ser} est inactivé, la double mutation avec *Trm2* ralentit fortement la croissance [168]. Le rôle stabilisateur de la T_{54} n'est donc visible que sur certains des ARNt isoaccepteurs sérine.

I.4.5 Dynamique des modifications

Il a longtemps été pensé que les modifications étaient un phénomène statique, l'ARNt à la fin de sa maturation portant toutes les modifications post-transcriptionnelles qui lui sont connues. Il est de plus en plus clair, que ces modifications font en réalité partie d'un processus dynamique, qui permet à un organisme de s'adapter aux changements de ses conditions environnementales.

Adaptation aux conditions environnementales

Thermus thermophilus est une bactérie thermophile extrême qui vit majoritairement dans les sources chaudes et peut se développer normalement dans une large gamme de température allant de 50 à 83°C. Trois modifications distinctes des ARNt sont liées à cette thermophilie : Gm₁₈, m⁵s²U₅₄, et m¹A₅₈, la combinaison de ces trois modifications augmentant la température de fusion des ARNt d'environ 10°C [169, 170]. Ces modifications ne sont pas toujours présentes, et leur présence varie en fonction de la température de culture (**Figure I-28 A**). Le taux de certaines modifications varie donc en fonction des conditions de vie de l'organisme, ce qui lui permet de maintenir sa synthèse protéique aux températures extrêmes [171, 172].

Par ailleurs, il a été montré que cette adaptation fonctionne en réseau de modification (**Figure I-28 B**): la délétion individuelle des gènes de la ψ_{55} ou de la m⁷G₄₆ provoque des changements forts du taux des autres modifications [173–176]. On peut en tirer ce schéma global [172] :

- À des températures en dessous de 65°C, le nucléotide ψ_{55} augmente la rigidité de l'ARNt et inhibe l'interaction de l'ARNt avec les enzymes responsables de l'adaptation aux plus hautes températures (Gm₁₈, m⁵s²U₅₄, et m¹A₅₈).
- À température supérieure à 65°C, la rigidification de l'ARN par la ψ_{55} n'est plus suffisante pour empêcher les autres modifications. Par ailleurs la modification m⁷G₄₆ est favorisée, celle-ci accélérant fortement l'apparition des autres modifications et déclenchant une cascade d'adaptation à la température.

Il ne semble donc pas y avoir d'activité génétique pour ces adaptations à la chaleur (pas d'augmentation de la production d'enzymes de modification), c'est directement la structure de l'ARNt en fonction de la température qui promeut ses propres modifications.

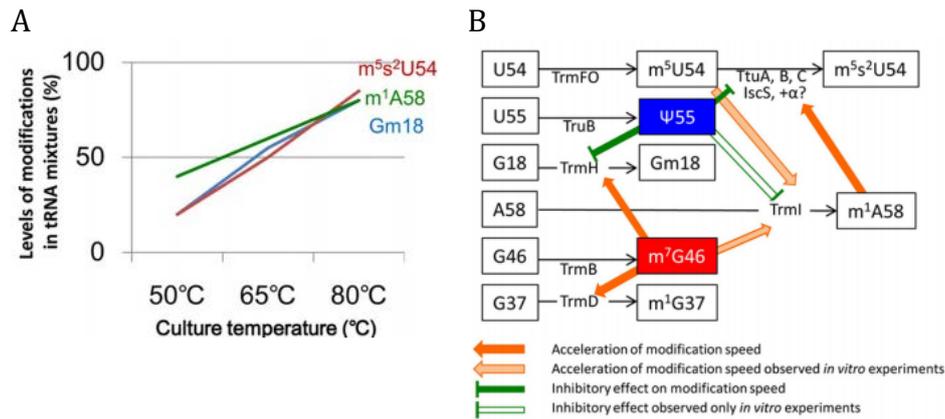


Figure I-28. Réseau de modifications et d'adaptation à la température de *T. thermophilus* (adapté de [172]). (A) Taux de modifications des ARNt de *T. thermophilus* en culture à différentes températures. Les ARNt sont plus modifiés avec l'augmentation de la température ce qui augmente leur température de fusion. (B) Réseau de modifications des ARNt mis en jeu dans l'adaptation au changement de température chez *T. thermophilus*.

Adaptation aux conditions cellulaires

La régulation de la traduction des protéines par le biais d'usage des codons a été déjà évoquée (*Biais d'usage des codons* en page 15). C'est un processus dynamique permettant à la cellule de répondre aux conditions environnementales. L'effet du biais d'usage des codons peut en effet être modulé par des modifications spécifiques des ARNt. Par exemple, la méthyltransférase Trm9 méthyle les uridines en position 34 des ARNt^{Arg}_{UCU} et ARNt^{Glu}_{UUC}. Elle catalyse en fait la dernière étape de la formation des bases mcm⁵U et mcm⁵s²U [177]. Plusieurs études impliquaient cette méthyltransférase dans la réparation de l'ADN [178, 179] (résistance aux produits chimiques transformants et aux radiations ionisantes). Ce gène est également présent chez l'Homme sur un locus chromosomique (8p22) fréquemment délété dans les cancers colorectaux [180]. Il a été montré que c'est un important suppresseur de tumeur dans ce type de cancer [181]. Une étude bio-informatique a permis de mieux comprendre les mécanismes en jeu [182]. Les auteurs ont identifié

425 gènes de *S. cerevisiae*, tous liés par les biais de codon AGA et AGG (correspondant aux anticodons spécifiquement modifié par Trm9), parmi lesquels d'importants gènes de réponse aux dommages dans l'ADN. Trm9 en améliorant significativement la reconnaissance codon/anticodon [183] module le biais d'usage des codons et augmente la traduction de ces protéines de protection de l'ADN.

Le même type de régulation a été montré dans le cadre de la réaction au stress oxydant. L'utilisation d'eau oxygénée sur des cultures de *S. cerevisiae* produit une augmentation notable de m⁵C notamment catalysée par Trm4. La délétion de ce même gène provoque une hypersensibilité à H₂O₂ [184]. Cette modification n'est retrouvée en position wobble que sur un seul ARNt : l'ARNt^{Leu}_{CAA}. Une étude récente a montré que l'exposition à H₂O₂ augmente le taux de cette m⁵C particulière de près de 70% [185]. Les auteurs ont pu montrer que cette hausse de modification provoquait l'augmentation de la traduction de protéine contenant le codon Leu_{TTC}. La levure possède pour l'une de ses protéines ribosomales (Rpl22), deux paralogues A et B, dont les leucines sont codées avec respectivement 100% ou 34% de codons Leu_{TTC}. Les auteurs ont également pu lier directement le paralogue Rpl22A avec la résistance au stress oxydant. Le fonctionnement d'une partie au moins de ce système de résistance est donc très clair : le stress oxydant augmente la modification m⁵C₃₄ de l'ARNt^{Leu}_{CAA}, ce qui augmente directement la traduction des protéines ayant un biais de codon fort pour le codon Leu_{TTC}, dont en particulier le paralogue de la protéine Rpl22A ayant un rôle dans la résistance au stress oxydant.

Ces deux exemples sont, à ce jour, les plus aboutis dans la compréhension des systèmes de régulations cellulaires impliquant des modifications des ARNt. Mais il a été montré que pour une grande variété de stress cellulaires, on peut détecter une signature spécifique au niveau des modifications des ARNt [184]. Les modifications ont donc des rôles importants dans la dynamique de réponses cellulaires, notamment par la modulation du biais de codon, mais très certainement par d'autres mécanismes qui restent à élucider.

I.4.1 Modifications post-transcriptionnelles des ARNt et pathologies humaines

De plus en plus de maladies génétiques sont reliées à des mutations d'enzymes de modifications post-transcriptionnelles des ARNt. Des déficits dans les modifications affectent en général l'efficacité et la fidélité de la traduction. Plusieurs exemples sont donnés dans les paragraphes suivants. On peut s'attendre à des phénotypes très aspécifiques pour ce type de mutations, on retrouve souvent en réalité des syndromes très précis, touchant une petite catégorie de processus biologiques, qui sont une fois encore à relier au biais d'usage des codons [186, 187].

Désordre neurologique

L'enzyme de modification FtsJ1, homologue de Trm7 de levure [188], est une méthyltransférase des positions 32 et 34 des ARNt^{Leu}, ARNt^{Phe} et ARNt^{Trp} [188]. Le gène FTSJ1 est situé sur le chromosome X et sa mutation est associée avec des retards mentaux non syndromiques liés au chromosome X [189, 190]. Des études ont montré que ce gène était principalement exprimé dans le cerveau fœtal [189], mais ne l'est plus à l'âge adulte [191]. Cette modification post-transcriptionnelle de ces ARNt a donc certainement un rôle dans l'expression de protéines liées au développement du cerveau [188].

La dysautonomie familiale est une maladie génétique affectant les systèmes nerveux autonomes et sensoriels et est causée par la mutation du complexe elongator [192, 193]. Elongator est un complexe protéique de six sous-unités, conservé des levures aux humains. Des défauts dans ce complexe affectent un grand nombre de fonctions cellulaires dont l'acétylation des histones, l'exocytose, et l'élongation dans la transcription [194]. L'action de ce complexe très conservé sur autant de voies cellulaires est restée longtemps mystérieuse, mais il est aujourd'hui clair que tous ces effets sont en fait liés à la production par elongator de la mcm⁵s²U en position 34 de plusieurs ARNt [194]. Il a été montré qu'elongator influe sur l'acétylation de l' α -tubuline dans les neurones, ce qui pourrait également engager sa dysfonction dans des maladies neurologiques impliquant le trafic cellulaire comme la maladie de

Huntington, d'Alzheimer ou de Parkinson [195]. La mutation de la protéine ELP4 d'elongator est associée avec l'épilepsie rolandique, la forme la plus commune d'épilepsie [196]. Plusieurs allèles d'elongator sont également associés à des formes de sclérose amyotrophique latérale. Il semble enfin qu'il soit associé à certaines formes d'asthme bronchique infantile [197].

Cancer

De nombreuses études proposent des liens directs entre certains cancers et des enzymes de modifications [198–202]. Il est néanmoins essentiel de rappeler que les processus de carcinogenèse sont (hors affection génétique très rare) multifactoriels, et il est indispensable de se garder de discours simpliste. Les études précitées impliquent des effets oncogènes de différents enzymes de modifications, et il devient en effet très clair que les modifications aberrantes des ARNt est un des facteurs importants caractérisant de nombreux cancers.

D'une manière générale, les ARNt sont plutôt hypomodifiés dans les tissus cancéreux [203]. Mais certaines enzymes sont fortement surexprimées, c'est le cas du gène TRMT12 qui est responsable de la production de wybutosine sur l'ARNt^{Phe} en position 37 [204]. Ce gène est surexprimé dans de nombreux cancers du sein. La protéine Trmt12p est même utilisée comme biomarqueur de ce type de cancer [205]. C'est également le cas de NSUN2 qui produit la m⁵C₃₄ et active la prolifération cellulaire, elle est dupliquée dans 35% des cancers du sein [202].

Maladies métaboliques

Le diabète de type II affecte plus de 200 millions de personnes dans le monde. Il est causé par des facteurs environnementaux et génétiques. Plusieurs études le relient avec des mutations du gène CDKAL1 [206–208]. Ce gène code pour une méthyltransférase qui intervient dans la synthèse de la modification complexe ms²t⁶A₃₇ de l'ARNt_{UUU}^{Lys} [209]. Il semble que l'absence de cette modification conduise à la mauvaise traduction de certaines protéines dont notamment la pré-pro-insuline, précurseur de l'insuline. La pré-pro-insuline s'accumulerait alors dans le

réticulum endoplasmique, provoquant le stress de celui-ci et l'intolérance au glucose (pour une revue détaillée voir [210]).

Il existe aussi deux mitochondriopathies qui sont associés à des modifications aberrantes des ARNt : les syndromes MELAS (Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episode), et MERRF (Myoclonic Epilepsy with Ragged Red Fibers) [211–214]. Ces deux syndromes sont causés par l'absence de modification sur l'uridine wobble d'ARNt mitochondriaux : $\tau\text{m}^5\text{U}^*$ de l'ARNt- $\text{mt}_{\text{UAA}}^{\text{Leu}}$ dans le cas du MELAS et $\tau\text{m}^5\text{s}^2\text{U}$ de l'ARNt- $\text{mt}_{\text{UUU}}^{\text{Lus}}$ pour le MERRF. Ces mutations se traduisent par une diminution globale de la traduction mitochondriale dans le cadre du MERRF et une diminution plus spécifique de la traduction de ND6 dans le MELAS. ND6 est l'un des composants du complexe I de la chaîne respiratoire, le gène de ND6 est enrichi en codon UUG que l'ARNt- $\text{mt}_{\text{UAA}}^{\text{Leu}}$ non modifié ne reconnaît pas (pour des revues voir [16, 215]).

* τ est l'abréviation pour taurine

II. La m¹A dans les ARN de transfert

Durant mon travail de thèse, j'ai étudié deux méthyltransférases de la position N₁ des adénines, TrmI et TrmK, qui sont toutes les deux des enzymes bactériennes. Cette partie de l'introduction présente les méthyltransférases des ARNt, leur cofacteur donneur de méthyle et se focalise ensuite sur la modification m¹A pour terminer par la présentation des deux enzymes que j'ai étudiées.

II.1 Les méthyltransférases

La méthylation est un processus extrêmement commun dans le métabolisme, on retrouve des méthyltransférases aussi bien dans le métabolisme des ARN de transfert, que dans celui du maintien de la chromatine, la biosynthèse de petites molécules, la transduction du signal, la détoxification... Il existe donc une grande variété de méthyltransférases présentant des structures peu ou pas apparentées, et utilisant divers cofacteurs comme donneur de groupement méthyle.

II.1.1 Donneurs de méthyle

Il existe une multitude de donneurs de méthyle, chacun donnant naissance à une classe de méthyltransférases différente. Les trois donneurs de méthyle principaux sont : les bétaines, le 5-méthyle-tétrahydrofolate (5CH₃THF), et la S-adénosyl-L-méthionine (SAM **Figure II-1**). Ces trois donneurs de méthyle sont liés entre eux par différents cycles métaboliques (**Figure II-2**). On trouve des donneurs de méthyle plus "exotiques" chez les archées méthanogènes méthylotrophes utilisant par exemple le méthanol, le sulfure de diméthyle, le méthanethiol, la méthylamine, diméthylamine, et triméthylamine, ou encore le tétraméthylammonium [216].

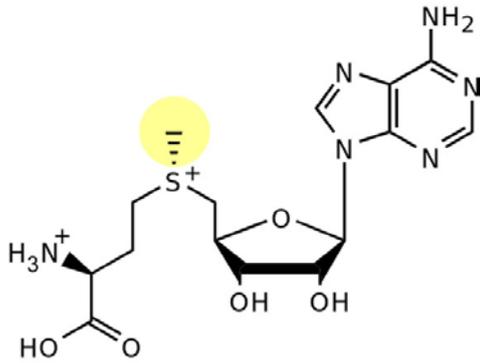


Figure II-1 Structure chimique de la S- adénosyl-L-méthionine. Le groupement sulfonium central porte un méthyle (en jaune) très électrophile du fait de la charge positive adjacente.

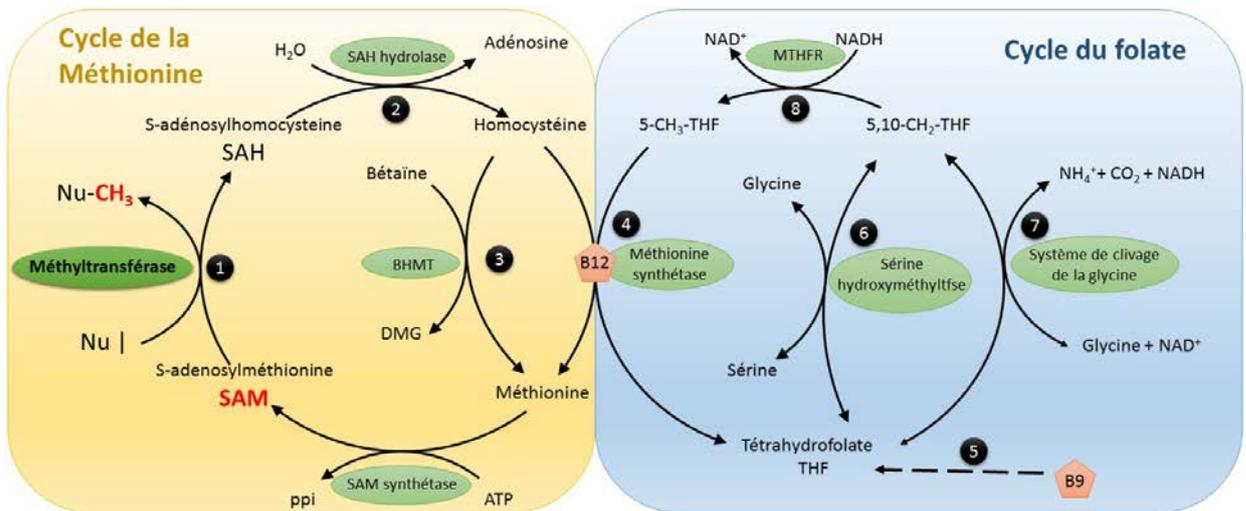


Figure II-2 Cycles métaboliques de la méthionine (encadré en jaune) et du folate (en bleu), sont indiquées certaines des enzymes impliquées (entourée en vert) et les vitamines B (hexagone rouge). BHMT: Bétaïne Homocystéine MéthylTransférase. MTHFR méthyl-tétrahydrofolate reductase.

Lors de la réaction de méthylation par une méthyltransférase SAM-dépendante, un groupement nucléophile d'une molécule cible (Nu|) va recevoir le méthyle très électrophile porté par le sulfonium de la SAM (voir **Figure II-1** et **Figure II-2** ①).

Le produit de la réaction est la S-adénosyl-L-homocystéine (SAH) qui est clivée en adénosine et homocystéine (**Figure II-2** ②).

Plusieurs voies sont possibles pour reconvertir l'homocystéine en méthionine, la voie prédominante étant celle de la méthionine synthase. Il existe deux formes de cette enzyme : l'une est cobalamine (vitamine B₁₂) dépendante, et l'autre non. Dans le cas de la forme B₁₂-dépendante, la vitamine reçoit temporairement le méthyle (méthylcobalamine), avant qu'il soit retransféré sur l'homocystéine. Le donneur de méthyle pour cette réaction est le 5CH₃THF, le produit étant le tétrahydrofolate ou THF (**Figure II-2** ③ ④). Le THF est un dérivé de la vitamine B₉ (l'acide folique), par deux étapes de synthèse non montrées sur le schéma par souci de clarté (**Figure II-2** ⑤).

Les deux voies majoritaires de régénération du 5CH₃THF sont des voies cataboliques, avec d'un côté la sérine hydroxyméthyltransférase qui dégrade la sérine en glycine et de l'autre le système de clivage de la glycine qui la dégrade en ammoniacque et gaz carbonique. Ces deux voies fonctionnent en général ensemble en un catabolisme général de la sérine (**Figure II-2** ⑥ ⑦). Le produit n'est pas directement le 5CH₃THF mais le méthylènetétrahydrofolate (5,10CH₂THF) qui est réduit (**Figure II-2** ⑧).

D'une façon générale, on peut dire que les méthyltransférases dépendantes du 5CH₃THF sont plutôt impliquées dans l'anabolisme des petites molécules, alors que la SAM est le donneur de méthyle de tous les types de réactions, des petites molécules aux histones. En effet, près de 90% des méthyltransférases connues sont SAM-dépendantes. Il a même été longtemps pensé que la SAM était le seul donneur de méthyle biologique [217]. Cette prédominance de la SAM par rapport aux autres donneurs s'explique par la très forte électrophilie du méthyle porté sur le sulfonium de la SAM (**Figure II-1**). Ceci en fait un donneur de méthyl particulièrement efficace. L'enthalpie standard de réaction $\Delta_r G^0$ pour la réaction de méthylation de l'homocystéine en méthionine (SAM + Hcy → SAH + Met) est de -71 kJ.mol⁻¹, soit plus du double de l'hydrolyse de l'ATP en ADP (-30 kJ.mol⁻¹) [217, 218].

La SAM est le second cofacteur enzymatique le plus utilisé dans les cellules juste après l'ATP [217], et sa concentration intracellulaire est très élevée (variant de 100 à 600 μM pour la SAM selon les types cellulaires et jusqu'à 10 mM pour l'ATP intramusculaire [219]). Les méthyltransférases ne sont donc jamais limitées par la concentration en SAM (en condition non pathologique). La SAM n'est pas uniquement un donneur de méthyle, elle entre en jeu dans d'autres réactions importantes comme la synthèse des polyamines (spermines, spermidines...), les voies de trans-sulfuration qui lient les métabolismes des acides aminés soufrés entre eux, et des réactions radicalaires impliquant la décomposition de la SAM en méthionine et 5'-déoxyadénosyl-5'radical par une enzyme à centre [4Fe-4S] [220, 221].

La SAH est un puissant inhibiteur des méthyltransférases avec une constante d'inhibition (K_i) souvent du même ordre de grandeur que la constante de Michaelis K_m pour la SAM de la plupart des méthyltransférases dépendantes de la SAM [222]. Elle est souvent utilisée dans les études structurales, car beaucoup plus stable en solution que la SAM dont la rupture (deux ruptures possibles) et l'inversion du centre chiral ont tendance à se produire en solution avec des constantes de temps non négligeables pour des études longues, de l'ordre de 1% à 3% de dégradation par jour pour chacune de ces trois réactions [220].

II.1.2 Les Méthyltransférases SAM-dépendantes

La première structure d'une méthyltransférase utilisant la SAM comme co-facteur a été résolu en 1993. Il s'agissait d'une méthyltransférase des cytosines de l'ADN M.HhaI [223]. Cette structure a été suivie de nombreuses autres arborant toutes le même repliement structural caractéristique. Ces structures ont été regroupées dans la classe I des méthyltransférases qui regroupent 90% de ces enzymes [218]. Par la suite, 4 autres classes structurales ont été découvertes. Les méthyltransférases de classe I à V utilisent le même mécanisme réactionnel S_N2 (substitution nucléophile bimoléculaire) qui demande que le substrat soit un minimum nucléophile pour

réagir avec le groupe méthyle très électrophile. Il existe un autre type de méthyltransférases SAM-dépendantes dite méthyltransférases à radical SAM (MTRS) capable de méthyler des substrats non nucléophiles en utilisant deux molécules de SAM. Ces enzymes seront abordées brièvement en fin de partie.

Méthyltransférases de classe I

La structure conservée est un feuillet β central de sept brins (six parallèles et un antiparallèle) flanqué d'hélices α formant un pli en sandwich $\alpha\beta\alpha$ (**Figure II-3**). On peut résumer la topologie en quelques caractéristiques :

- La structure est formée par l'alternance entre une hélice α et un brin β
- Le feuillet peut être séparé en deux groupes de brins : 1 à 3 et 4 à 7. Ces deux groupes partant dans des directions opposées ; on a l'alternance suivante : 6↑7↓5↑4↑1↑2↑3↑.
- La zone d'inversion topologique entre les deux groupes de brins forme un sillon où la SAM vient se loger (**Figure II-3 B**).
- Le brin 7 est antiparallèle aux autres et vient s'insérer entre les brins 5 et 6.
- Il n'y a qu'une seule épingle à cheveux β entre les brins 6 et 7 et ces deux brins sont en général plus longs que les autres.

Ce repliement est très comparable au pli de Rossmann décrit bien avant la première structure de méthyltransférase (1974) [224] qui pointait alors l'identité structurale entre de nombreuses enzymes ayant pour trait commun de lier des nucléotides (kinase avec l'ATP, déshydrogénase avec le FAD...) [224] ; et ce, alors qu'aucune homologie de séquence ne les relie. Seule la présence du brin β_7 antiparallèle aux autres brins diffère du pli de Rossmann qui avait alors été décrit.

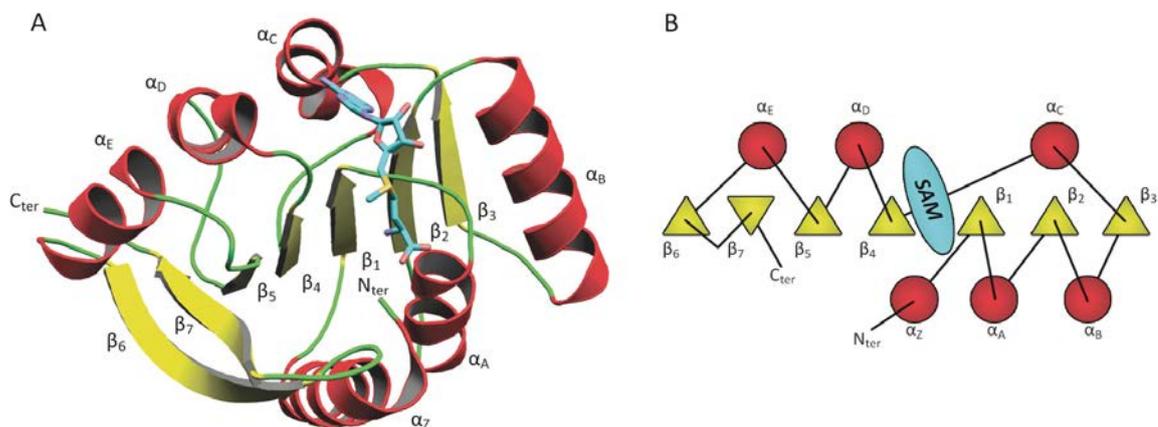


Figure II-3 Méthyltransférase de classe I. (A) Site actif de la catécol-O-méthyltransférase (COMT), les brins β sont en jaune, les hélices α en rouge et les boucles en vert, la SAM est en bleu clair. On repère aisément le pli classique des méthyltransférases de classe I formés d'un feuillet central flanqué de six hélices (code d'accèsion PDB: 1VID [225]) (B) Schéma topologique correspondant, le même code couleur est employé, la pointe des triangles indique la direction des brins. On remarque que le brin β_7 est antiparallèle aux autres.

La partie c-terminale du domaine catalytique des méthyltransférases de classe I montre une variabilité structurale plus élevée que la partie N-terminal. C'est en effet cette partie qui lie le substrat spécifique à chaque enzyme ce qui demande une plus forte adaptabilité que la partie commune liant la SAM. Les PIMT (protéine isoaspartyl méthyltransférases) qui sont impliquées dans la régénération des isoaspartates en aspartates ont, par exemple, leurs brins β_7 et β_6 inversés [226]. Certaines méthyltransférases des arginines ne possèdent même pas du tout ces brins [227]. Enfin, si le repliement du domaine catalytique est conservé, la majorité des enzymes possèdent un domaine supplémentaire en N_{ter} ou en C_{ter} qui permet d'accueillir le substrat [228].

De façon générale, ces protéines sont le plus souvent monomériques, mais des dimères et des tétramères existent [218].

Si la structure tridimensionnelle est très bien conservée, ce n'est pas le cas de la séquence. Le site de fixation de la SAM est souvent très différent d'une protéine à l'autre. La SAM est toujours retrouvée en contact avec les mêmes boucles, mais les résidus qui forment ces contacts varient énormément d'une famille à l'autre [228].

Les autres méthyltransférases SAM-dépendantes

Il existe quatre autres classes de méthyltransférases dépendantes de la SAM, elles seront décrites plus brièvement. Elles comptent en effet à elles quatre seulement 10% des méthyltransférases dépendantes de la SAM [218], et les deux protéines étudiées lors de ma thèse sont de classe I.

La classe II des méthyltransférases ne comporte en fait qu'un seul domaine protéique : le domaine de réactivation de la méthionine synthase. Cette enzyme est une méthyltransférase THF-dépendante. Dans la réaction ⁴ de la **Figure II-2**, le méthyle est tout d'abord donné au cobalt de la vitamine B₁₂ donnant le cofacteur méthyl-Co(III)-cobalamine, qui va lui-même transférer son méthyle et redevenir le Co(I)-cobalamine initial. Mais, en condition aérobie [229, 230], le cofacteur peut périodiquement être oxydé en Co(II)-cobalamine, rendant l'enzyme inactive. La réactivation passe par le transfert de méthyle de la SAM sur le cobalt et la réduction simultanée par le NADPH (et l'enzyme méthionine synthase réductase), ceci malgré un potentiel de réduction du cobalt(II) très défavorable [230–233]. La structuration de ce domaine est très éloignée de la classe I (**Figure II-4 A**).

Il est notable que l'aspect énergétiquement défavorable de cette réaction conduit certains auteurs à penser que le mécanisme en jeu ici n'est pas une méthylation de type S_N2, mais utiliserait en fait les capacités de réactif radicalaire du cofacteur pour réactiver la cobalamine [233]. La classe II des méthyltransférases n'existerait alors pas.

Les méthyltransférases de **classe III** sont impliquées dans la biosynthèse des précorines, intermédiaire de la formation de la vitamine B₁₂ [234]. La diptine

synthase qui modifie une histidine dans la voie de biosynthèse de la diphtamine.* aurait également ce repliement, basé sur un alignement de séquence. On peut voir dans la structure de ces protéines l'équivalent de deux domaines de classe I accolés l'un à l'autre (**Figure II-4 B**).

La **classe IV** ne comprend que des méthyltransférases des ARN de la famille SPOUT. Leurs structures comportent un motif extrêmement rare dans les protéines : un nœud formé par une trentaine de résidus dans le domaine C_{ter} [235]. Hormis cette particularité topologique, leurs structures ressemblent aux méthyltransférases de classe I (**Figure II-4 C**), mais sans aucune homologie de séquence discernable. Ces deux classes pourraient avoir un ancêtre commun très lointain [236].

Les protéines de **classe V** sont des protéines dites à domaine SET. Elles sont associées avec la méthylation des protéines, et sont de ce fait des régulatrices de l'activité protéique. Elles sont impliquées en particulier dans la régulation de la transcription par la méthylation des histones [237–239], mais aussi de p53 [240], ou encore du cytochrome C [241]. Elles ont en commun avec les méthyltransférases de la classe IV le nœud topologique en C_{ter} (**Figure II-4 D**).

* La diphtamine est un acide aminé modifié retrouvé uniquement en une position du facteur d'élongation EF2. Sa présence est essentielle, mais son rôle spécifique reste discuté. C'est également la cible de la toxine diphtérique.

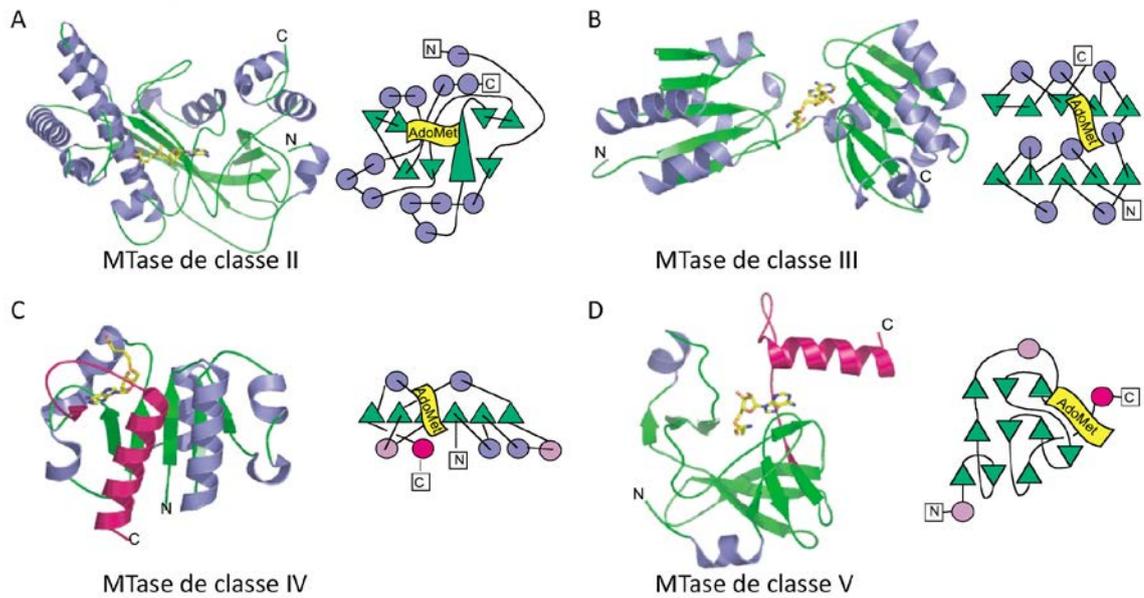


Figure II-4 Structures tridimensionnelles et topologies des méthyltransférases de classe II à V, d'après [218]. Sur les schémas topologiques et les structures, les brins sont en vert et les hélices en bleu. Les nœuds topologiques des classes IV et V sont en rose. La classe III (B) ressemble à deux domaines de classe I apposés. La classe IV (C) est très proche de la classe I en dehors du nœud topologique terminal.

Méthyltransférases à radical SAM

Les enzymes à radical SAM sont une superfamille pour laquelle 113 000 séquences uniques ont été détectées [242], dont 10% au moins seraient des méthyltransférases [243, 244]. Ces méthyltransférases sont regroupées en quatre classes notées A, B, C et D (**Figure II-5**). Elles possèdent toutes le domaine radical-SAM de la superfamille. Elles se différencient par leur domaine supplémentaire (absent de la classe A), les cofacteurs requis, et les mécanismes réactionnels mis en jeu. Ces méthyltransférases produisent un radical 5'-déoxyadénosine à partir de la SAM, pour activer un atome cible en arrachant un hydrogène. Le radical ainsi formé est ensuite méthylé par une seconde molécule de SAM (classe A, B et C) ou par le CH₂-THF (classe D). Ce mécanisme permet de s'affranchir de l'obligation d'avoir un substrat nucléophile des méthyltransférases non radicalaires. Le domaine radical

SAM est un domaine métalloprotéique à centre [4Fe-4S]. Par exemple, la protéine RlmN est une MTRS de classe A qui méthyle les adénines en C2 sur l'ARN ribosomal 23S (m²A₂₅₀₃) [245], et sur les ARNt en position 37 (m²A₃₇) [246] (pour des revues voir [244, 247, 248]).

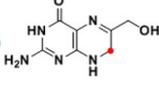
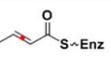
Class A		Class B	
Radical SAM		Cobalamin-BD	Radical SAM
RlmN	rRNA and tRNA modification	Fom3	Fosfomycin biosynthesis
Cfr	rRNA modification	Fms7	Fortimicin A biosynthesis
		GenE and GenK	Gentamicin biosynthesis
		ThnK, L, and P	Thienamycin biosynthesis
		Pacl, N, and O	Pactamycin biosynthesis
		CndI	Chondrochlorens biosynthesis
		Swb9	Quinomycin biosynthesis
		HpnP	2-methylhopanoid biosynthesis
		BchQ, BchR	Chlorophyll biosynthesis
		PoyB and C	Polytheonamide biosynthesis
			
		CloN6	Clorobiocin biosynthesis
		CouN6	Coumermycin biosynthesis
		TsrM	Thiostrepton biosynthesis
		SioT	Siomycin biosynthesis
		BcpD/PhpK	Bialaphos biosynthesis
			
Class C		Class D	
Radical SAM		Radical SAM	CH ₂ H ₄ Folate
	HemN		
NosN	Nosiheptide biosynthesis		
NocN	Nocathiacin biosynthesis		
TpdI	Thiomuracin biosynthesis		
TpdL and U	GE2270 biosynthesis		
Blm-Orf8	Bleomycin biosynthesis	MJ0619	Methanopterin modification
Tlm-Orf11	Tallysomycin biosynthesis		
Zpm-Orf26	Zorbamycin biosynthesis		
YtkT	Yatakemycin biosynthesis		
Jaw5	Jawsamycin biosynthesis		
			

Figure II-5 : Les quatre classes de méthyltransférases à radical SAM d'après [243]. Elles comportent toutes le domaine radical SAM de la superfamille.

II.2 La méthylation des adénines en sur l'azote 1 : m¹A

II.2.1 Les méthylations dans les ARN de transfert

Les méthylations sont les modifications les plus communes dans les ARNt. Elles sont aussi bien endo qu'exocyclique, et peuvent concerner les quatre bases canoniques (A,C,G,U) donnant ainsi les bases méthylées : m¹A, m²A, m⁶A, m⁵U, m³C, m⁵C, m¹G, m²G, m⁷G et m²₂G. Elles peuvent également viser des bases qui portent d'autres modifications et on a par exemple la m¹ψ ou la ms²t⁶A. On peut citer également la méthylation du ribose en 2' rencontrée sur tout type de nucléotide (**Figure II-6**).

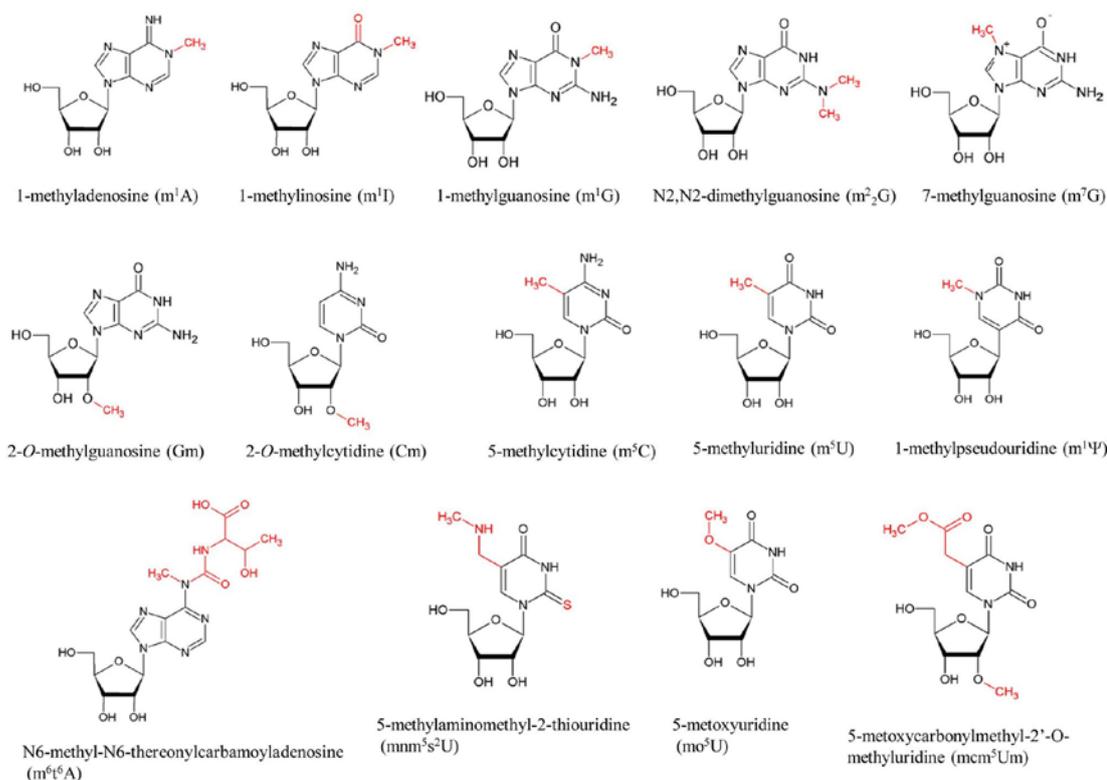


Figure II-6 Exemples de nucléotides méthylés trouvés dans les ARNt [172]. Les sites de modification sont en rouge.

Les méthyltransférases des ARNt sont de trois sortes : la classe I pour la vaste majorité, la classe IV dans une moindre mesure et les MTRS. La découverte de ces

dernières étant relativement récente, on n'en connaît à ce jour qu'une seule pour les ARNt, il s'agit de RlmN responsable de la m²A₃₇ et déjà présentée [246]. Il reste néanmoins de nombreuses méthylation pour lesquelles la transférase n'a pas été identifiée, dont certaines sur des atomes non nucléophiles (en particulier C8 des purines et C2 de l'adénine). L'implication d'une MTRS est alors fort probable [172]. Les années à venir verront donc sûrement la découverte de nombreuses ARNt-MTRS.

II.2.2 Importance de la modification m¹A

La méthylation sur l'azote 1 des adénines (m¹A) est une modification qui a été mise en évidence au début des années 60 [249, 250]. Cette modification est réalisée en une seule étape de méthylation sur une adénine. Cette base modifiée est caractérisée par la présence d'une charge positive portée par l'azote qui est méthylé (**Figure II-7**). C'est l'une des trois bases modifiées portant une charge avec la m⁷G et la m³C.

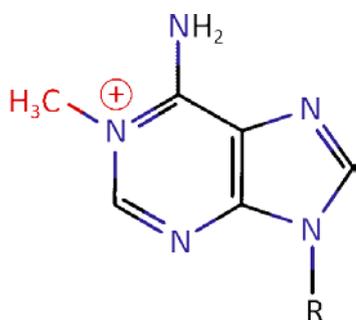


Figure II-7 Structure de la base modifiée m¹A. La méthylation confère à cette base une charge positive.

La m¹A est retrouvée aux positions 9, 14, 22, 57 et 58 des ARNt (**Figure II-8**).

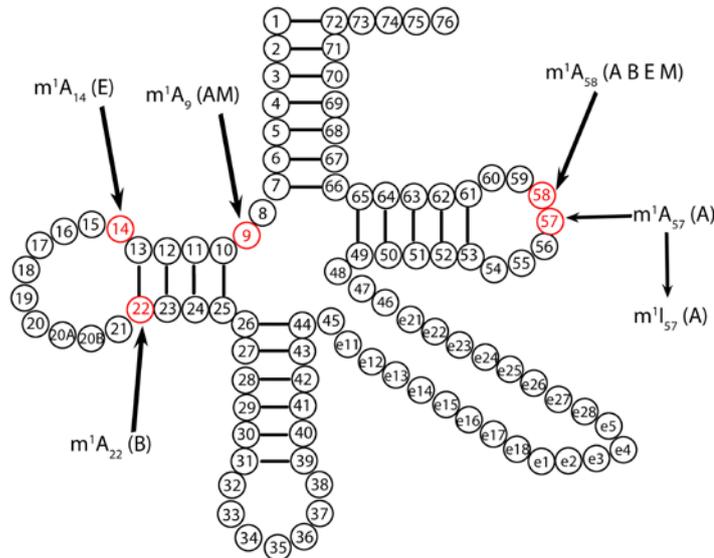


Figure II-8 Position de la m^1A dans les ARNt. Les lettres correspondent aux règnes dans lesquels on retrouve la modification (A archées, E eucaryotes, B bactéries, M mitochondrie). La modification m^1A_{58} est la seule commune à tous les règnes.

Position 9

La modification m^1A_9 est retrouvée dans les mitochondries et une méthyltransférase m^1A_9 a été récemment découverte chez les archées [251]. Les ARNt mitochondriaux ont la particularité d'avoir assez fréquemment des structures secondaires et tertiaires non canoniques ce qui leur confère notamment une température de fusion très basse [211, 252]. Ces particularités font que certaines modifications sont indispensables à leur repliement correct. C'est le cas de la m^1A_9 ; l'ARNt^{Lys} mitochondrial non modifié ne se replie pas en feuille de trèfle, mais forme une tige-boucle étendue. La modification en position 9 empêchant l'appariement A_9-U_{64} est suffisante et indispensable pour former la structure en trèfle de l'ARNt [156] (**Figure II-9**). Elle est également indispensable pour le recrutement du facteur d'élongation Tu par les ARNt n'ayant pas de bras T [253].

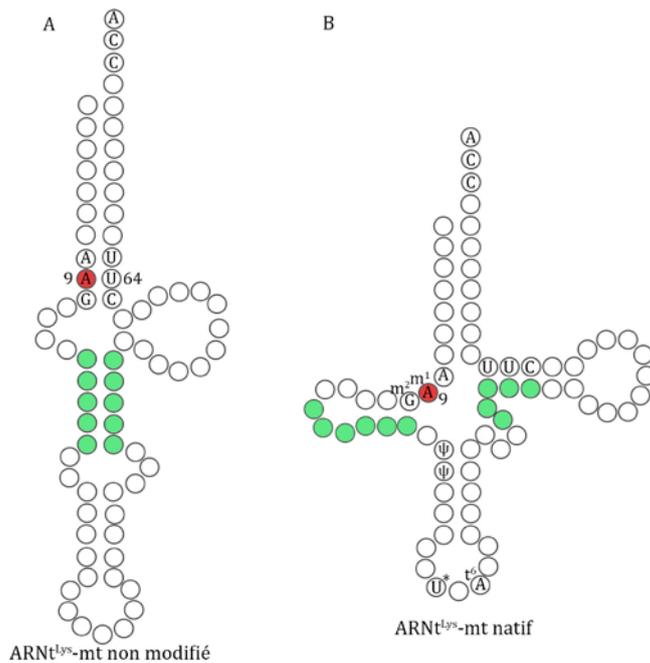


Figure II-9 Structure de l'ARNt^{Lys} mitochondrial (adapté de [156]) A) L'ARNt^{Lys} non modifié adopte une structure en tige-boucle étendue [156]. B) La structure de l'ARNt complètement modifié est la feuille de trèfle canonique. Les nucléotides étendant la tige-boucle acceptrice sont montrés en vert. Les modifications m¹A₉ (en rouge) et m²G₁₀ empêchent les appariements avec les nucléotides 64 et 63 respectivement. La modification m¹A₉ est nécessaire et suffisante pour former la structure en feuille de trèfle [156].

Position 14

La modification en position 14 est rare, et n'a été décrite que pour l'ARNt^{Phe} cytoplasmique de certains mammifères [137, 254]. L'activité m¹A₁₄:ARNt méthyltransférase a été identifiée dans des extraits de cerveau de rat, et l'enzyme partiellement purifiée [255], mais le gène de l'enzyme reste inconnu à ce jour.

Position 57

La modification m¹A₅₇ a pu être détectée chez les archées, mais il ne s'agit que d'un état transitoire. La m¹A₅₇ subit une déamination pour donner la m¹I₅₇ (**Figure II-10**), base qui ne possède plus la charge positive caractéristique de la m¹A [256].

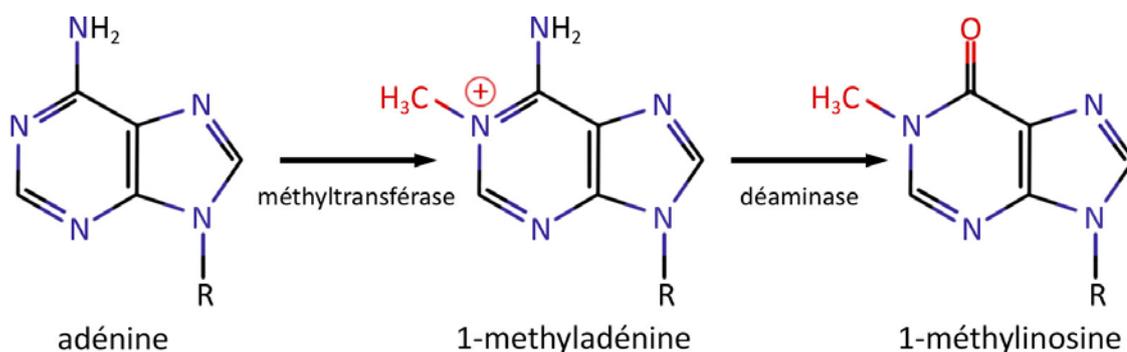


Figure II-10 Formation de la m^1I à partir d'une adénine. L'étape intermédiaire est la m^1A .

Positions 22 et 58

La méthyltransférase de la position 22 est bactérienne uniquement, la méthylation en position 58 est retrouvée dans tous les règnes du vivant. J'ai étudié les méthyltransférases de ces deux positions au cours de ma thèse et elles sont présentées dans les paragraphes II.2.1 et II.2.1.

II.2.1 Position 22 : TrmK

La méthylation en position 22 des ARNt bactériens a été découverte en 1977 [257], sur l'observation que des extraits de *B. subtilis* avaient la capacité de produire une m^1A sur des ARNt eucaryotes portant déjà la m^1A_{58} . La méthylation devait alors avoir lieu sur un autre nucléotide qui s'est révélé être le A_{22} [257].

Ce n'est qu'en 2008 que le gène responsable de cette activité m^1A_{22} a été découvert par nos collaborateurs de l'Université Libre de Bruxelles [258]. Il s'agit du gène *yqfN* dont la protéine YqfN a été nommée TrmK. Les protéines de la famille TrmK appartient à la famille de protéines COG2384 (cluster of ortholog group database) jusque-là non caractérisées, qui sont présentes en particulier, mais pas uniquement, chez les *Firmicutes* avec notamment les pathogènes du genre *Streptococcus*, *Staphylococcus*, *Listeria*...

Fonction

Chez *B. subtilis*, la croissance de souches délétées du gène *yqfN* n'est pas modifiée. Aucun phénotype n'a pu être mis en évidence, que ce soit dans des conditions de croissance normale, de croissance en complétions, ou différence de sporulation [258] Une étude à haut débit de disruption de gènes chez *S. pneumoniae* a, par contre, montré que le gène *sp1610* de TrmK (homologue de *yqfN*) est essentiel à la survie de la bactérie [259].

Une étude de 2007 du phosphoprotéome de *B. subtilis* a identifié la protéine TrmK comme étant phosphorylé sur la sérine 69 (peptide identifié : SGLNpSHISVR) [260]. Par ailleurs des résultats non publiés (Thèse de H. Nguyen chez J.M. Jault [261]) identifieraient la kinase responsable de cette phosphorylation comme YdiB, une protéine essentielle de la famille des shikimate déhydrogénases*, montrant une activité ATPase [262]. Les phosphoprotéines sont relativement rares chez les bactéries. En général, une centaine sont identifiées par organisme, ce chiffre variant en fonction des conditions expérimentales [260, 263–266]. Ces protéines sont souvent au centre de métabolismes clés dont la régulation rapide est indispensable (métabolisme du carbone par exemple), ou de système de transduction du signal [267]. Néanmoins, d'autres études du phosphoprotéome de *B. subtilis* n'ont pas détecté la protéine TrmK comme étant phosphorylée [264–266]. La protéine a été correctement détectée, mais non phosphorylée, dans la dernière étude en date [266]. Ces résultats contradictoires ne permettent pas de tirer de conclusions définitives sur la phosphorylation de TrmK qui pourrait être dépendante des conditions de culture des bactéries, et la question reste ouverte pour l'instant tant sur son existence que sur son rôle.

* Les shikimate déhydrogénases sont des enzymes impliquées dans le métabolisme de l'acide shikimique chez les plantes et les microorganismes qui est, entre autre, un précurseur du tryptophane, de la tyrosine, de la phénylalanine, mais aussi d'alcaloïde, ou encore de l'acide salicylique.

Structure de TrmK

Trois structures de protéines du COG2384 ont été publiées à ce jour, celles de *S. pneumoniae* avec et sans SAM (code d'accèsion PDB3KR9 et 3KU1 respectivement) [268], de *S. agalactiae* (3LEC) et *L. monocytogenes* (3GNL). Les deux dernières ayant été résolues lors de la campagne de génomique structurale Protein Structure Initiative (PSI) ne font pas l'objet de publication. La fonction méthyltransférase de ces trois protéines n'a pas été démontrée. Ces trois structures montrent un domaine méthyltransférase de classe I typique en N_{Ter} (**Figure II-11**), ce domaine étant parfaitement superposable pour les trois homologues. Un domaine C_{Ter} est également présent, il est composé de 4 hélices dont deux très longues formant une structure en "coiled-coil" (superenroulement α). La superposition des structures sur ce domaine est moins parfaite, les deux dernières hélices étant plus ou moins longue, et l'angle entre la structure en coiled-coil et le domaine méthyltransférase est plus ou moins ouvert.

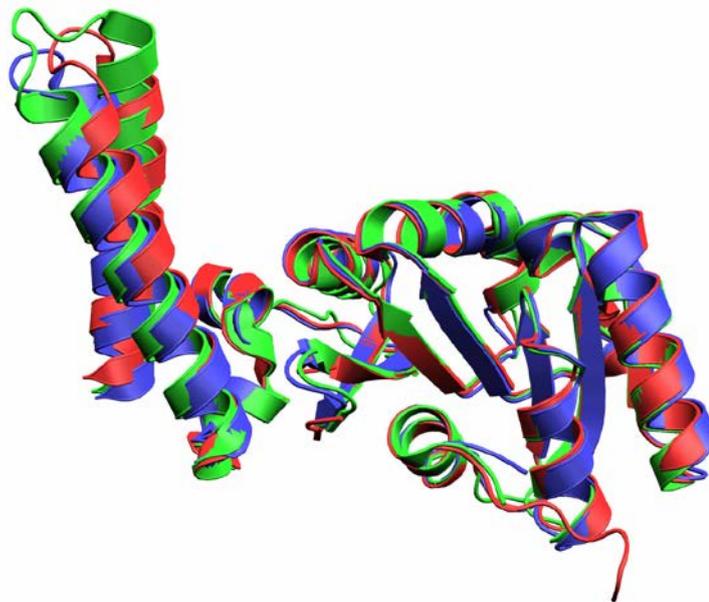


Figure II-11 Superposition des structures de protéines du COG2384 de *S. pneumoniae* sans SAM (en bleu code d'accèsion PDB : 3KR9) [268], de *S. agalactiae* (en rouge 3LEC) et *Listeria monocytogenes* (en vert 3GNL).

L'alignement de séquence de protéines du COG2384 (**Figure II-12**) permet de retrouver les deux domaines. Le domaine C_{Ter} (méthyltransférase) est beaucoup mieux conservé que le domaine N_{Ter}. Les boucles β_1 - α_A et β_4 - α_D correspondant au site actif sont très conservées, ainsi que la boucle β_5 - α_E dans une moindre mesure.

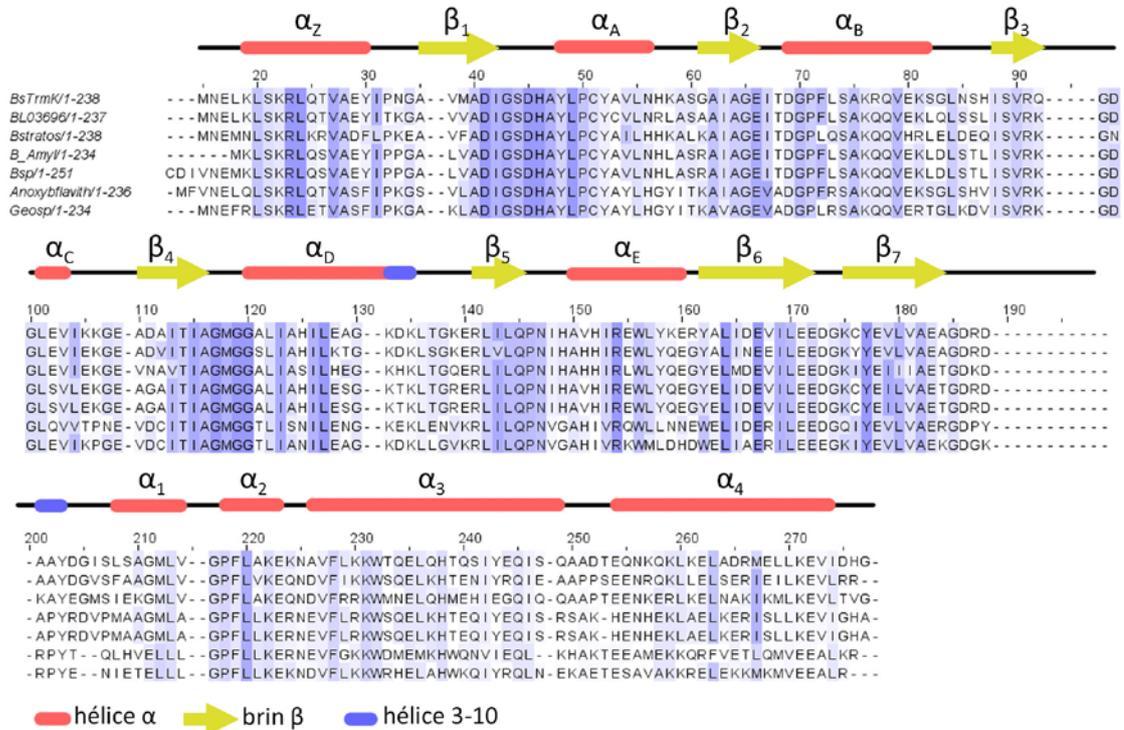


Figure II-12 Alignement de séquences des protéines du COG2384 de 32 organismes (seuls 7 sont montrés). Les éléments de structures secondaires de TrmK de *B. subtilis* sont indiqués au-dessus de la séquence. Le domaine C_{Ter} (méthyltransférase) est beaucoup mieux conservé que le domaine N_{Ter}. Les boucles β_1 - α_A et β_4 - α_D sont particulièrement bien conservées, elles correspondent au site actif de la protéine.

Substrats

Les grandes lignes de spécificité du substrat de la méthyltransférase m¹A₂₂ ont été décrites [269]. La séquence consensus du bras D nécessaire à l'activité m¹A₂₂ méthyltransférase semble être du type A₂₁ A₂₂⁺ G₂₃ G₂₄ C₂₅, mais d'autres séquences

sont possibles. Les ARNt de classe I et II sont modifiés ; et des ARNt dénaturés ou sous forme de duplex ne sont pas substrat de la réaction.

Les spécificités et mode de reconnaissance du substrat ARNt ne sont pas compris. Cette reconnaissance du substrat ARNt par l'enzyme TrmK a été le sujet principal de mon travail de thèse, qui sera présenté dans le Chapitre III.

II.2.2 Position 58

La méthylation m^1A_{58} est commune à tous les règnes du vivant, et est aussi retrouvé dans les mitochondries et les plastes. L'existence de cette modification m^1A_{58} est connue de longue date, et la protéine responsable de cette méthylation avait été purifiée en 1991, mais le gène non identifié [270]. La m^1A_{58} méthyltransférase de *S. cerevisiae* a été identifiée en 1998 [271], il s'agit d'un complexe de deux protéines Gcd10p-Gcd14p (codé par les gènes GCD10 et GCD14), renommées par la suite Trm6p-Trm61p.

La sous-unité Trm61p est responsable de la fixation de la SAM et de la catalyse, et exhibe le repliement des méthyltransférases de classe I. La catalyse nécessite néanmoins la présence de Trm6p, Trm61p ne peut en effet pas fixer seule l'ARNt [271–273].

Des homologues de Trm61p ont par la suite été identifiés chez les procaryotes [274] (**Figure II-13**), et la présence d'activité m^1A_{58} méthyltransférase chez *T. thermophilus* [275], *P. abyssi* [276] et *M. tuberculosis* [277] a permis de valider l'homologie de fonction entre Trm61p et ses homologues procaryotes renommés alors TrmI. La protéine Trm6p n'a par contre pas d'homologues procaryotes, TrmI étant active sous forme d'homotétramère. La présence de deux protéines pour la modification chez les eucaryotes contre une seule chez les procaryotes est un phénomène courant dans les enzymes de modifications [278]. Cette augmentation de la complexité chez les eucaryotes peut être reliée au besoin accru de système de

régulation chez les organismes au métabolisme complexe [278]. Chez les archées, TrmI catalyse aussi la formation de la m¹A₅₇ en plus de la m¹A₅₈ [279, 280].

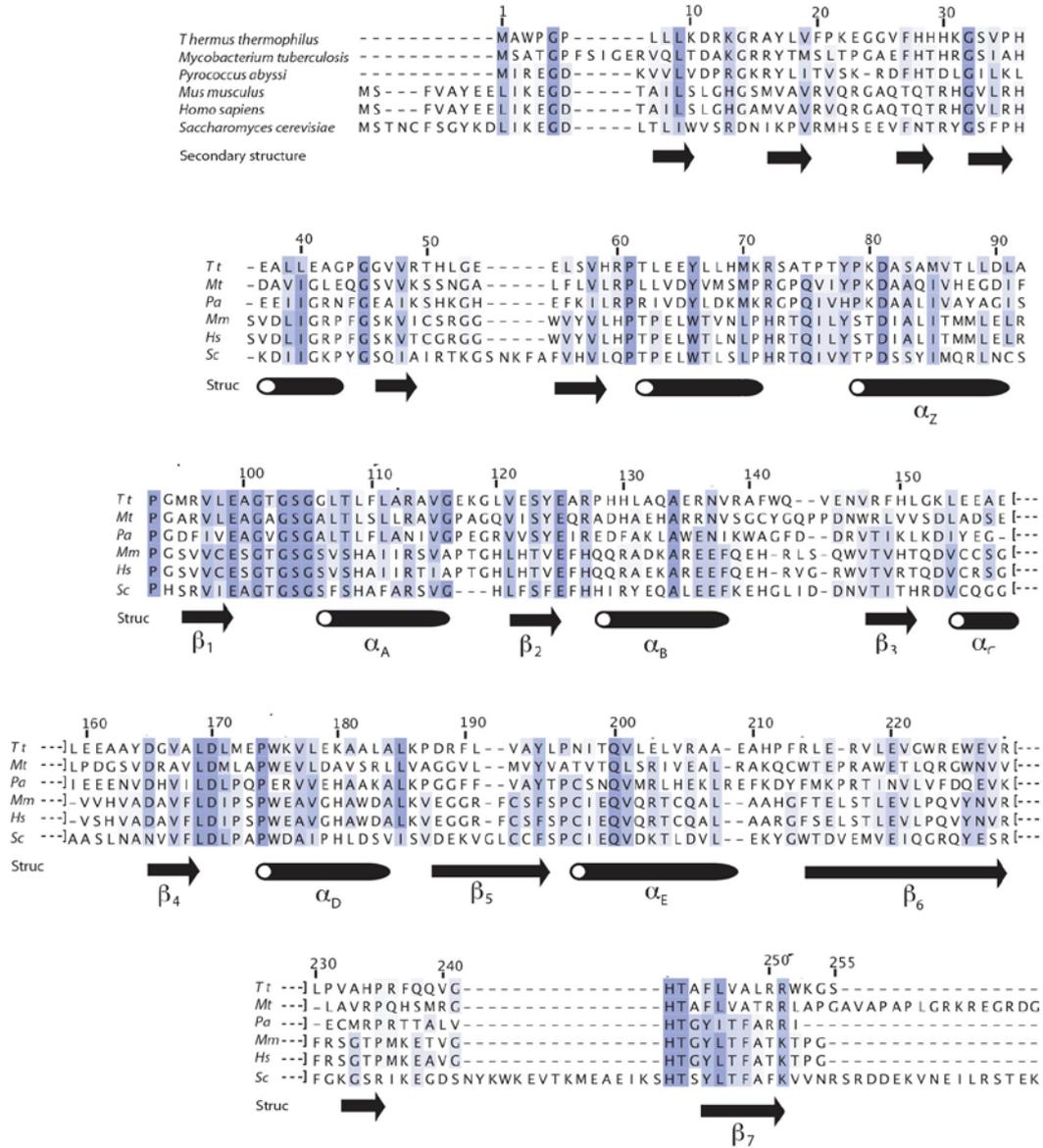


Figure II-13 Alignement de séquence de méthyltransférases m¹A₅₈ (adapté de Barraud et al [281]). Les séquences de TrmI des procaryotes *T. thermophilus*, *M. tuberculosis* et *P. abyssi* sont alignées avec celles de Trm61p de *M. musculus*, *H. sapiens* et *S. cerevisiae*. Les éléments de structure secondaire de TrmI de *T. thermophilus* sont représentés par des flèches pour les brins β et des cylindres pour les hélices α.

Chapitre I

La m¹A dans les ARN de transfert – La méthylation des adénines en sur l'azote 1 : m¹A

Fonctions

Nous avons déjà évoqué la fonction de la m¹A₅₈ dans la stabilisation thermique des ARNt à haute température chez *T. thermophilus* (*Modification : Adaptation aux conditions environnementales en page 47*), et également son importance dans la structuration particulière de la boucle T de l'ARNt_i^{Met} de *S. cerevisiae* (*Modification : Structuration et stabilisation des ARN de transfert en page 42 ; Figure I-27*).

Ces deux cas particuliers montrent que cette modification participe d'une part à la stabilité générale des ARNt en augmentant sensiblement leur résistance à la dénaturation thermique [171, 282], mais a aussi un rôle primordial dans la structuration et le contrôle qualité de l'ARNt_i^{Met} chez les eucaryotes.

Cette modification a également un rôle important à jouer dans la réplication du VIH-1 comme évoqué dans la partie *Implications des ARNt dans les infections virales en page 24*.

Structure de TrmI

Plusieurs structures d'enzymes TrmI ont été résolues [279–281], avec ou sans cofacteurs. La structure de TrmI de *T. thermophilus* en complexe avec la SAH a précédemment été résolue dans l'équipe par Pierre Barraud lors de sa thèse [283]. Les enzymes de la famille TrmI sont organisées en tétramère [284]. Chaque monomère est composé de deux domaines distincts connectés par un linker contenant une hélice (**Figure II-14 a** ; hélice 2). Le domaine C_{ter} adopte le repliement des méthyltransférases de classe I. Le domaine N_{ter} est plus petit et il contient majoritairement des brins β et une courte hélice. Le tétramère est composé d'un dimère de dimères. Les dimères sont formés par l'interaction de deux monomères *via* leurs brins β₆ pour former un feuillet central antiparallèle, des interactions avec les hélices α₂ et α₃ stabilisant ce dimère (**Figure II-14 b**, en bleu foncé et bleu clair). Le tétramère est ensuite formé par l'interaction entre les feuillets β₆ de deux dimères. La SAH est liée au site actif par de nombreux résidus

très conservés dans la famille de TrmI, et est dans la conformation typique des méthyltransférases de classe I (**Figure II-14** c et d).

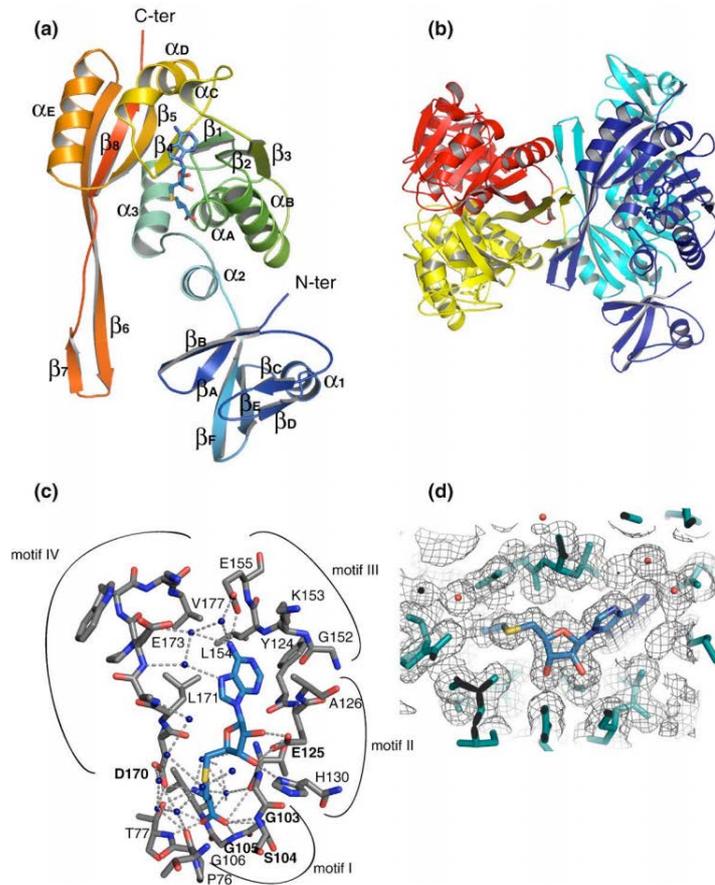


Figure II-14 Structure de TrmI de *T. thermophilus* d'après [281]. (a) monomère de TrmI (b) tétramère formé par deux dimères (c) site actif contenant la SAH, les liaisons hydrogène sont indiquées par des pointillés, les résidus strictement conservés dans la famille sont indiqués en gras (d) densité électronique autour de la SAH.

Les résidus impliqués dans la stabilisation de ce tétramère varient énormément selon les organismes [284]. Les protéines TrmI des organismes thermophiles faisant plus d'interaction ioniques intermoléculaires, et les hyperthermophiles présentant plus d'interactions hydrophobes. De plus chez *P. abyssi* qui est un thermophile extrême (100-103°C) quatre ponts disulfures intermoléculaires maintiennent le tétramère. (**Figure II-15**).

Chapitre I

La m1A dans les ARN de transfert – La méthylation des adénines en sur l'azote 1 : m1A

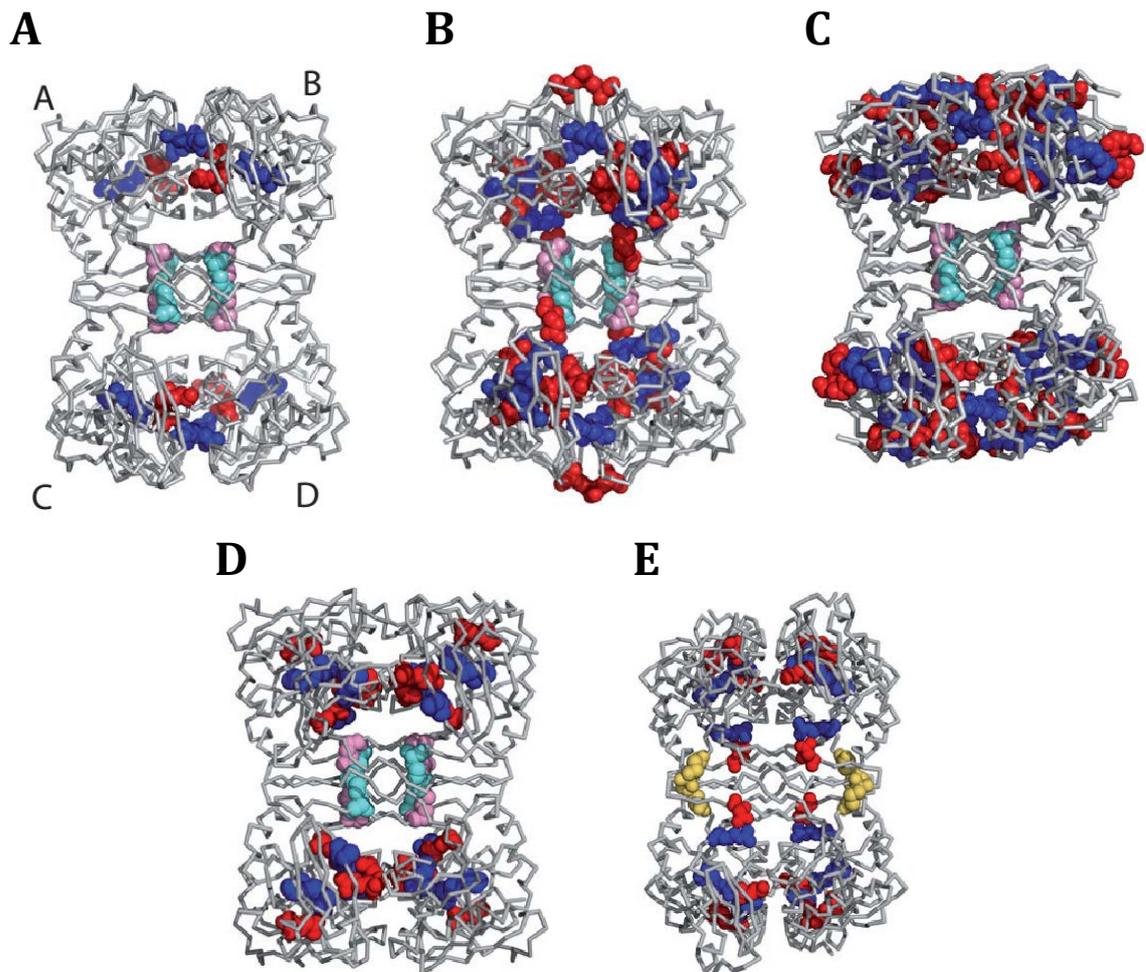


Figure II-15 Organisation tétramérique des protéines TrmI chez différents organismes [284]. Les interfaces dimériques sont situées entre les dimères AB et CD, et les interfaces tétramériques entre les dimères AC et BD. Les résidus impliqués dans des interactions ioniques sont en bleu ou en rouge pour les résidus basiques ou acides respectivement. Le Glu226 et l'Arg230 qui forme le pont salin central de l'interaction tétramérique sont en rose et cyan respectivement. Les ponts disulfures entre les cystéines 196 et 133 sont en jaune. Les organismes sont : (A) *M. tuberculosis* ; (B) *T. thermophilus* ; (C) *T. maritima* ; (D) *A. aeolicus* ; (E) *P. abyssi*.

Substrats

Différentes études [285–287] ont montrées que TrmI peut modifier une tige-boucle contenant uniquement le bras T d'un ARNt substrat de TrmI. Ces résultats impliquent que la reconnaissance principale des ARNt par les méthyltransférases m¹A₅₈ eucaryotes comme procaryotes se fait par le bras T, et que la structure générale de l'ARNt n'est que peu reconnue. Il a par ailleurs été montré dans notre équipe que TrmI reste sous forme tétramérique lorsqu'elle fixe l'ARNt et peut fixer deux molécules d'ARNt [281].

Une étude récente a permis de comprendre de façon détaillée les déterminants de reconnaissance de l'enzyme pour son substrat ARNt [287]. Cette étude a montré que :

- TrmI reconnaît de façon stricte la base modifiée (A₅₈) qui doit absolument être une adénine, et en particulier le groupe 6-amino est reconnu de façon spécifique.
- La position 56 doit être un C, la 57 est une purine et la position 60 une uridine.
- TrmI modifie mieux les ARNt si ceux-ci présentent des modifications préalables.
- Le bras accepteur est reconnu spécifiquement par l'enzyme : un ARNt avec le "bon" bras T mais un bras accepteur d'un ARNt non-substrat est mal modifié.
- La région variable participe à l'interaction.

Ces données de spécificité, les structures cristallographiques et les données de docking de l'adénine dans le site actif [281], apportent des informations sur l'interaction entre TrmI et son substrat. Néanmoins, le mécanisme de la réaction et le mode de reconnaissance entre l'ARNt et son substrat ne sont pas encore bien compris. Si les zones d'interaction peuvent être partiellement visualisées par le potentiel électrostatique de surface de la protéine [281] (**Figure II-16**), il est impossible de modéliser un ARNt en contact avec TrmI sans d'importantes gênes stériques. Une déformation de la protéine et/ou de l'ARNt sont nécessaires pour

cette interaction. Nos recherches en cours sur ce complexe font l'objet du Chapitre II.

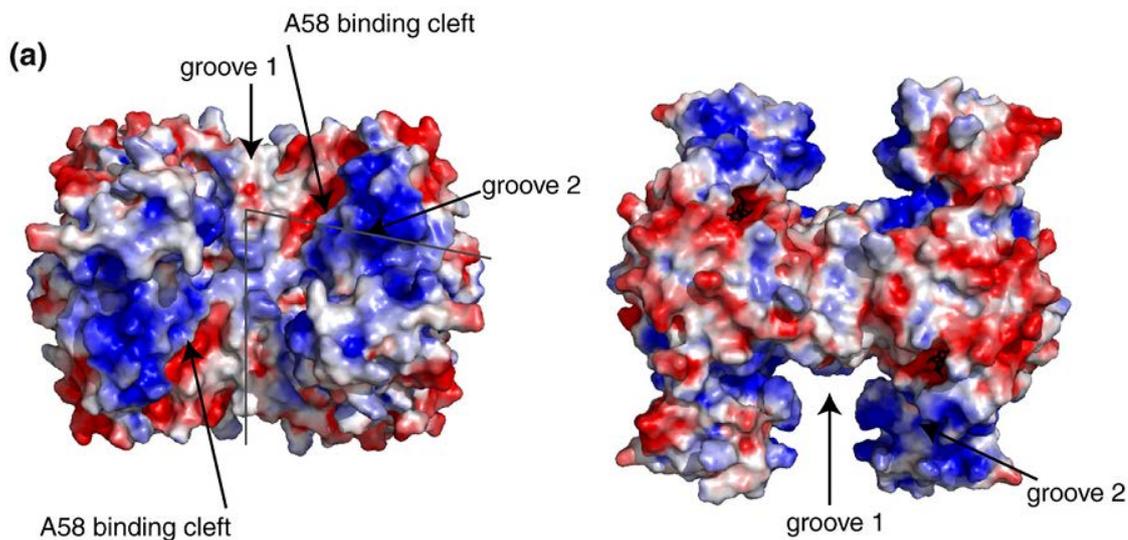


Figure II-16 : Surface électrostatique de TrmI (*T. thermophilus*), adapté de [281]. Le bleu indique une charge positive et le rouge négative. (échelle de -3 kT/e à $+3$ kT/e). Le trait gris indique la position des deux sillons pouvant accommoder la structure en "L" d'un ARNt.

La structure de TrmI de *T. thermophilus* avait permis d'identifier trois résidus du site actif qui sont importants pour la catalyse et la reconnaissance de l'adénine modifiée : D170, Y78 et Y194. En effet, l'aspartate 170 lie à la fois le groupe amino de la SAM (observé dans la structure cristallographique) et l'azote exocyclique N₆ de l'adénosine modifiée (suggéré par le docking de l'adénine dans la poche catalytique). Les tyrosines Y78 et Y194 forment respectivement le fond et le côté de la poche de fixation de l'adénosine. Les trois mutants D170A, Y78A et Y194A conservent une affinité pour le substrat ARNt proche de la protéine sauvage, mais les activités catalytiques (k_{cat} et k_M) sont très affectées [281]. Afin d'appuyer les résultats de cette étude, des études cristallographiques ont été entreprises sur les trois mutants. Les résultats de cette étude sont développés dans le chapitre suivant.

Chapitre I

La m1A dans les ARN de transfert –
La méthylation des adénines en sur l'azote 1 : m1A

Chapitre II Étude de la méthyltransférase m¹A₅₈ des ARNt : TrmI

L'étude des enzymes de modifications des ARNt s'articule autour de deux questionnements principaux : le mécanisme de la catalyse et les déterminants de la reconnaissance du substrat ARNt par l'enzyme.

La cristallographie aux rayons X est une méthode de choix pour répondre à ces questions lorsque la structure du complexe enzyme/substrat peut être résolue. Souvent la connaissance du fonctionnement d'une enzyme permet d'inférer celui d'autres enzymes catalysant la même réaction. Néanmoins, les méthyltransférases de classe I n'ont que très peu d'homologie de séquences, seul le repliement général étant conservé. Les sites catalytiques sont donc très différents d'une enzyme à l'autre, on ne retrouve pas de résidus catalytiques conservés. Ainsi la question du mécanisme catalytique d'une enzyme ne peut pas être trivialement résolue par comparaison avec d'autres méthyltransférases connues. Afin d'apporter des informations sur le mécanisme réactionnel de la méthylation m¹A₅₈ catalysée par TrmI, nous avons résolu la structure cristallographique de mutants de son site actif. Ce travail fait l'objet de la première partie de ce chapitre et a donné lieu à une publication dans le journal *Biophysical chemistry*.

Le mode reconnaissance du substrat pour les enzymes de modifications des ARN est un sujet passionnant mais difficile. Seule une dizaine de complexes entre enzymes de modification des ARNt et substrat ARN a été résolue. Chaque enzyme a un mode

d'interaction avec son substrat qui lui est propre. De plus, elles n'ont pas de domaines spécifiques d'interaction avec les ARN tels que les doigts de zinc, les RRM ou encore le domaine PIWI (pour une revue sur ce type de domaines voir [288]). Ces modes d'interaction particuliers peuvent être en partie expliqués par le besoin pour l'enzyme de déformer son substrat pour atteindre le nucléotide à modifier et l'amener dans son site actif. En effet, la modification implique souvent la rupture d'une ou plusieurs paires de bases, le retournement de la base à modifier, voire la déformation complète du substrat comme cela a été observé pour l'enzyme arcTGT et la forme "λ" de l'ARNt (voir Plasticité structurale en page 7). De l'étude structurale des enzymes de modification ressort néanmoins un trait commun : l'organisation très fréquente en deux domaines, l'un catalytique et l'autre étant vu comme un domaine fixant l'ARN. Ces domaines sont spécifiques à chaque enzyme et n'ont pas d'origine commune. L'étude de complexe entre ARNt et enzyme de modification permet donc d'apporter des éléments nouveaux concernant les modes de reconnaissance ARN-protéine. L'obtention de structure de ce type de complexe est difficile du fait notamment des affinités relativement faibles entre enzyme et substrat. Les complexes sont facilement détruits par la très haute teneur en sels des solutions de cristallisation. Par ailleurs, la grande flexibilité des ARN est un frein considérable à la cristallisation. On utilise donc souvent des fragments d'un ARNt substrat, un des bras de l'ARNt lorsque celui-ci est substrat par exemple. La technique la plus efficace pour maintenir un ARN dans le site actif de l'enzyme est de former un complexe covalent entre cet ARN et l'enzyme. Nous avons fait ce choix pour essayer d'obtenir des cristaux entre l'enzyme TrmI de *T. thermophilus* et une tige-boucle correspondant au bras T de l'ARNt^{ASP} de cet organisme. Ces travaux sont exposés dans la seconde partie de ce chapitre.

Chapitre II

La m1A dans les ARN de transfert – *La méthylation des adénines en sur l'azote 1 : m1A*

III. Mutants du site actif de TrmI

III.1 Problématique et résumé de l'article 1

Cet article décrit la structure cristallographique de deux mutants du site actif de TrmI. Il s'agit des mutants D170A et Y78A qui avaient été précédemment identifiés comme important pour la catalyse [281]. Ces deux résidus sont situés à la jonction entre le site de fixation de la SAM et celui du substrat adénine. Ces structures nous ont permis de discuter le rôle catalytique de D170 et de proposer deux mécanismes réactionnels possibles pour la réaction de transfert de méthyle m^1A_{58} . L'analyse des séquences des bras T des ARNt portant cette modification et des données récentes de la littérature nous a permis de définir une séquence optimale reconnue par TrmI. Enfin, une analyse des modes normaux des structures de ces mutants nous a permis de montrer l'importante mobilité relative du domaine N_{Ter} par rapport au reste de la protéine, ce qui semble être nécessaire à ce domaine pour "attraper" l'ARNt.

Dans cet article, j'ai réalisé l'affinement des structures cristallographiques ainsi que l'analyse des modes normaux de la structure de la protéine sauvage.

The m¹A₅₈ modification in eubacterial tRNA: An overview of tRNA recognition and mechanism of catalysis by TrmI

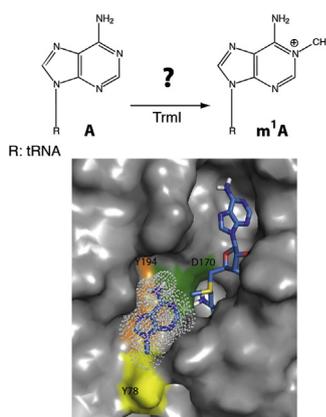
Clément Dégut¹, Luc Ponchon¹, Marcia Folly-Klan, Pierre Barraud, Carine Tisné*

Laboratoire de Cristallographie et RMN biologiques, CNRS, Université Paris Descartes, Paris Sorbonne Cité, 4 avenue de l'Observatoire, 75006, Paris

HIGHLIGHTS

- TrmI D170A mutant has a highly flexible catalytic pocket.
- TrmI Y78A mutant presents the same structure of the catalytic pocket.
- D170 is crucial to grasp the SAM and the adenine in a reactive position.
- Y78A stabilizes the flipped adenosine into the active site of TrmI.
- Two mechanisms of methyl transfer are in agreement with data accumulated to date.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 30 April 2015
Received in revised form 16 June 2015
Accepted 27 June 2015
Available online xxxx

Keywords:

TrmI
m¹A
Catalysis
Methylation
tRNA
Structure

ABSTRACT

The enzymes of the TrmI family catalyze the formation of the m¹A₅₈ modification in tRNA. We previously solved the crystal structure of the *Thermus thermophilus* enzyme and conducted a biophysical study to characterize the interaction between TrmI and tRNA. TrmI enzymes are active as a tetramer and up to two tRNAs can bind to TrmI simultaneously. In this paper, we present the structures of two TrmI mutants (D170A and Y78A). These residues are conserved in the active site of TrmIs and their mutations result in a dramatic alteration of TrmI activity. Both structures of TrmI mutants revealed the flexibility of the N-terminal domain that is probably important to bind tRNA. The structure of TrmI Y78A catalytic domain is unmodified regarding the binding of the SAM co-factor and the conformation of residues potentially interacting with the substrate adenine. This structure reinforces the previously proposed role of Y78, *i.e.* stabilize the conformation of the A₅₈ ribose needed to hold the adenosine in the active site. The structure of the D170A mutant shows a flexible active site with one loop occupying in part the place of the co-factor and the second loop moving at the entrance to the active site. This structure and recent data confirms the central role of D170 residue binding the amino moiety of SAM and the exocyclic amino group of adenine. Possible mechanisms for methyl transfer are then discussed.

© 2015 Elsevier B.V. All rights reserved.

Abbreviations: MTase, Methyltransferase; SAM, S-adenosyl-L-methionine; SAH, S-adenosyl-L-homocysteine.

* Corresponding author.

E-mail address: carine.tisne@parisdescartes.fr (C. Tisné).

¹ These authors contributed equally to this work.

1. Introduction

Transfer RNAs (tRNAs) are crucial biomolecules that play essential roles in many aspects of various cellular processes. Their primary function resides in the translation process, where they constitute adaptor

<http://dx.doi.org/10.1016/j.bpc.2015.06.012>

0301-4622/© 2015 Elsevier B.V. All rights reserved.

Chapitre II

Mutants du site actif de TrmI –
Article 1 : Dégut *et al.* (2015) *Biophys. Chem.*

molecules between messenger RNAs (mRNAs) and elongating peptide chains on active ribosomes.

The biosynthesis and maturation process of tRNAs is a complex and multistep process (for reviews [1–5]). Although synthesis and maturation of tRNA vary significantly among organelles, organisms, and kingdoms [2], many common features have been retained during evolution. It involves the transcription of a tDNA gene followed by processing steps to remove the 5' leader and the 3' trailer, and by the addition of the -CCA amino-acid accepting extremity. When present, introns are also spliced out during the maturation process. In eukaryotic cells, some steps occur within the nucleus and the nucleolus, while other steps occur in the cytoplasm. One of the most interesting and regulated processes of tRNA maturation consists in the introduction of a large variety of modified nucleotides via the enzymatic modification of the initially naked transcript.

About 110 modified nucleotides have been described in the different RNA molecules, of which no less than 91 are present in tRNAs [6,7] (<http://genesilico.pl/trnamodviz>). tRNA is the family of RNA molecules displaying the highest variety of nucleotides modifications. In addition, each tRNA contains a large number of modifications (in average 10 modifications for a typical tRNA length of 76 nucleotides). About one third of these modifications are found in the anticodon stem-loop. They often include hypermodified nucleotides, such as Q₃₄, t⁶A₃₇ or Y₃₇, the biosynthesis of which is a multi-step process. They are commonly thought to play an important role on the ribosome during protein synthesis (accuracy and/or efficiency) [8,9]. Modifications of nucleotides located outside the anticodon region are generally simple ones such as methylations (m¹A₅₈, m²Gm₂₆, m⁵C₄₉...), thiolation (s⁴U₈), reduction (D), and isomerisation (Ψ). Interestingly, some of these simple modifications are catalyzed by specific enzymes acting on the precursor tRNA at early steps of tRNA maturation and before intron splicing, while others are produced as final maturation steps after 5'- and 3'-end processing and intron splicing. Overall, modifications outside the anticodon region have been for a long time believed to be implicated almost exclusively in the stabilization of the three-dimensional architecture of tRNAs. However, the idea that tRNA modifications and their respective pathways can be connected with regulatory functions has recently emerged. Indeed, new roles of tRNA modifications as connectors between translation, metabolism and stress response have been recently uncovered [10–17]. It became clear that tRNA modifications are central to a global quality control of the cell integrity and that tRNA maturation pathways might be placed at the cross-road of various regulation pathways such as the response to oxidative stress or metabolite starvation [1,19–23]. Moreover, variations in the modification level of tRNAs occur in eukaryotes, from organ to organ, in differentiating cells, during aging, starvation, in response to hormones, to certain drugs and to carcinogens, in malignant cells and after infection with viruses (reviewed in [24–27]).

In this context, we have been studying structural aspects of the m¹A modification at position 58 of tRNAs [28,29]. This modification consists in the incorporation of a methyl group on the nitrogen at position 1 of adenines, thus bringing a positive charge on this atom. The m¹A nucleotide is found at positions 9, 14, 22 and 57 (transiently as an intermediate to m¹I) and 58 of tRNA, with a major occurrence at position 58 in the T-loop of tRNA. The m¹A₅₈ is found in tRNAs from the three domains of life and was proved to be an essential modification for different processes. First, this modification is present in most eukaryotic tRNAs, including in three mitochondrial tRNAs. siRNA knockdown of either subunit of the human m¹A₅₈-methyltransferase results in a slow-growth phenotype [30]. Secondly, in *Saccharomyces cerevisiae*, it was shown that mutants that are defective in N1-methylation of A₅₈ are non-viable because this modification is required for the stability of the initiator tRNA^{Met} [31]. Thirdly, this modification is present in all retroviral tRNA primers (including tRNA^{Lys} the primer of HIV-1 reverse transcription), and is crucial for reverse transcription fidelity and efficiency [32]. Finally, m¹A₅₈ was shown to stabilize tRNA structure at

high temperature since gene disruption studies have shown that m¹A₅₈ modification is required for the growth of *Thermus thermophilus* at high temperatures [33]. Interestingly, for this bacterium, it was shown that changes in the structural flexibility of tRNA through multiple nucleoside modifications are used to adapt protein synthesis to environmental changes. The combination of three distinct modifications (Gm₁₈, m⁵s²U₅₄ and m¹A₅₈) found in *T. thermophilus* tRNAs was shown to increase the melting temperature of tRNA by approximately 10 °C as compared with the unmodified transcript [34–37]. Although, it was known that the level of these three modifications in tRNA control the elongation of translation via the flexibility of tRNA since 1987 [38], only recently, the common modifications, m⁷G₄₆ and Ψ₅₅, found in *T. thermophilus* tRNA in addition to these three modifications, were shown to be responsible for this control [37,39]. Indeed, at high temperature (above 65 °C), m⁷G₄₆ functions as a marker of precursor tRNA and increases the reaction rates of other modification enzymes. In contrast at low temperatures, Ψ₅₅ confers local structural rigidity and slows down the rate of formation of other modifications around Ψ₅₅ (Gm₁₈, U₅₄, A₅₈).

The enzymes TrmI that catalyze the formation of m¹A₅₈ at position 58 were described in the early 2000s [33,40,41]. They make use of S-adenosyl-L-methionine (SAM or AdoMet) as a methyl donor. We previously solved the crystal structure of the *T. thermophilus* TrmI and conducted a biophysical study to characterize its interaction with tRNA [28]. TrmI is a methyltransferase of class I, i.e. its catalytic domain is composed of 7 β-strands surrounded by 6 α-helices (Fig. 1A). This domain contains the SAM-binding pocket and is encompassed in the large C-terminal part of TrmI. TrmI is active as a tetramer and the tetramer is organized as a dimer of tight dimers [28,29] (Fig. 1B). The TrmI monomer is composed of two distinct domains connected by an extended linker containing a short α-helix (Fig. 1A). The N-terminal domain is largely a β structure composed of 6 β-strands and one small α helix. Using mass-spectrometry in native conditions, we were able to show that TrmI binds up to two tRNAs per tetramer and remains tetrameric upon binding. We uncovered two grooves that have the appropriate electrostatic potentials and sizes to accommodate one molecule of tRNA per face of tetramer. In addition, we identified three conserved residues (Y78, D170, Y194) important for the methylation reaction and we showed using mutagenesis, kinetic assay and molecular modeling that they fashion an optimal shape of the catalytic pocket in order to position the adenine and the methyl donor in a reactive position.

In this paper, we report the crystal structures of two mutants of the TrmI catalytic pocket (D170A and Y78A). We then discuss tRNA recognition and m¹A catalytic mechanism by TrmIs in the lens of these structures and of newly published data describing tRNA determinants for TrmI recognition [42].

2. Experimental

2.1. Production, purification and crystallization of TrmI mutants (D170A and Y78A)

TrmI D170A and Y78A were expressed and purified as previously described [28]. Protein samples were concentrated at 10 mg/mL in 20 mM Tris-HCl buffer pH 8.0, 100 mM KCl with 2 mM SAM. Crystallization was performed at 18 °C by vapor diffusion using the hanging drop method. The conditions of crystallization of the wild-type TrmI were used as a starting point (2.0 M ammonium sulfate, 5% isopropanol (v/v)).

Crystals of TrmI D170A were obtained in 2.4 M ammonium sulfate and 10% isopropanol (v/v). Crystals of TrmI Y78A were obtained in 2.1 M ammonium sulfate and 8% isopropanol (v/v). The crystals were harvested, soaked in a cryoprotectant solution containing 20% glycerol and flash-frozen in liquid nitrogen before data collection.

TrmI D170A diffraction data were collected at beamline PX1 of the Synchrotron SOLEIL (Gif sur Yvette, France) and TrmI Y78A diffraction data were collected at beamline ID23-2 of the European Synchrotron Radiation Facility (Grenoble, France).

Chapitre II

Mutants du site actif de TrmI –

Article 1 : Dégut et al. (2015) Biophys. Chem.

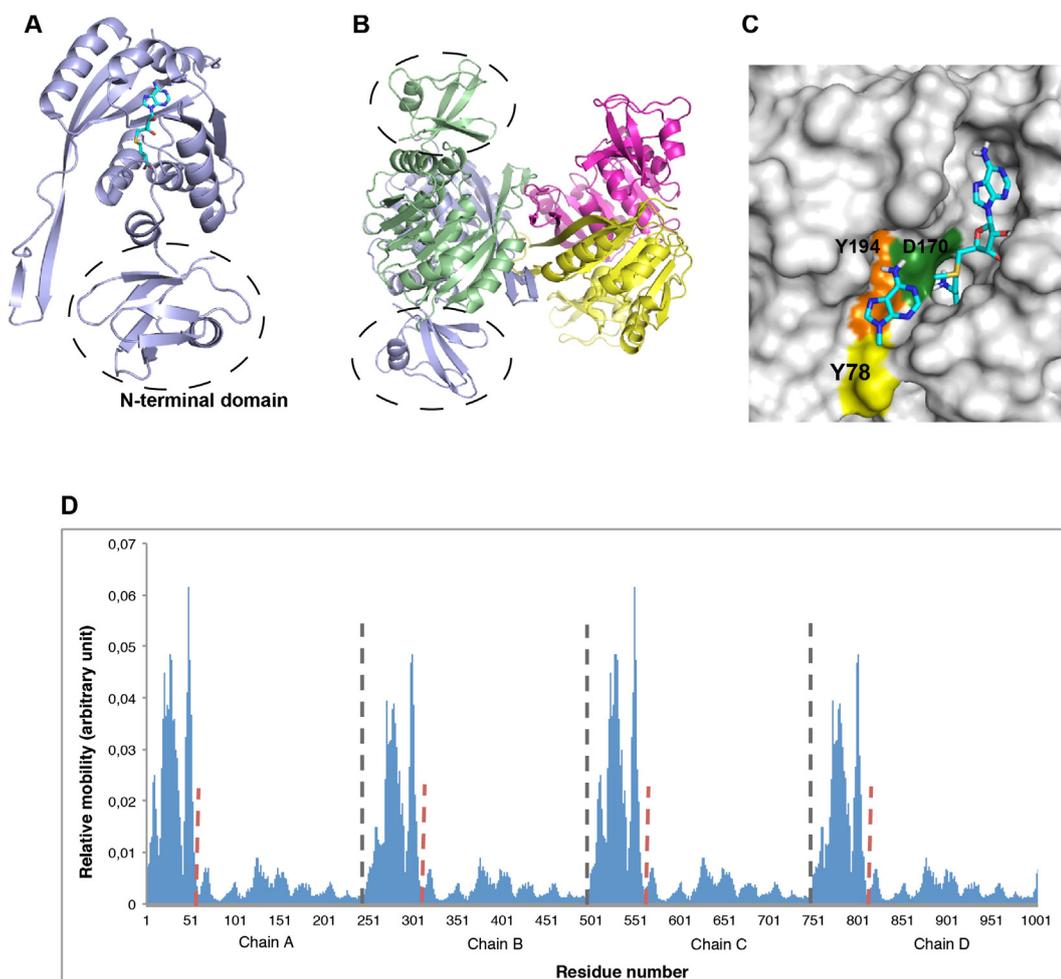


Fig. 1. *T. thermophilus* TrmI structure (PDB code 2PWY): location of catalytic mutants and dynamics of the protein. (A) Structure of the *T. thermophilus* TrmI monomer showing the fold of the N-terminal domain composed of 6 β -strands and one α -helix. The N-terminal domain is surrounded in dotted line. (B) Architecture of the TrmI tetramer showing the N-terminal domains surrounded in dotted line. (C) Modeling of an adenine in the catalytic pocket of TrmI, average docked structure of the adenine ring inside the catalytic pocket of TrmI previously published [28]. The enzyme surface is shown in gray. The AdoMet ligand and the docked adenine ring are shown as sticks (carbon atoms in light blue, nitrogen atoms in dark blue, oxygen atoms in red, sulphur atom in yellow and hydrogen atoms in white). The Van der Waals surface of the adenine ring is shown as white dots. The positions of the key residues are highlighted in color on the surface (D170 in green, Y194 in orange and Y78 in yellow). (D) Normal mode analysis of TrmI showing the mobility of the N-terminal domain. The mean mobility of residues in the six first non-trivial computed normal modes is plotted as a function of the residue number for the four molecules in the TrmI tetramer. The dotted lines in gray indicate the first residue of each monomer whereas the red dotted lines delineate the end of the N-terminal domain of each monomer. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Crystallization assays of TrmI Y194A led to poorly diffracting crystals.

2.2. Crystal structure determination

X-ray data were processed with MOSFLM and scaled with SCALA of the CCP4 suite [43]. The structures of the mutants were solved by molecular replacement using the program Phaser [44] and the structure of the wild-type TrmI (PDB code 2PWY). Restrained refinements of the structures were performed with the program phenix.refine in the Phenix suite [45]. Model and map visualizations for manual reconstructions were performed with the program Coot [46].

Structures have been deposited to the Protein Data Bank under the codes 5C11 and 5C00, for the D170A and Y78A mutants, respectively. Analysis of the SAM binding site in TrmI Y78A mutant was performed using the lidia extension of Coot [47].

2.3. Normal mode analysis

Normal mode analysis for the wild-type TrmI structure (PDB code 2PWY) was conducted with the program ProDy [48] and visualized

using VMD [49]. To analyze the mobility of this domain, the mean relative mobility of residues in the six first non-trivial computed normal modes was calculated and plotted as a function of the residue number for the four molecules in the TrmI tetramer.

2.4. *T. thermophilus* tRNA sequence analysis

The available sequences of tRNA from *T. thermophilus* of the web site (<http://genesilico.pl/trnamodviz>) were used to discuss the preferred sequence of the T-loop of *T. thermophilus* TrmI (see paragraph 3.2).

3. Results and discussion

3.1. Structures of TrmI catalytic mutants (D170A, Y78A)

In a previous study, we solved the crystal structure of *T. thermophilus* TrmI [28] and identified three conserved residues D170, Y78 and Y194 in the active site of TrmI important for the methyl transfer. D170 forms the back of the catalytic pocket (Fig. 1C) whereas Y78 is located at the floor of the adenosine-binding pocket (Fig. 1C). By mutagenesis and steady-state kinetic assays, we showed that the D170A and Y78A

Chapitre II

variants of TrmI are severely altered in their catalytic activity. Indeed, the D170A mutant catalytic turnover was severely reduced by a factor 300 compared to the wild-type TrmI, whereas that of the Y78A variant experienced a 20-fold decrease [28]. Using molecular docking, we showed that the D170 residue could act as a bidentate ligand binding the amino terminal moiety of SAM and the N₆ amino group of the adenine ring substrate.

To investigate at the atomic level the impact of the mutations on the structure of the catalytic pocket of TrmI, we solved the crystal structures of these two mutants. Each mutant crystallized in a different space group compared to wild-type TrmI (Table 1) although the crystallization conditions to get crystal are almost identical. The structure of TrmI D170A was solved and refined to 3.1 Å to a final R_{free} factor of 27% whereas the TrmI Y78A model was refined to 2.6 Å to a final R_{free} of 28% (Table 1). In both structures, four molecules of protein are present in the asymmetric unit. Globally, the overall structures of TrmI D170A and of TrmI Y78A mutants closely resemble that of the wild-type TrmI (Fig. 1A and B). Indeed, the tetramer architecture and the fold of the N and C-terminal domains are conserved. The major differences in structures lie in the definition of the N-terminal domain in the density map and in the structures of the catalytic pocket due to the mutation. First, the N-terminal domain is folded but is not very well defined in the density map, its B-factors are higher than the C-terminal part in both mutant crystals. On the contrary, it was very well defined and not very flexible in the wild-type structure due to packing contacts in the crystal. These differences highlight the mobility and the dynamics of this domain in the tetramer architecture. Interestingly, the normal mode analysis conducted on wild-type TrmI (Fig. 1D) confirms this feature and shows that the N-terminal domain exhibits the highest relative mobility compared to the C-terminal domain that contains the catalytic pocket and the tetramer interfaces. Moreover, the mean B-factor of the C-terminal domain in wild-type TrmI was 9 Å² whereas it reaches 25 Å² in the N-terminal domain. The N-terminal

domain forms one groove of the tetramer and was proposed to be in contact with the variable loop or the D/anticodon stems of the tRNA [28,50]. A recent paper shows that an RNA containing the 3'-part of the acceptor stem, the T arm and the variable loop is a very good substrate of TrmI whereas one containing the D/T/anticodons arms and the variable loop was a poor substrate [42]. Therefore, the N-terminal domain most probably grasp the variable loop of substrate tRNAs and its mobility is probably essential in this interaction to keep the adenine in the catalytic pocket.

Secondly, the comparison of the structures of these mutants to the wild-type is more informative when one considers the structure of the catalytic pocket (Figs. 2 and 3). TrmI Y78A was successfully co-crystallized with SAM ligand. Fig. 2A shows the superimposition of the Y78A structure with that of the wild-type at the level of the catalytic center. The structures are closely related and the conformation of the

Table 1

Data collection and refinement statistics for the determination of TrmI structure. Values in parenthesis are for the outer resolution shell.

	TrmI D170A	TrmI Y78A
A. Data collection statistics		
Beamline	PX1 (SOLEIL)	ID23-2 (ESRF)
Space group	P6 ₁ 22	P 2 ₁ 2 ₁ 2 ₁
Unit-cell parameters (Å)		
a b c (Å)	110.4 110.4 306.4	76.5 79.9 184.0
α β γ (°)	90 90 120	90 90 90
No. of measured reflections	41,449 (4054)	66,077 (6377)
No. of unique reflections	20,952 (2035)	34,672 (3339)
Multiplicity	2.0 (2.0)	1.9 (1.9)
Resolution limits (Å)	51.94–3.10 (3.21–3.10)	52.92–2.62 (2.71–2.62)
Data completeness (%)	99.95 (100.0)	99.88 (99.88)
Mean I/σ(I)	9.6 (2.2)	10.78 (2.68)
Wilson B-factor	84.05	28.35
R _{merge} (%)	5.6 (39.3)	4.7 (25.9)
B. Refinement statistics		
R-factor (%)	20.0 (24.2)	23.5 (24.0)
R _{free} (%)	26.9 (33.4)	28.2 (30.0)
Number of non-hydrogen atoms	6987	7264
Protein	6950	7127
Ligands	15	5
Waters	22	132
Average B-factors (Å ²)		
Macromolecules	74.0	41.2
Solvent	61.2	36.8
RMSD (bonds) (Å)	0.017	0.005
RMSD (angles) (°)	1.78	1.04
Ramachandran favored (%)	93.0	94.0
Ramachandran allowed (%)	5.1	3.1
Ramachandran outliers (%)	1.9	0.9

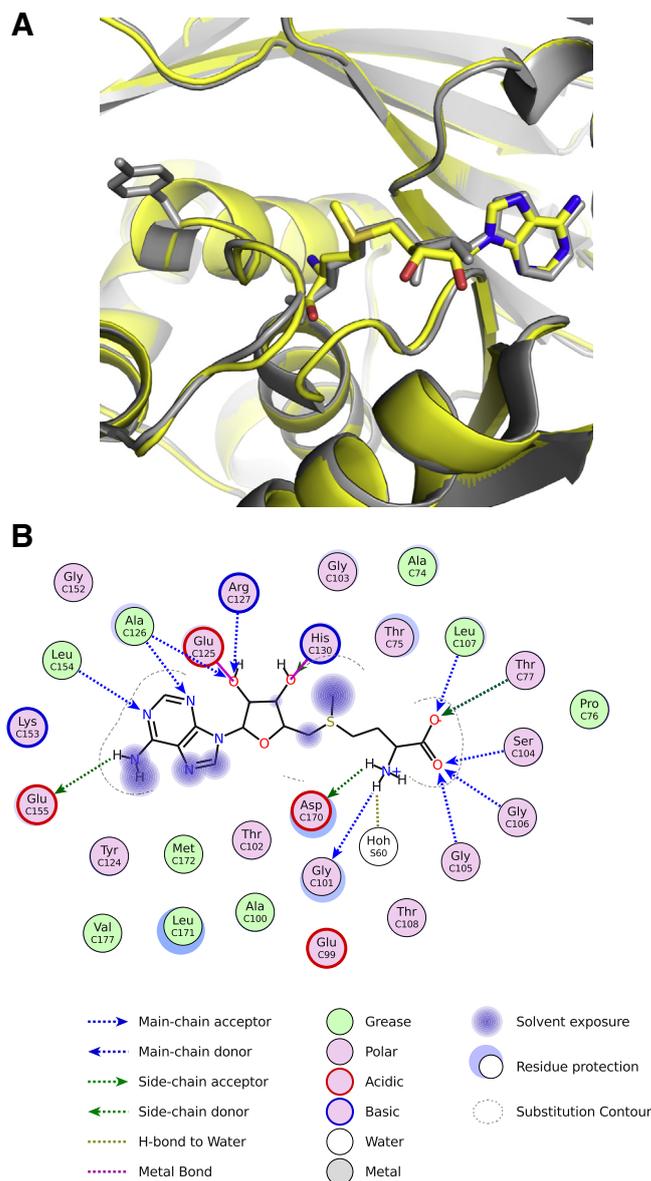


Fig. 2. Structure of the catalytic pocket of TrmI Y78A mutant. A) Superimposition of the main chain of the wild-type protein (in gray, PDB code 2PWY) with that of the Y78A mutant (in yellow). The tyrosine Y78 and the SAM/SAH are represented as sticks. B) Analysis of the SAM binding site in TrmI Y78A mutant. The symbols are those used by the lidia extension of Coot [47]. Hydrogen bonds are indicated by arrows whereas hydrophobic interactions are represented by circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Chapitre II

Mutants du site actif de TrmI – Article 1 : Dégut et al. (2015) Biophys. Chem.

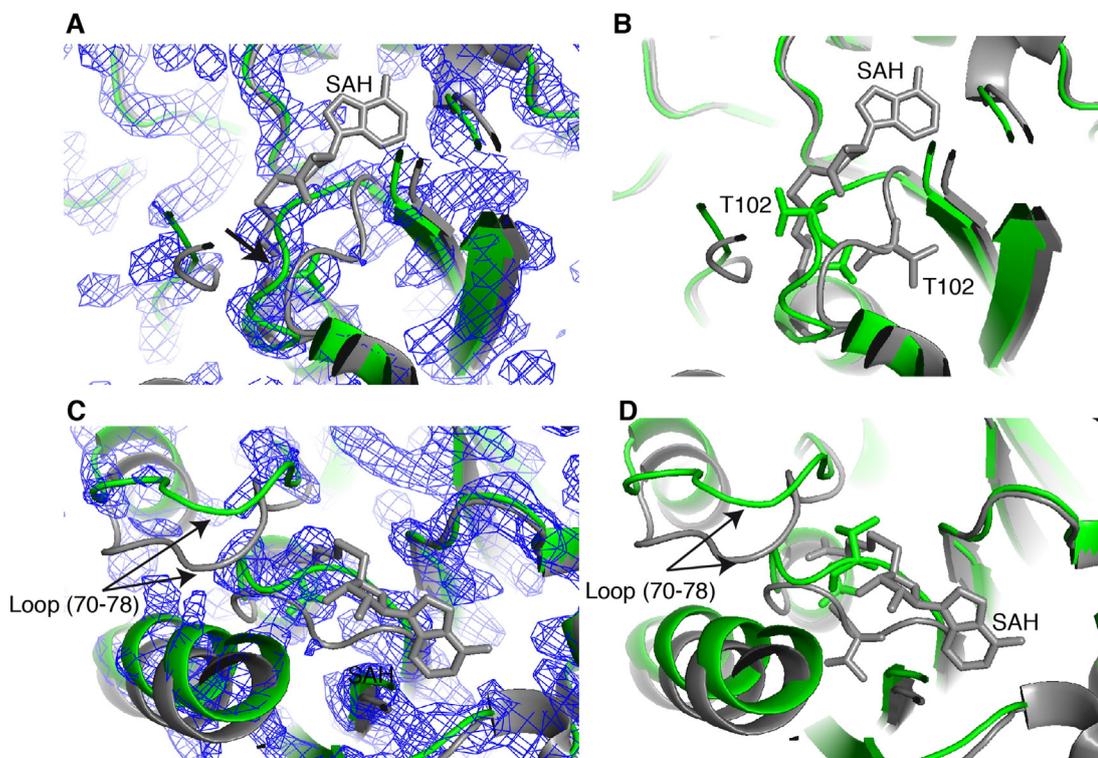


Fig. 3. Structure of the catalytic pocket of TrmI D170A mutant (in green). Four molecules are present in the asymmetric unit of the TrmI D170A crystal. Chains B and D are the domains for which the loop 101–105 containing T102 is the best defined in the density map ($2F_0 - F_C$ at 1.5σ) of the TrmI D170A mutant (in blue, A). Chain B (in green) was used for A and B. T102 is shown as sticks in B. Chains A and C are the domains for which the loop 70–78 is the best defined in the density map ($2F_0 - F_C$ at 1.5σ) of the TrmI D170A mutant (in blue, C). Chain C (in green) was used in C and D. The wild-type TrmI (PDB code 2PWY) in complex with SAH is represented in green. The black arrows indicate the loops that exhibit large movement in the D170A mutant compared to wild-type TrmI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ligand is the same in both structures. The network of residues that make contacts with SAM in the structure of Y78A, either by hydrogen bonds or hydrophobic contacts, is indicated in Fig. 2B. This network is the same than that identified for the wild-type protein [28]. Notably, the carboxylate chain of D170 is involved in a hydrogen bond with the ammonium group of SAM and the residue at position 78 does not participate to the binding of the co-factor. Therefore, the mutation Y78A does not perturb the organization of the catalytic pocket around the SAM binding pocket. The only and main differences lies in the hole introduced by the mutation at the floor of the catalytic pocket due to the absence of the aromatic ring. This residue is conserved as an aromatic (F or Y) among the family of TrmI enzymes. This residue cannot be involved in the methyl transfer due to its position in the active site. The structure of the Y78A mutant thus reinforced our hypothesis that Y78 is, most probably, essential to stabilize the flipped adenine in the catalytic center of TrmI, by stacking interaction with the ribose of the adenosine 58. Y78 is a good candidate for the stabilization of the flipped-out adenine given its position in the catalytic pocket, the structure and the dramatic decrease of k_{cat} of the TrmI Y78A mutant.

D170 is a crucial residue, universally conserved, that makes hydrogen bonds with both the SAM co-factor and the adenine to be methylated [28,42]. We could not obtain co-crystal of the D170A mutant with either SAM or SAH whereas SAM was present in the catalytic pocket in the crystal structure of Y78A mutant and SAH in the wild-type protein structure. This feature indicates that, among the numerous contacts between the co-factor and TrmI (Fig. 2B), the hydrogen bond between D170 and the co-factor is essential to maintain the co-factor in the catalytic pocket of TrmI. Interestingly, when the co-factor is not bound to TrmI, the active site of the protein is occupied by two loops, namely the loop made of residues 70–78 (Fig. 3A, B) and the second one made of residues 99–105 with a pivotal motion at glycine residues 101 and 105 (Fig. 3A, B). Notably, this motion positions T102 in the binding

site of the co-factor near the mutation (Fig. 3B). The loop 70–78, that is positioned at the entrance of the TrmI active site, is very flexible and thus not very well defined (Fig. 3C). However, it is clear that the orientation of this loop is different from that in the wild-type structure. Interestingly, the orientation of the loop 70–78 is similar in the TrmI D170A mutant and in the structure of *Thermogota maritima* TrmI that was crystallized without any co-factor (PDB code 1O54). In contrast, in the *T. maritima* TrmI structure, the loop 99–105 is positioned like in the wild-type *T. thermophilus* TrmI. Therefore, in the TrmI D170A structure, the positioning of the loop 70–78 in the active site most probably reflects the absence of the co-factor, whereas the location of the loop 99–105 is a consequence of the mutation itself. As a conclusion, the active site of TrmI, in absence of co-factor, is occupied by the loop 70–78 that probably prevents the inappropriate binding of a substrate tRNA prior to the accommodation of the co-factor. Moreover, D170 is an essential residue to maintain an operational active site and to position both the co-factor and the substrate adenine.

3.2. Recognition of tRNA substrate

It was recently shown that the activity of *T. thermophilus* TrmI is dependent on precise sequence determinants of its tRNA substrate [42] and that the sequence of tRNA will impact the m^1A_{58} activity of TrmI. For instance, TrmI from *T. thermophilus* preferentially methylates *T. thermophilus* tRNA^{Phe} than *S. cerevisiae* tRNA^{Phe}. It was also shown that the access of TrmI to A_{58} requires disruption of the interaction between T and D loops, and indeed, disruption of T and D loop interactions accelerates the methylation [42]. In Fig. 4, we have analyzed the T-arm sequence of *T. thermophilus* tRNAs and summarized on these sequences the determinants of TrmI recognition that are known to date. For instance, the T-stem structure is essential for TrmI recognition [42] and notably, the two first base pairs closing the T-loop are conserved

Chapitre II

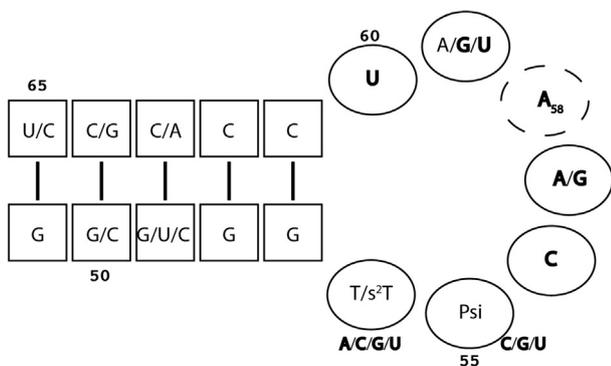


Fig. 4. Sequences of tRNA T-arms that are substrate of TrmI. The sequences in square and ellipse are the sequence of *T. thermophilus* tRNAs. The letters in bold indicate nucleotides that were mutated in the study of Takuma et al. [42] and for which the tRNA transcripts were efficiently methylated at position 58 by TrmI. Nucleotides in the T-stem are indicated in square whereas nucleotides in T-loop are in ellipse. The modified nucleotide A₅₈ is in a dotted ellipse. A₅₈ cannot be mutated into an inosine nor a 2-aminopurine [42].

as G–C base pairs in *T. thermophilus* tRNA (Fig. 4). Point mutations revealed that C56 and U60 are involved in recognition of the substrate tRNA by TrmI [42]. Indeed, these two nucleotides are conserved in *T. thermophilus* tRNA. Moreover, nucleotide at position 57 should be a purine. The position 55 does not tolerate an adenine whereas the position 59 does not tolerate a cytosine. In contrast, the nucleotide 54 can be mutated in any nucleotide without altering TrmI activity. Therefore, the U₅₄–A₅₈ reverse Hoogsteen base pair is not required for TrmI recognition. However, this nucleotide is conserved as a T or s²T in *T. thermophilus* for insuring thermal stability of tRNA. Last but not least, the replacement of A₅₈ by a 2-aminopurine or an inosine abolishes the methylation reaction at position 1, therefore confirming our model showing that the exocyclic amino group at position 6 of A₅₈ is recognized by TrmI through a hydrogen bond with D170.

These determinants of TrmI recognition to efficiently recognize and modify a tRNA substrate are crucial for the rapid response to environmental changes of temperature that *T. thermophilus* has to face.

3.3. Mechanism of m¹A methyl transfer

We pursued the structural study of TrmI protein variants in order to better characterize the methyl transfer mechanism in tRNA m¹A MTase. The structure of TrmI Y78A mutant suggests that Y78 cannot be involved in methyl transfer but is, most probably, involved in the stabilization of the adenosine in the active site of TrmI. The universally conserved D170 was previously identified as a pivotal residue for the methyl transfer reaction [28]. However, we could not resolve whether D170 acts as a general base catalyst or whether D170 only acts as a bidentate ligand that properly positions the two substrates of TrmI in a proper reactive orientation [28]. The structure of D170A mutant shows the crucial role of this residue to organize the active site of TrmI, in positioning or in maintaining the SAM-cofactor and in leaving the active site accessible to tRNA substrate (Fig. 3). We propose here to discuss in detail the possible chemical mechanisms of methyl transfer by SAM-dependent MTases at the N₁ position of adenines in RNAs. Overall, it is well admitted that SAM-dependent N-methylation reactions proceed via a direct S_N2-like transfer, in which the attacking nucleophile and the methyl donor are involved in a single transition state structure [51,52]. Importantly, the thermodynamic of methyl transfer by SAM is extremely favorable and SAM-dependent methylation reactions are known to be irreversible. Indeed, for instance, the free energy of methionine formation via methyl transfer of SAM to homocysteine is more than twice the free energy of ATP hydrolysis (i.e. 71 kJ·mol⁻¹ vs. 30 kJ·mol⁻¹) [53]. For that reason, the catalytic efficiency of SAM-dependent MTases is predominantly linked to their ability to bring together the methyl-donor and the methyl-acceptor in the proper

orientation. In addition to the proximity and orientation requirements, the enzyme could be involved in the stabilization of the transition state of the reaction, three parameters that overall increase the rate of methyl-transfer [52]. Even though the mechanism of m¹A methylation has not been thoroughly investigated on its own, some information can be inferred from the study of the N-methylation mechanism of other compounds such as nicotinamide [54] or intermediate products of the caffeine pathway [55,56]. In these methylation reactions, the endocyclic nitrogens are sometimes initially protonated (e.g. in the intermediate products of the caffeine pathway), or free of proton (e.g. in nicotinamide). The mechanism is therefore quite different when a proton is present, since it must be removed before or during the methyl transfer step. This proton removal, if present, is absolutely required and increases the nucleophilicity of the nitrogen for a proper methyl transfer following an S_N2-like mechanism [51,52,57,58]. In the case of the N₁ position in adenines, there is no proton attached to the endocyclic nitrogen, meaning the methyl transfer could occur without a preliminary deprotonation. However, in the field of RNA N-methylation, the best-studied mechanisms are for the N₁ of guanines [59–61] and the N₆ of adenines [62,63], two cases where a preliminary deprotonation is needed to reveal the nucleophilicity of the corresponding centers. Similarly, but without experimental and/or computational arguments, it was proposed that the N₁-methylation of adenines would proceed via the deprotonation of the N₆ exocyclic nitrogen [58]. Although the need for a deprotonation is still in question, we had previously proposed that, if present at all, the deprotonation could be performed by D170 [28]. In that case, D170 would act as a base catalyst and the methyl transfer would lead to the deprotonated imino tautomer form of the m¹A nucleotide (Fig. 5A, 1st step). As we discussed earlier [28], since the deprotonation of N₆ of m¹A is easier than the deprotonation of the N₆ of the non-methylated adenine (pK_a = 9.3 for m¹A vs. pK_a = 16.7 for A) [64,65], the deprotonation must occur in the same elementary step as the methyl transfer. The formation of the acidic form of m¹A, which is the one stable at physiological pH (pK_a = 9.3), requires then the protonation of the exocyclic imino group. The proton donor could be either D170 or a solvent molecule (Fig. 5A, 2nd step). However, it is also perfectly possible that the N₁-methylation proceeds without deprotonation, in a single elementary step following an S_N2-like mechanism (Fig. 5B). In this mechanism, there is no intermediate and the methyl transfer gives directly the acidic form of the m¹A nucleotide. In this case, the enzyme would provide the appropriate proximity and orientation to the methyl donor and acceptor. In addition, since in this mechanism the transition state is obviously cationic, the enzyme could also be involved in the stabilization of the cationic transition state. We proposed earlier that Y194 could serve that role [28].

The most interesting point regarding the m¹A mechanism comes from the fact that the D170A variant is very affected in its catalytic properties, but is still active (its k_{cat} is about 300 times lower than the one of the wild-type enzyme). In this variant, the potential basic catalyst is not present anymore, and therefore mechanism (A) presented on Fig. 5A is no longer possible. The methyl transfer must therefore occur via a mechanism similar to mechanism (B) presented on Fig. 5B. The mechanism occurring in the D170A variant is therefore 300 times less efficient than the one occurring in the wild-type TrmI enzyme. But we cannot conclude whether the mechanism (A) or the mechanism (B) is the one occurring in the wild-type enzyme. Indeed, both of them perfectly agree with our previous kinetic results [28], and the recent literature [42]. The difference in activity between the wild-type enzyme and the D170A variant could mean two things: (i) the mechanism occurring in the wild-type enzyme is mechanism (A) and the difference in activity is due to a default in the deprotonation of the substrate which prevent proceeding through the imino tautomer form of m¹A; (ii) the mechanism occurring in the wild-type enzyme is mechanism (B) and the difference in activity is due to an inappropriate orientation of the two substrate for efficient methyl transfer. The fact that modified adenine rings, e.g. hypoxanthine or 2-aminopurine, at position 58 of tRNAs are

Chapitre II

Mutants du site actif de TrmI –

Article 1 : Dégut et al. (2015) Biophys. Chem.

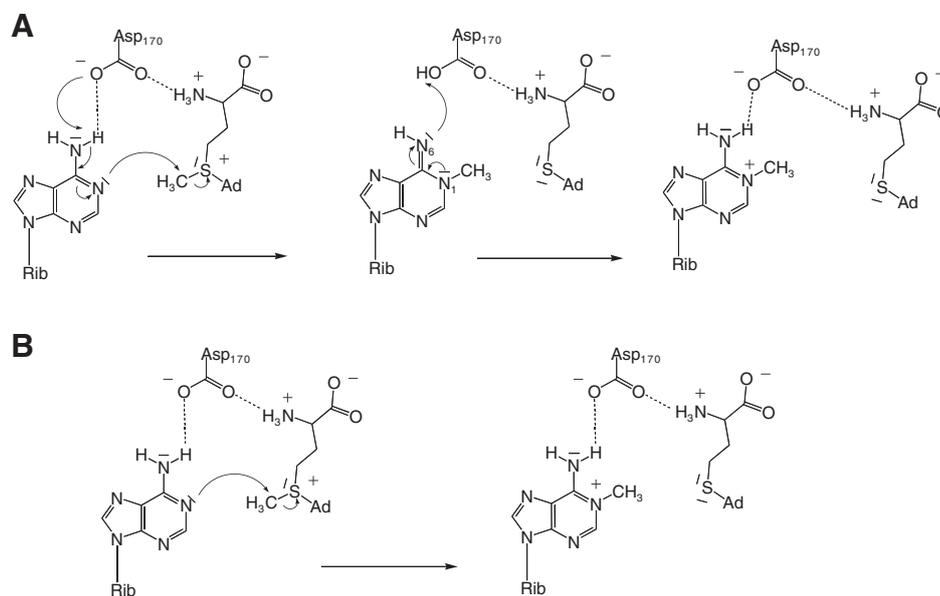


Fig. 5. Possible mechanisms for the methyl transfer at position 1 of adenines. (A and B) The two possible mechanisms for the formation of m^1A catalyzed by the wild-type TrmI enzyme. (A) With deprotonation of the amino exocyclic group of the adenine ring and *via* the intermediate imino tautomer of m^1A . (B) Direct transfer of the methyl, without intermediate.

not methylated by TrmI [42] is in accordance with either of the mechanisms. Indeed, the default in methyl transfer activity could be due to a default in the deprotonation of the substrate at position N_6 (mechanism **A**) or to an inappropriate orientation in the catalytic pocket of a substrate lacking the exocyclic amino-group at position N_6 (mechanism **B**).

In conclusion, the actual chemical mechanism that occurs in the enzymatic pockets of m^1A MTases is still a source of debate: whether or not it needs a deprotonation of the N_6 position of the adenine ring has not been clearly established. To our knowledge, the two mechanisms discussed here (Fig. 5A and B) are perfectly in accordance with the available literature. Although we favor mechanism **B** that satisfies best the principle of simplicity, further studies, either experimental or computational would be needed to definitely determine and strengthen the chemical mechanism of m^1A MTases.

4. Conclusions

First, we report here crystal structures of two TrmI mutants, namely D170A and Y78A. Interestingly, these structures point the dynamics of the N-terminal domain that is thought to bind the tRNA substrate. A normal mode analysis confirms the mobility of this domain. The data available on TrmI enzymes seem to converge to stress that the N-terminal domain of TrmI binds to the variable loop of tRNA substrates. In this context, the mobility of the N-terminal domain could be an advantage to grasp this loop.

The structure of the Y78A mutant is identical to that of wild-type TrmI except at the exact position of the residue. Indeed, the mutation makes a hole at the floor of the active site that can explain the defect of activity of this mutant. The space arising from the absence of the aromatic ring probably destabilizes the adenosine 58 flipped in the active site. The aromatic ring is therefore crucial for the stabilization of the sugar ring conformation. This structure reinforces our previous assumption that Y78 is, probably, not involved in the methyl transfer.

The structure of D170A revealed a partial disorganization of the active pocket. The co-factor could not be co-crystallized with the mutant protein. As a consequence, one loop of the protein obstructs the binding site of SAM and another moves at the entrance of the catalytic pocket. This designates D170 to act as a latch that holds the co-factor in the catalytic pocket and locks the adenine in a reactive position.

Secondly, the idea that the sequence of tRNA could have an importance for TrmI activity was recently shown. We analyzed the available sequences of *T. thermophilus* tRNA and these new data to discuss the determinants of recognition between TrmI and its substrate tRNA at the level of the T-arms.

Finally, we discuss the possible mechanism of methyl transfer catalyzed by TrmI. D170 is a central residue for the methyl transfer reaction, and two mechanisms, with or without deprotonation of the N_6 exocyclic amino group, are in accordance with the available information. Further studies will be necessary to decipher between these two mechanisms.

Acknowledgments

These experiments were performed on the ID23-2 beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and PX1 (Soleil, Gif S/Yvette). We are grateful to Local Contact at ESRF and Soleil for providing assistance in using the beamlines. The authors also thank Philippe Benas for the collection of data at ESRF.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bpc.2015.06.012>.

References

- [1] E.M. Phizicky, A.K. Hopper, tRNA biology charges to the front, *Genes Dev.* 24 (2010) 1832–1860.
- [2] A.K. Hopper, E.M. Phizicky, tRNA transfers to the limelight, *Genes Dev.* 17 (2003) 162–180.
- [3] K. Nakanishi, S. Fukai, Y. Ikeuchi, A. Soma, Y. Sekine, T. Suzuki, O. Nureki, Structural basis for lysidine formation by ATP pyrophosphatase accompanied by a lysine-specific loop and a tRNA-recognition domain, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 7487–7492.
- [4] H. Li, Complexes of tRNA and maturation enzymes: shaping up for translation, *Curr. Opin. Struct. Biol.* 17 (2007) 293–301.
- [5] Y. Motorin, M. Helm, tRNA stabilization by modified nucleotides, *Biochemistry* 49 (2010) 4934–4944.
- [6] H. Hori, Methylated nucleosides in tRNA and tRNA methyltransferases, *Front. Genet.* 5 (2014) 144.
- [7] M.A. Machnicka, A. Olchowiak, H. Grosjean, J.M. Bujnicki, Distribution and frequencies of post-transcriptional modifications in tRNAs, *RNA Biol.* 11 (2014) 1619–1629.

Chapitre II

Mutants du site actif de TrmI – Article 1 : Dégut et al. (2015) *Biophys. Chem.*

- [8] J. Urbonavicius, Q. Qian, J.M. Durand, T.G. Hagervall, G.R. Bjork, Improvement of reading frame maintenance is a common function for several tRNA modifications, *EMBO J.* 20 (2001) 4863–4873.
- [9] P.F. Agris, F.A. Vendeix, W.D. Graham, tRNA's wobble decoding of the genome: 40 years of modification, *J. Mol. Biol.* 366 (2007) 1–13.
- [10] U. Begley, M. Dyavaiah, A. Patil, J.P. Rooney, D. DiRenzo, C.M. Young, D.S. Conklin, R.S. Zitomer, T.J. Begley, Trm9-catalyzed tRNA modifications link translation to the DNA damage response, *Mol. Cell* 28 (2007) 860–870.
- [11] C.T. Chan, M. Dyavaiah, M.S. DeMott, K. Taghizadeh, P.C. Dedon, T.J. Begley, A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress, *PLoS Genet.* 6 (2010) e1001247.
- [12] C. Chen, B. Huang, M. Eliasson, P. Ryden, A.S. Bystrom, Elongator complex influences telomeric gene silencing and DNA damage response by its role in wobble uridine tRNA modification, *PLoS Genet.* 7 (2011) e1002258.
- [13] B. Nawrot, E. Sochacka, M. Duchler, tRNA structural and functional changes induced by oxidative stress, *Cell. Mol. Life Sci.* 663 (2011) 4023–4032.
- [14] D.M. Thompson, C. Lu, P.J. Green, R. Parker, tRNA cleavage is a conserved response to oxidative stress in eukaryotes, *RNA* 14 (2008) 2095–2103.
- [15] D.M. Thompson, R. Parker, Stressing out over tRNA cleavage, *Cell* 138 (2009) 215–219.
- [16] W.L. Towns, T.J. Begley, Transfer RNA methyltransferases and their corresponding modifications in budding yeast and humans: activities, predications, and potential roles in human health, *DNA Cell Biol.* 31 (2012) 434–454.
- [17] S. Gehrig, M.E. Eberle, F. Botschen, K. Rimbach, F. Eberle, T. Eigenbrod, S. Kaiser, W.M. Holmes, V.A. Erdmann, M. Sprinzl, G. Bec, G. Keith, A.H. Dalpke, M. Helm, Identification of modifications in microbial, native tRNA that suppress immunostimulatory activity, *J. Exp. Med.* 209 (2012) 225–233.
- [19] S. Laxman, B.M. Sutter, X. Wu, S. Kumar, X. Guo, D.C. Trudgian, H. Mirzaei, B.P. Tu, Sulfur amino acids regulate translational capacity and metabolic homeostasis through modulation of tRNA thiolation, *Cell* 154 (2013) 416–429.
- [20] B. Zinshteyn, W.V. Gilbert, Loss of a conserved tRNA anticodon modification perturbs cellular signaling, *PLoS Genet.* 9 (2013) e1003675.
- [21] I. Moukadiiri, M.J. Garzon, G.R. Bjork, M.E. Armengod, The output of the tRNA modification pathways controlled by the *Escherichia coli* MnmEG and MnmC enzymes depends on the growth conditions and the tRNA species, *Nucleic Acids Res.* 42 (2014) 2602–2623.
- [22] P.C. Dedon, T.J. Begley, A system of RNA modifications and biased codon use controls cellular stress response at the level of translation, *Chem. Res. Toxicol.* 27 (2014) 330–337.
- [23] C.T. Chan, Y.L. Pang, W. Deng, I.R. Babu, M. Dyavaiah, T.J. Begley, P.C. Dedon, Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins, *Nat. Commun.* 3 (2012) 937.
- [24] G. Dirheimer, W. Baranowski, G. Keith, Variations in tRNA modifications, particularly of their queuine content in higher eukaryotes. Its relation to malignancy grading, *Biochimie* 77 (1995) 99–103.
- [25] T. Suzuki, A. Nagao, Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases, *Annu. Rev. Genet.* 45 (2011) 299–329.
- [26] B. El Yacoubi, M. Bailly, V. de Crecy-Lagard, Methyltransferase and function of posttranscriptional modifications of transfer RNAs, *Annu. Rev. Genet.* 46 (2012) 69–95.
- [27] A.G. Torres, E. Batlle, L. Ribas de Pouplana, Role of tRNA modifications in human diseases, *Trends Mol. Med.* 20 (2014) 306–314.
- [28] P. Barraud, B. Golinelli-Pimpaneau, C. Atmanene, S. Sanglier, A. Van Dorssele, L. Droogmans, F. Dardel, C. Tisne, Crystal structure of *Thermus thermophilus* tRNA m¹A₅₈ methyltransferase and biophysical characterization of its interaction with tRNA, *J. Mol. Biol.* 377 (2008) 535–550.
- [29] A. Guelorget, P. Barraud, C. Tisne, B. Golinelli-Pimpaneau, Structural comparison of tRNA m¹A₅₈ methyltransferases revealed different molecular strategies to maintain their oligomeric architecture under extreme conditions, *BMC Struct. Biol.* 11 (2011) 48.
- [30] M. Saikia, Y. Fu, M. Pavon-Eternod, C. He, T. Pan, Genome-wide analysis of N1-methyl-adenosine modification in human tRNAs, *RNA* 16 (2010) 1317–1327.
- [31] J. Anderson, L. Phan, A.G. Hinnebusch, The Gcd10p/Gcd14p complex is the essential two-subunit tRNA (1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 5173–5178.
- [32] S. Auxilien, G. Keith, S.F. Le Grice, J.L. Darlix, Role of post-transcriptional modifications of primer tRNA_{Lys,3} in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription, *J. Biol. Chem.* 274 (1999) 4412–4420.
- [33] L. Droogmans, M. Roovers, J.M. Bujnicki, C. Tricot, T. Hartsch, V. Stalon, H. Grosjean, Cloning and characterization of tRNA (m¹A₅₈) methyltransferase (TrmI) from *Thermus thermophilus* HB27, a protein required for cell growth at extreme temperatures, *Nucleic Acids Res.* 31 (2003) 2148–2156.
- [34] K. Watanabe, M. Shinma, T. Oshima, S. Nishimura, Heat-induced stability of tRNA from an extreme thermophile, *Thermus thermophilus*, *Biochem. Biophys. Res. Commun.* 72 (1976) 1137–1144.
- [35] N. Horie, M. Hara-Yokoyama, S. Yokoyama, K. Watanabe, Y. Kuchino, S. Nishimura, T. Miyazawa, Two tRNA^{Ala1} species from an extreme thermophile, *Thermus thermophilus* HB8: effect of 2-thiolation of ribothymidine on the thermostability of tRNA, *Biochemistry* 24 (1985) 5711–5715.
- [36] N. Shigi, T. Suzuki, T. Terada, M. Shirouzu, S. Yokoyama, K. Watanabe, Temperature-dependent biosynthesis of 2-thioribothymidine of *Thermus thermophilus* tRNA, *J. Biol. Chem.* 281 (2006) 2104–2113.
- [37] C. Tomikawa, T. Yokogawa, T. Kanai, H. Hori, N7-methylguanine at position 46 (m⁷G46) in tRNA from *Thermus thermophilus* is required for cell viability at high temperatures through a tRNA modification network, *Nucleic Acids Res.* 38 (2010) 942–957.
- [38] S. Yokoyama, K. Watanabe, T. Miyazawa, Dynamic structures and functions of transfer ribonucleic acids from extreme thermophiles, *Adv. Biophys.* 23 (1987) 115–147.
- [39] K. Ishida, T. Kunibayashi, C. Tomikawa, A. Ochi, T. Kanai, A. Hirata, C. Iwashita, H. Hori, Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*, *Nucleic Acids Res.* 39 (2011) 2304–2318.
- [40] A. Gupta, P.H. Kumar, T.K. Dineshkumar, U. Varshney, H.S. Subramanya, Crystal structure of Rv2118c: an AdoMet-dependent methyltransferase from *Mycobacterium tuberculosis* H37Rv, *J. Mol. Biol.* 312 (2001) 381–391.
- [41] U. Varshney, V. Ramesh, A. Madabushi, R. Gaur, H.S. Subramanya, U.L. RajBhandary, *Mycobacterium tuberculosis* Rv2118c codes for a single-component homotetrameric m¹A₅₈ tRNA methyltransferase, *Nucleic Acids Res.* 32 (2004) 1018–1027.
- [42] H. Takuma, N. Ushio, M. Minoji, A. Kazayama, N. Shigi, A. Tomikawa, A. Ochi, H. Hori, Substrate tRNA recognition mechanism of eubacterial tRNA (m¹A₅₈) methyltransferase (TrmI), *J. Biol. Chem.* 290 (2015) 5912–5925.
- [43] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, *Acta Crystallogr. D Biol. Crystallogr.* 67 (2011) 235–242.
- [44] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (2007) 658–674.
- [45] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 213–221.
- [46] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 486–501.
- [47] J.E. Debrezzeni, P. Emsley, Handling ligands with Coot, *Acta Crystallogr. D Biol. Crystallogr.* 68 (2012) 425–430.
- [48] A. Bakan, L.M. Meireles, I. Bahar, ProDy: protein dynamics inferred from theory and experiments, *Bioinformatics* 27 (2011) 1575–1577.
- [49] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38 (27–38).
- [50] D. Hamdane, A. Guelorget, V. Guerneau, B. Golinelli-Pimpaneau, Dynamics of RNA modification by a multi-site-specific tRNA methyltransferase, *Nucleic Acids Res.* 42 (2014) 11697–11706.
- [51] S. Klimasauskas, G. Lukinavicius, AdoMet-dependent methyltransferases, chemistry of, *Wiley Encyclopedia of Chemical Biology*, John Wiley & Sons, Inc. 2008, pp. 1–10.
- [52] S. Boschi-Muller, Y. Motorin, Chemistry enters nucleic acids biology: enzymatic mechanisms of RNA modification, *Biokhimiya/Biochemistry* 78 (2013) 1392–1404.
- [53] H.L. Schubert, R.M. Blumenthal, X. Cheng, Many paths to methyltransfer: a chronicle of convergence, *Trends Biochem. Sci.* 28 (2003) 329–335.
- [54] Y. Peng, D. Sartini, V. Pozzi, D. Wilk, M. Emanuelli, V.C. Yee, Structural basis of substrate recognition in human nicotinamide N-methyltransferase, *Biochemistry* 50 (2011) 7800–7808.
- [55] Y. Yue, H. Guo, Quantum mechanical/molecular mechanical study of catalytic mechanism and role of key residues in methylation reactions catalyzed by dimethylxanthine methyltransferase in caffeine biosynthesis, *J. Chem. Inf. Model.* 54 (2014) 593–600.
- [56] A.A. McCarthy, J.G. McCarthy, The structure of two N-methyltransferases from the caffeine biosynthetic pathway, *Plant Physiol.* 144 (2007) 879–889.
- [57] Y.M. Hou, J.J. Perona, Stereochemical mechanisms of tRNA methyltransferases, *FEBS Lett.* 584 (2010) 278–286.
- [58] W.E. Swinehart, J.E. Jackman, Diversity in mechanism and function of tRNA methyltransferases, *RNA Biol.* 12 (2015) 398–411.
- [59] P.A. Elkins, J.M. Watts, M. Zalacain, A. van Thiel, P.R. Vitazka, M. Redlak, C. Andraos-Selim, F. Rastinejad, W.M. Holmes, Insights into catalysis by a knotted TrmD tRNA methyltransferase, *J. Mol. Biol.* 333 (2003) 931–949.
- [60] T. Christian, G. Lahoud, C. Liu, K. Hoffmann, J.J. Perona, Y.M. Hou, Mechanism of N-methylation by the tRNA m¹G37 methyltransferase Trm5, *RNA* 16 (2010) 2484–2492.
- [61] Z. Shao, W. Yan, J. Peng, X. Zuo, Y. Zou, F. Li, D. Gong, R. Ma, J. Wu, Y. Shi, Z. Zhang, M. Teng, X. Li, Q. Gong, Crystal structure of tRNA m¹G9 methyltransferase Trm10: insight into the catalytic mechanism and recognition of tRNA substrate, *Nucleic Acids Res.* 42 (2014) 509–525.
- [62] G. Schluckebier, P. Zhong, K.D. Stewart, T.J. Kavanaugh, C. Abad-Zapatero, The 2.2 Å structure of the rRNA methyltransferase ErmC and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism, *J. Mol. Biol.* 289 (1999) 277–291.
- [63] A.S. Punekar, J. Liljeruhm, T.R. Shepherd, A.C. Forster, M. Selmer, Structural and functional insights into the molecular mechanism of rRNA m⁶A methyltransferase RlmJ, *Nucleic Acids Res.* 41 (2013) 9537–9548.
- [64] A. Kettani, M. Gueron, J.-L. Leroy, Amino proton exchange processes in mononucleosides, *J. Am. Chem. Soc.* 119 (1997) 1108–1115.
- [65] R. Stewart, M.G. Harris, Amino group acidity in nucleotide bases, *Can. J. Chem.* 55 (1977) 3807–3814.

Chapitre II

Mutants du site actif de TrmI –

Article 1 : Dégut et al. (2015) *Biophys. Chem.*

IV. Complexe covalent TrmI/bras T de l'ARNt

IV.1 Introduction

De nombreux essais de co-cristallisation entre TrmI et des ARNt pleine longueur ou des mini-substrats mimant le bras T des mêmes ARNt ont été menés dans l'équipe, mais n'ont pas donné de résultats jusqu'à présent. Afin de maximiser nos chances d'obtenir des cristaux du complexe, nous avons décidé de produire des complexes covalents entre TrmI et des tiges-boucles mimant le bras T de l'ARNt^{ASP} de *T. thermophilus*. Une collaboration avec l'équipe de Ronald Micura (Université d'Innsbruck) a été développée. Il est spécialisé dans la synthèse chimique d'ARN et de leurs dérivés. Son groupe a développé la synthèse d'ARN modifiés. Une chaîne cystamine peut être greffée sur l'azote exocyclique des guanosines et des cytidines (**Figure IV-1**). Cette méthode a été décrite pour cristalliser le complexe entre la transcriptase inverse du VIH et un ADN servant d'amorce [289, 290]. L'idée est que le groupement cystamine peut faire un échange de pont disulfure avec une cystéine libre de la protéine, proche du nucléotide portant ce groupement. L'utilisation de soufre oxydé sous la forme d'un pont disulfure (S-S) permet d'éviter la formation de complexe ARN-S-S-ARN, qui nuirait à la formation du complexe qui nous intéresse. Trois ARN contenant la séquence du bras T de l'ARNt^{ASP} de *T. thermophilus* ont été testés chacun portant une des modifications indiquées sur la **Figure IV-1**. Cette méthode implique que l'enzyme ne possède qu'une seule cystéine afin d'obtenir un complexe covalent homogène. Le choix des mutants ponctuels en cystéine de TrmI (nommé après en mutant cystéine) est développé dans le paragraphe suivant. Les ARN réactif modifiés seront appelés dans ce chapitre ARN* de façon générique et NXXN* de façon spécifique, où N est le nom du nucléotide modifié et XX sa position dans le bras T. Par exemple G56G*.

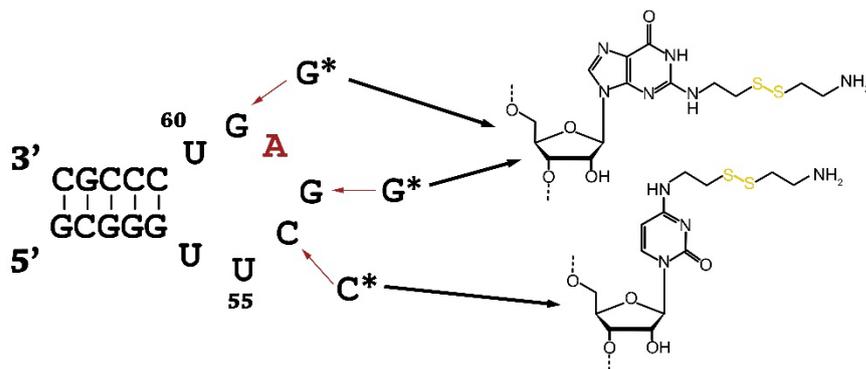
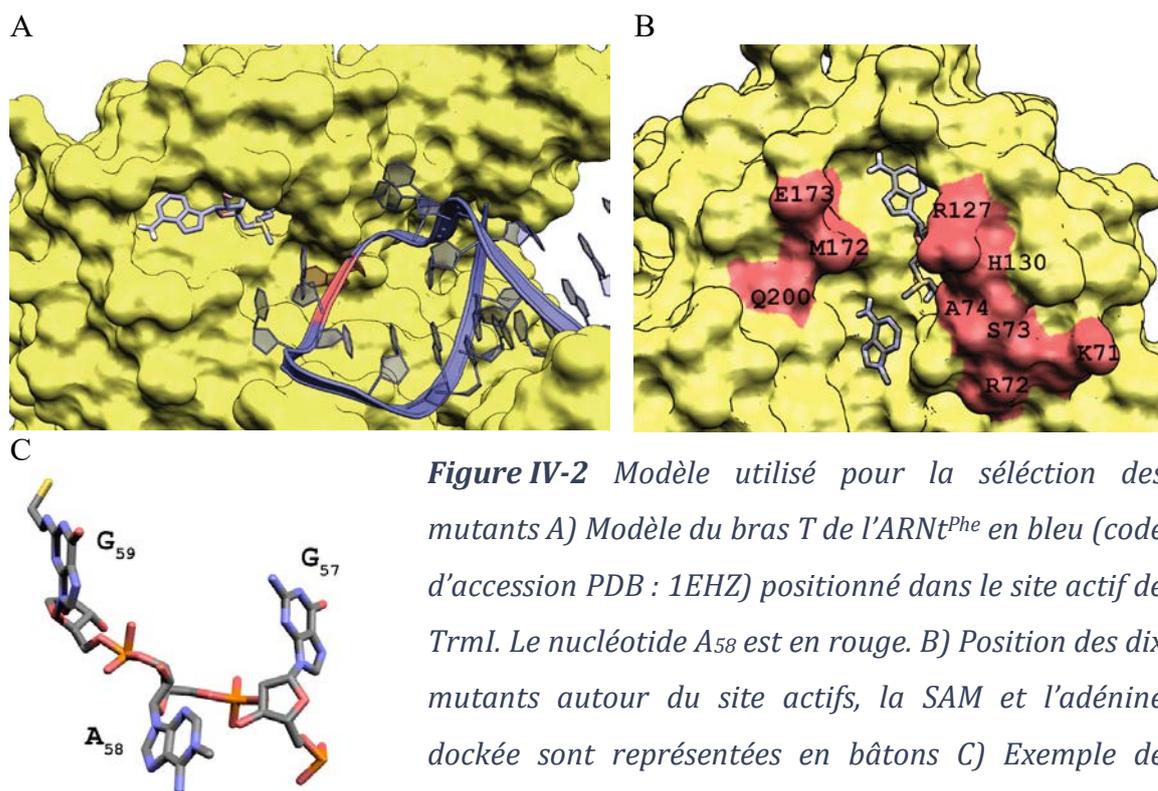


Figure IV-1 : Séquences d'ARN utilisées pour former des complexes covalents avec TrmI. L'adénine modifiée par TrmI est indiquée en rouge. Les flèches indiquent les 3 nucléotides de la boucle T qui peuvent porter le groupement cystamine.

IV.2 Mutations ponctuelles en cystéines

IV.2.1 Sélection des mutants cystéines de TrmI

TrmI de *T. thermophilus* ne possède pas de cystéine. Nous avons décidé de sélectionner une dizaine de résidus à muter en cystéine autour du site actif. Pour sélectionner les mutants, j'ai créé un modèle simple de la tige-boucle à partir de la structure cristallographique de l'ARNt^{Phe} de levure (qui porte une m¹A₅₈) [13]. J'ai récupéré le bras T de cette structure, et ai minimisé l'énergie de la structure avec AVOGADRO [291] pour obtenir un modèle de tige-boucle simple. J'ai ensuite tourné l'adénine 58 pour la placer dans la même position que dans le résultat de docking réalisé précédemment avec HADDOCK [292, 293] (**Figure IV-2 A**). J'ai ajouté avec AVOGADRO le groupement thioethyl (-CH₂-CH₂-SH) sur les azotes exocycliques des guanines, le tout dans une dizaine de conformères différents (**Figure IV-2**). Ces modèles permettent de voir facilement quels résidus seraient, ou ne seraient pas accessibles au pont disulfure de la fonction cystamine. Les dix mutants sélectionnés sont : K71C, R72C, S73C, A74C, R127C, H130C, M172C, E173C, Q200C et P231C (**Figure IV-2 B**). Les plasmides de ces mutants ont été construits par mutagenèse dirigée par la société Genecust.



IV.2.2 Production et purification des différents mutants de TrmI

Les dix mutants ont été testés produit et purifié avec le même protocole que la protéine sauvage [281]. L'état oligomérique des différents mutants a été analysé par chromatographie d'exclusion. La présence ou l'absence de pont disulfure intramoléculaire non souhaité ont été analysées par gels SDS-PAGE sans agent réducteur. Les résultats sont résumés dans le **Tableau IV-4**. Seuls deux mutants (M172C et R127C) sur les dix testés n'ont pas été surexprimés dans *E. coli*. Tous les mutants produits montraient une bonne solubilité et pas de production en corps d'inclusion. L'état oligomérique est variable selon les mutants, certaine mutation semble avoir déstabilisé l'architecture du tétramère. Enfin on retrouve assez peu de

pont dissulfure intermoléculaire. Nous avons de préférence travaillé avec les protéines tétramériques en solution, les mutants S72C (dimère) et E173C (monomère) ont été écartés.

N°	Mutation	expression	solubilité	État oligomérique	Pont disulfure
1	H130C	oui	oui	tétramère	peu
2	S73C	oui	oui	Dimère/tétramère	peu
3	E173C	oui	oui	monomère	non
4	A74C	oui	oui	tétramère	peu
5	R72C	oui	oui	Dimère	peu
6	Q200C	oui	oui	tétramère	non
7	M172C	non	/	/	/
8	K71C	oui	oui	tétramère	peu
9	P231C	oui	oui	tétramère	peu
10	R127C	non	/	/	/

Tableau IV-4 Résumé des tests d'expressions et de purification des différents mutants cystéines de la protéine. La surexpression par la bactérie *E. coli* (BL21(DE3)), la solubilité (absence de corps d'inclusion), l'état oligomérique en solution et la présence de pont disulfure intermoléculaire ont été testés. Expression et solubilité indique que le mutant est surexprimé dans *E. coli* dans la fraction soluble.

IV.3 Formation du complexe covalent TrmI/BrasT

IV.3.1 Tests de formation des complexes covalents en microvolume

Comme exposé précédemment, les ARN modifiés sont produit dans l'équipe de Ronald Micura. Ils ont rencontré certaines difficultés de synthèses qui nous ont conduit à travailler sur de très faibles quantités d'ARN réactifs modifiés (6 nmol pour la plus faible quantité, 54 la plus grande).

ARN*	Séquence	Quantité
17nt_C56C*	5'-GCG GGU UC*G AGU CCC GC-3'	6 nmol
17nt_G57G*	5'-GCG GGU UCG* AGU CCC GC-3'	54 nmol
17nt_G59G*	5'-GCG GGU UCG AG*U CCC GC-3'	34 nmol

Tableau IV-5 Noms, séquences, et quantité des ARN disponibles*

Six mutants ont été testés avec les différents ARN. Les ARN* ont été repris dans un tampon 50 mM Tris pH 8,0 contenant 250 mM KCl pour une concentration finale de 500 μ M en ARN*. Pour chaque protéine, un stock à 300 μ M a été préparé dans le même tampon. Les réactions ont été faites dans 5 μ L, avec une concentration finale de 100 μ M en protéine et 200 μ M en ARN*, la présence de SAM (500 μ M) dans le milieu réactionnel a également été testée. Les réactions ont été faites en micro tube dans un thermocycler pendant 1h à 40 °C. Les résultats des réactions ont été analysés par gels SDS-PAGE sans agent réducteur (**Figure IV-3**).

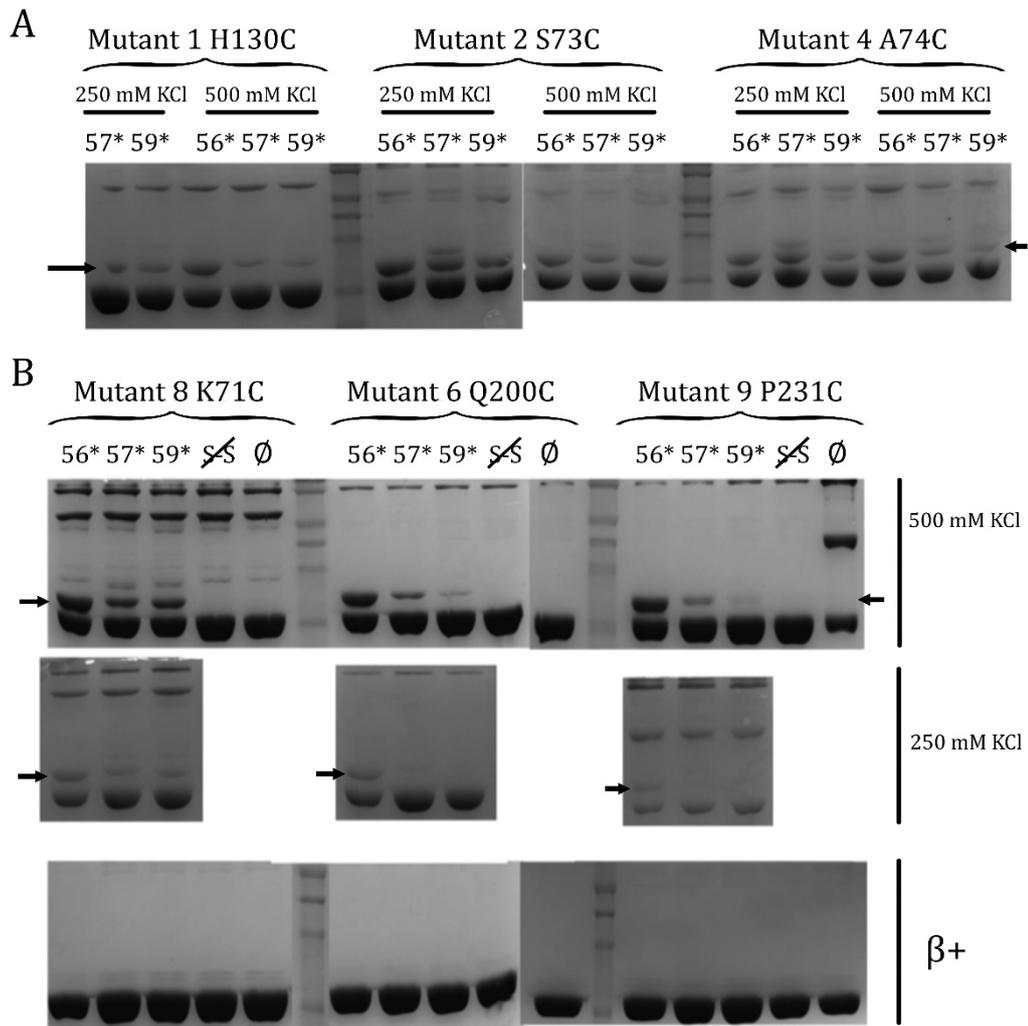


Figure IV-3 Gels SDS-PAGE (sans agent réducteur) analysant les réactions en micro volume. Les différents mutants de TrmI ont été testés à 2 concentrations de sels (250 et 500 mM KCl) et avec les trois ARN (noté 56* 57* et 59*). Sur les gels, la bande la plus basse correspond à la protéine libre, la présence d'un ARN lié à la protéine provoque un retard de migration pointé par les flèches. Pour le jeu de réaction en (B), un ARN non modifié a été testé comme témoin (~~S-S~~), le Ø correspond à la protéine seule. Les trois derniers gels correspondent aux gels témoins, les échantillons ont été mélangés à un agent réducteur (β -mercaptoéthanol à 15 mM), pour vérifier que les effets observé étaient dus à la formation de ponts disulfure.

Trois résultats importants ont été obtenus à partir de ces réactions tests :

- l'ARN C56C* réagit avec tous les mutants avec un très bon rendement, mais c'est l'ARN que nous avons en plus faible quantité.
- l'ARN G59G* a un bon taux de réaction avec les mutants 8 (K71C) et 2 (S73C)
- l'ARN G57G* a plutôt un mauvais rendement, mais avec le mutant 6 (Q200C) la réaction a un rendement correcte, avec peu d'agrégation (bande tout en haut du gel)
- La présence de SAM ne change pas le taux de réaction (non montré ici)

Ces résultats nous ont amenés à poursuivre les tests de cristallogenèse en utilisant l'ARN G57G*, que nous avons en plus grande quantité, avec le mutant Q200C de TrmI. Quatre autres tests, avec des volumes réactionnels de 2,5 μ L ont été fait sur ce couple ARN/mutant. A deux températures différentes 20 et 30 °C et avec 150 et 250 mM KCl, mais aucune différence notable a été observée.

IV.3.2 Essais de cristallogenèse du complexe covalent TrmI Q200C/ARN G57G*

J'ai donc entrepris de faire la réaction entre l'ARN G57G* et le mutant TrmI Q200C en plus grand volume : 400 μ L de réaction avec la protéine et l'ARN à 100 μ M dans un tampon 50 mM Tris pH 8,0 avec 250 mM KCl, pendant 1h à 35 °C. La purification a été conduite à 4 °C sur un système de purification Aktapurifier et une colonne MonoQ de 8 mL (GE Healthcare). Le gradient allait de 100 mM KCl à 500 mM en 500 mL à 4mL/min.

Un pic, correspondant au complexe, a pu être séparé du reste du mélange réactionnel et concentré sur amicon (cutoff 15 kDa) (**Figure IV-4**). Le spectre UV, et le gel nous ont permis d'évaluer la stœchiométrie de l'ARN par tétramère : il est visible sur le gel qu'une large part de la protéine n'est pas en complexe avec l'ARN même après purification. Par ailleurs le spectre UV montre une intensité à peu près équivalente pour l'ARN ($E_{260nm} = 151\,500 \text{ L.mol}^{-1}.\text{cm}^{-1}$) et la protéine ($E_{280nm} = 41\,940 \text{ L.mol}^{-1}.\text{cm}^{-1}$), ce qui correspond à un ratio de 1 ARN pour 4 protéines.

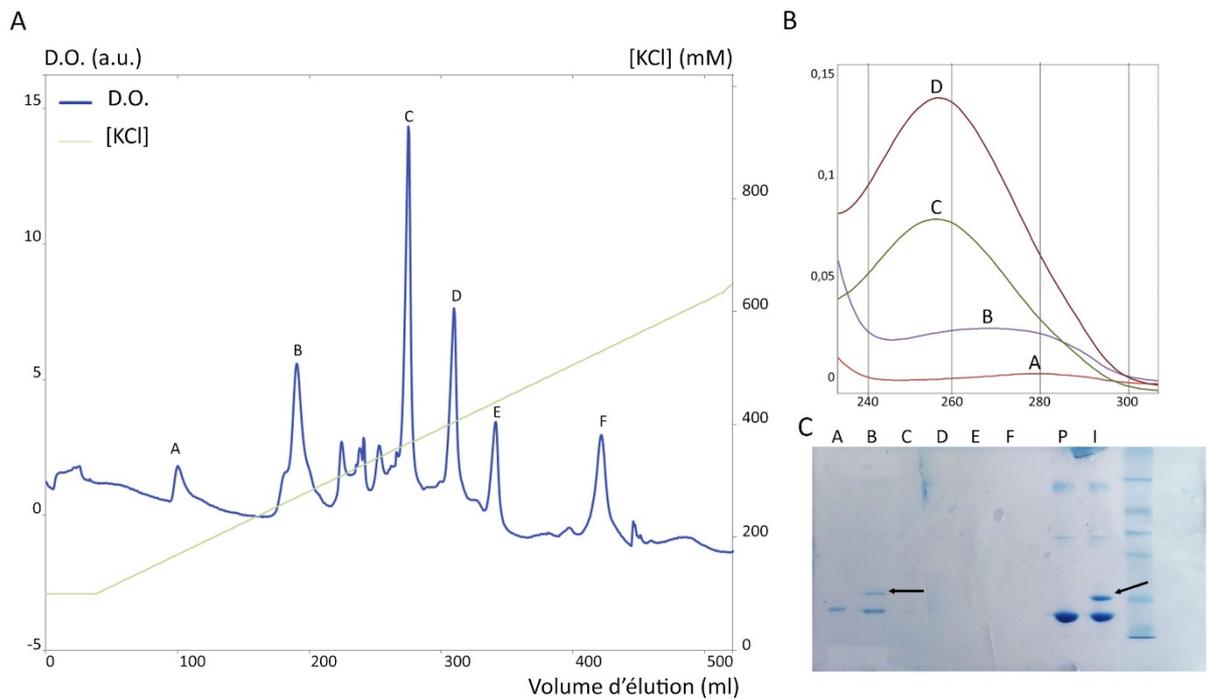


Figure IV-4 Purification du complexe covalent *TrmI-Q200C/ARN-G57G**, A) Profil de purification obtenu après injection de la réaction *TrmI-Q200C/ARN-G57G** sur une colonne MonoQ, la protéine libre est retrouvée dans le pic A, le complexe est élué en B, les pics C, D, E et F sont constitués d'ARN seul. B) profils spectrophotométriques UV des différents pics du chromatogramme, la courbe du pic B présente un maximum aux alentours de 270 nm caractéristique du complexe. C) gel SDS-PAGE à 12% dénaturant mais sans agent réducteur, correspondant aux différents pics (A,B,C,D,E,F), le témoin de la protéine seul est en P, le mélange injecté sur la colonne en I. La formation du complexe ARN-protéine provoque un décalage bien visible sur le gel (flèches bande B et I).

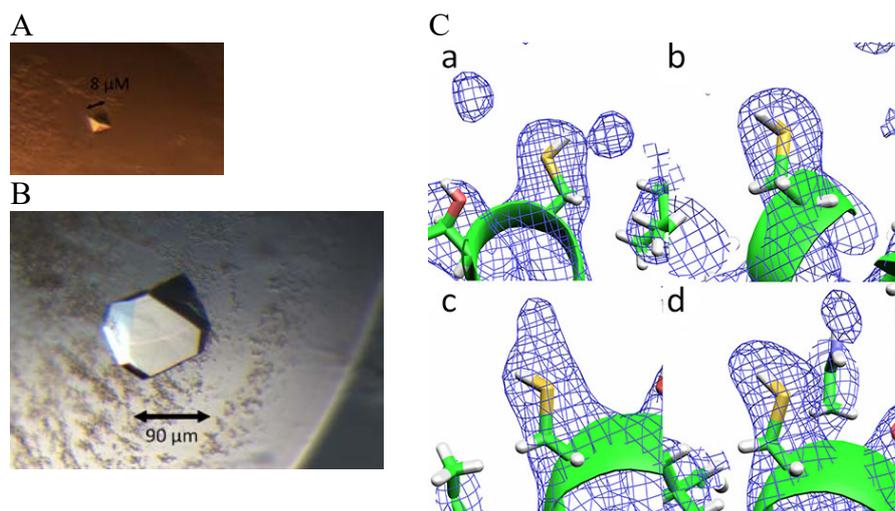
Les kits de criblage de conditions de cristallogénèse, Nucleix (Qiagen) et Crystal Screen (Hampton), ont été utilisés à l'aide d'un robot Mosquito (ttplabtech) afin d'essayer de trouver des conditions de cristallisation. Les tests ont été conduits à 18°C. La concentration en complexe est faible avec une $DO_{270nm} = 4$ soit environ 2 à 4 mg/mL, la superposition des spectres UV de l'ARN et de la protéine rend la

Chapitre II

Complexe covalent *TrmI*/bras T de l'ARNt – Formation du complexe covalent *TrmI/BrasT*

concentration difficile à estimer. Des petits cristaux sont apparus (inférieur à $10\ \mu\text{m}$ **Figure IV-5 A**), mais n'ont pas permis d'enregistrer des jeux de données. Par ailleurs, des cristaux ont été produits dans les conditions de cristallisation de la protéine sauvage (2 M sulfate d'ammonium et 5% isopropanol **Figure IV-5 B**). Un jeu de données a pu être enregistré sur la ligne id23-2 à l'ESRF à une résolution de $2.0\ \text{\AA}$. J'ai résolu la structure par remplacement moléculaire en utilisant Phaser de la suite PHENIX et la structure de TrmI sauvage (code d'accèsion PDB 2PWY). Il y a 4 molécules dans l'unité asymétrique, et le groupe d'espace est un $P2_12_12_1$ (dimension de l'ua : 74.42 79.02 184.53 90 90 90). La densité électronique, pourtant bien définie, ne contient pas de trace de l'ARN, il n'y a aucune densité visible aux abords des cystéines, après affinement jusqu'à $R_{\text{Free}} = 0,34$, et sur la carte calculée sans aplanissement du solvant désordonné (**Figure IV-5 C**).

On sait que l'on a une stœchiométrie de un ARN pour quatre protéines, donc un ARN par tétramère en solution. L'explication la plus probable pour l'absence d'ARN semble être un échange des monomères : le tétramère s'associe et se dissocie de façon dynamique en solution, avec reformation de tétramère libre. Les conditions de cristallisation de la protéine sauvage favorisant très fortement la cristallisation de tétramères libres au détriment de ceux covalamment lié à un ARN. Cette condition est donc à exclure pour l'obtention de cristaux du complexe.



Chapitre II

Complexe covalent TrmI/bras T de l'ARNt – Formation du complexe covalent TrmI/BrasT

Figure IV-5 Cristaux du complexe TrmI/Bras T. (A) Cristaux obtenus dans des nouvelles conditions (50 mM Na Cacodylate pH 6.0 ; 200 mM KCl ; 10 mM CaCl 10% PEG 4000) aucune diffraction n'a été obtenue, les cristaux étant trop petits. (B) Cristaux du complexe obtenus dans les conditions de la protéine sauvage (2 M sulfate d'ammonium et 5% isopropanol). (C) Les quatre cystéines mutées de l'unité asymétrique, aucune densité ne laisse entrevoir la présence de l'ARN (carte F0-2Fc à 2.0σ)

Un second lot d'ARN G57G* de (60 nmol) nous a permis de reproduire la réaction de formation du complexe avec un bien meilleur rendement. En se basant sur la connaissance de la stœchiométrie et afin de limiter la présence d'ARN* libre, j'ai augmenté la quantité de protéine dans le milieu réactionnel (200 μ M de protéine et 100 μ M d'ARN). Le profil de purification montre alors un pic de complexe plus important, et moins d'ARN* libre. Cette purification nous a permis de refaire les mêmes kits de criblage de conditions de cristallogénèse que précédemment, mais avec des concentrations beaucoup plus élevées ($DO_{270nm} = 20$ et $DO_{270nm} = 10$). Aucun cristaux n'ont pour l'instant été obtenu à partir de cette dernière purification.

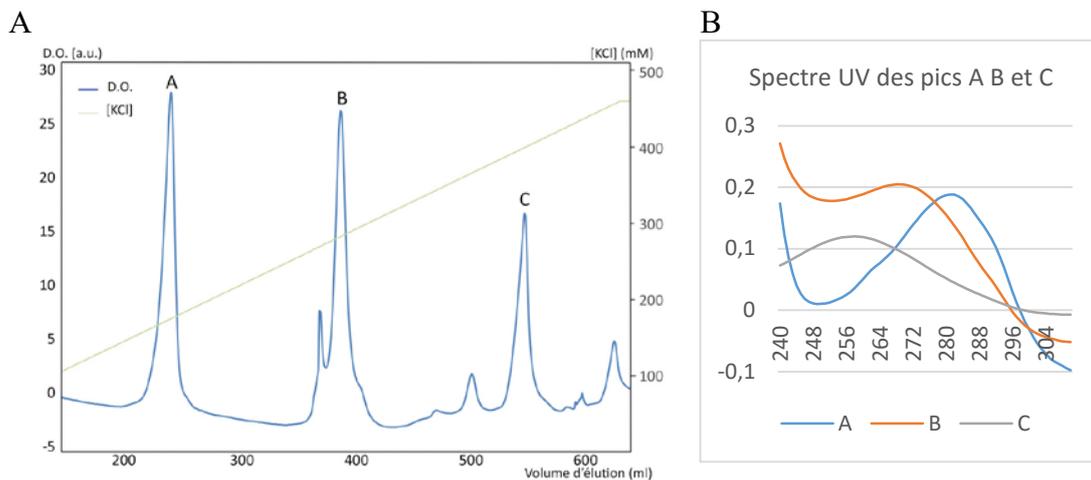


Figure IV-6 Purification de la seconde réaction de formation du complexe TrmI Q200C/ARN G57G*. (A) Pic A protéine seule, B complexe C ARN* seul. (B) Spectres UV

correspondant aux différents pics. La protéine seule a un maximum d'absorption à 280nm, l'ARN seul à 260nm et le complexe ARN/protéine à 270 nm, Le rendement de la réaction de formation du complexe (ratio Complexe/ARN seul) est bien meilleur lorsque la protéine est en excès dans le milieu.

IV.4 Perspectives

Nos travaux nous ont permis d'obtenir des complexes covalents entre plusieurs mutants cystéines de TrmI et des tiges-boucles modifiées. Il a été possible de purifier ce complexe de façon reproductible et en quantité suffisante, pour mener des premiers essais de cristallogénèse. Des cristaux dans les conditions de la protéine sauvage ont été obtenus, mais ils ne contiennent pas l'ARN. Par ailleurs, des cristaux dans des conditions différentes ont été obtenus, mais pour l'instant sans diffraction, les cristaux étant trop petit. Des tests restent à faire avec le mutant Q200C de TrmI et l'ARN G57G* (kits différents, concentration initiale en sel différente...). Par ailleurs d'autres mutants semblaient intéressants lors de nos tests notamment le mutant K71C et des essais de cristallisation sur ces complexes sont envisagés en fonction des ARN disponibles. Nos résultats sont très encourageants, mais le manque d'ARN nous a obligé à faire un choix sur les tests à réaliser, là où des tests systématiques seraient nécessaire.

Chapitre III Étude de la méthyltransférase m^1A_{22} des ARNt : TrmK

Les problématiques qui se posent pour TrmI, exposées dans le chapitre précédent, se posent également pour TrmK. Ainsi, l'étude du mécanisme réactionnel et des déterminants de la reconnaissance du substrat ARNt par TrmK sont les questionnements qui ont motivé cette étude.

Bien que la RMN soit une technique de choix pour observer l'interaction entre deux molécules en solution, cette technique n'était pas adaptée pour l'étude de l'interaction TrmI/ARNt, la taille du complexe (un tétramère et deux ARNt) étant trop importante pour une utilisation de la RMN avec un marquage classique. La protéine TrmK en revanche possède une taille qui la rend parfaitement accessible aux analyses RMN. En l'absence de cristaux du complexe TrmK/ARNt, nous avons utilisé la RMN pour étudier le complexe entre TrmK et son substrat ARNt.

L'essentiel de ce chapitre se rapporte à cette étude. La première partie est consacrée à la production d'ARNt *in vivo* et *in vitro* et est l'objet d'un chapitre d'ouvrage (article 2). La seconde partie est centrée sur l'étude RMN de TrmK et concerne l'attribution des signaux RMN de son squelette peptidique qui était nécessaire pour ensuite étudier l'interaction de TrmK avec un ARNt. Cette attribution fait l'objet d'une publication en préparation (article 3). La troisième partie aborde en particulier la structure cristallographique de TrmK que j'ai obtenu durant ma thèse, ainsi que l'utilisation de la technique d'empreintes RMN pour analyser l'interaction

entre TrmK et son substrat l'ARNt_{GGA}^{Ser}. L'ensemble des résultats obtenus sont présentés sous forme d'article (article 4).

La dernière partie de ce chapitre décrit des expériences en cours qui utilisent de petits analogues bisubstrats avec la SAM et l'adénine covalamment liée. Cette étude devrait nous permettre de progresser dans la compréhension du mécanisme de méthylation et de la reconnaissance de la base à modifier.

V. Production d'un substrat de TrmK : l'ARNt^{Ser} de *B. subtilis*

V.1 Problématique et résumé de l'article 2

La caractérisation de la fonction méthyltransférase de TrmK ayant été démontrée en utilisant l'ARNt_{GGA}^{Ser} de *B. subtilis* [258], nous avons mené nos études d'interaction avec cet ARNt. Deux méthodes principales ont été utilisées pour la production de l'ARNt^{Ser} : la transcription *in vitro*, et la production *in vivo* chez *E. coli* (**Figure V-1**). La différence majeure entre ces deux approches tient à la présence de modifications post-transcriptionnelles sur les ARNt produits *in vivo*. Ces ARN sont en effet pris en charge par la machinerie de maturation de la cellule hôte. On retrouvera donc sur cet ARNt recombinant toutes les modifications post-transcriptionnelles que les enzymes de la cellule hôte sont capable d'incorporer.

L'ARNt_{GGA}^{Ser} n'est produit qu'en faible quantité par transcription *in vitro*. Un travail d'optimisation conséquent a été nécessaire pour obtenir des quantités suffisantes pour des études structurales. Mes travaux pour l'optimisation de la production et la purification de cet ARNt sont présentés dans le chapitre 4 du livre *Methods in molecular biology : Nucleic Acid Crystallography* [294]. Le chapitre en question est intégralement restitué dans la partie V.2 en page 105.

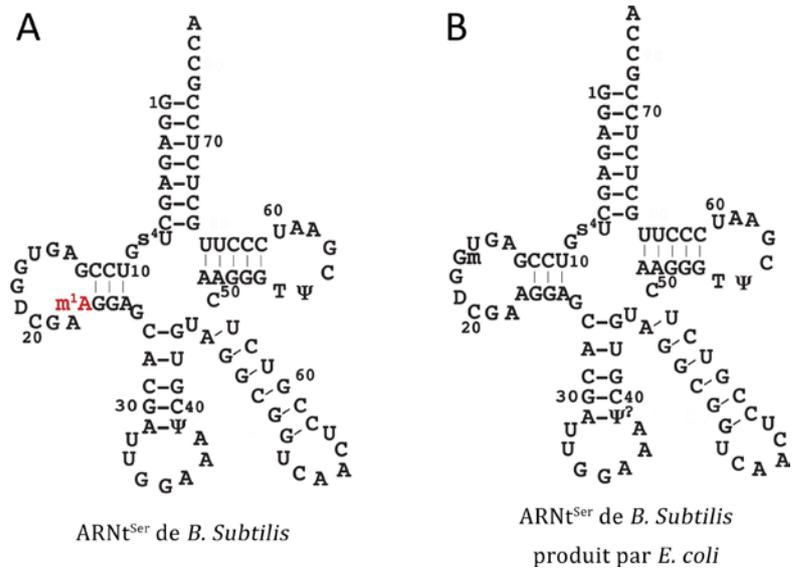


Figure V-1 Lors de la production *in vivo* d'ARNt, les ARN sont maturés et modifiés par les enzymes cellulaires. A), séquence de l'ARNt^{Ser} produit par *B. Subtilis*; B) le même ARNt, avec les modifications attendues lors de sa production dans *E. coli*.

V.2 Article 2 : Dégut *et al.* (2016)

Clément Dégut, Alexandre Monod, Franck Brachet, Thibaut Crépin, and Carine Tisé

In Vitro/In Vivo Production of tRNA for X-Ray Studies Methods in molecular biology: Nucleic Acid Crystallography [294]

In Vitro/In Vivo Production of tRNA for X-Ray Studies

Clément Dégut, Alexandre Monod, Franck Brachet,
Thibaut Crépin, and Carine Tisné

Abstract

tRNAs occupy a central role in the cellular life, and they are involved in a broad range of biological processes that relies on their interaction with proteins and RNA. Crystallization and structure resolution of tRNA or/and tRNA/partner complexes can yield in valuable information on structural organizations of key elements of cellular machinery. However, crystallization of RNA, is often challenging. Here we review two methods to produce and purify tRNA in quantity and quality to perform X-ray studies.

Key words In vitro transcription, In vivo production, tRNA purification, Crystallization

1 Introduction

tRNAs are universally conserved key components of the translation machinery, linking the genetic code with the amino acid sequence of proteins. tRNAs have a length of 73–90 nucleotides (nt) and have a characteristic secondary structure in cloverleaf and a tertiary structure in the so-called “L-shape” (Fig. 1). tRNA represents one of the most abundant cellular transcripts and amount up to 15 % of the total cellular RNAs. They belong to the most stable RNAs in a cell [1] with half-lives of hours to days. Folding and stability of tRNA molecules is further controlled by numerous posttranscriptional modifications that cluster primarily in the anticodon-, the D- and T-arms [2]. Since structural and functional integrity of tRNA molecules is crucial for proper cellular functioning, many control pathways evolved, which recognize and degrade misfolded or hypomodified tRNA molecules rapidly [3].

2014 celebrates respectively the 40th anniversary of the first high resolution tRNA crystal structure [4, 5] and the 25th anniversary of the first crystal structure of a complex between an aminoacyl-tRNA synthetase and its cognate tRNA [6]. Forty years later, they are still matter of research in structural biology since advances in biochemistry, crystallization, collection and treatment of X-ray data

Eric Ennifar (ed.), *Nucleic Acid Crystallography: Methods and Protocols*, Methods in Molecular Biology, vol. 1320, DOI 10.1007/978-1-4939-2763-0_4, © Springer Science+Business Media New York 2016

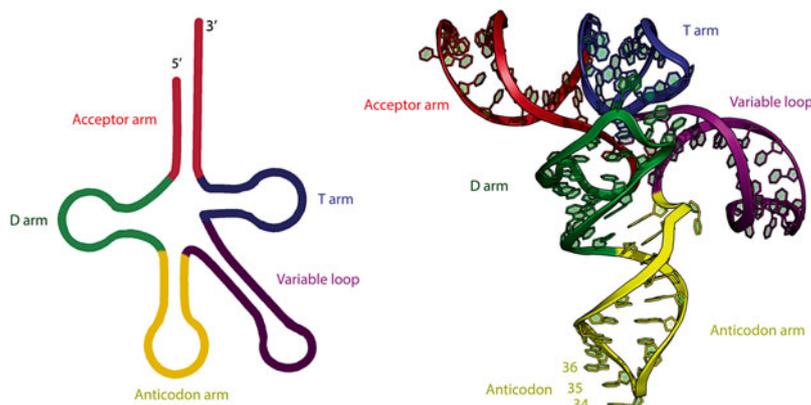


Fig. 1 Secondary and tertiary structures of tRNA^{Sec} (PDB ID 3W3S) showing the cloverleaf and the “L-shape” organization of tRNA

allow to tackle more complex biological assembly involving, for instance, tRNA in the ribosome [7–10], or in complex with enzymes involved in the processing and/or maturation of tRNA. Moreover, newly discovered functions of tRNA have boosted new structural studies with tRNA molecules. For instance, tRNAs have been shown to act as stress-sensors by regulating numerous metabolic and cellular processes, in both prokaryotes and eukaryotes. For example, in *E. coli*, uncharged tRNAs act as signaling molecules under nutritional stress, activating the stringent response pathway and thereby promoting cell survival [11]. A similar function in gene expression has been described in yeast and human cells [12]. Upon binding to T-box elements located in the 5' UTR of cistronic bacterial mRNAs, tRNAs are capable of regulating aminoacyl-tRNA synthetase genes or genes involved in amino acid biosynthesis and uptake [13, 14]. Besides, tRNAs can also serve as primers for the replication of viruses, participate in cell wall synthesis, and in the biosynthesis of certain porphyrins and amino acids [15].

These newly discovered functions involve newly discovered partners for tRNA and make that tRNA and/or tRNA co-crystallization is still relevant nowadays. In this chapter, we detail protocols used to produce tRNA by in vivo or in vitro strategies ready for crystallization assay and X-ray studies [16–19]. We will focus on the production of *B. subtilis* tRNA^{Ser} (GGA) that is 92 nucleotides long since it has a large variable loop. This tRNA is under study in one of our group as the substrate of a modifying enzyme.

2 Materials

2.1 In Vitro

Production:

Preparation of T7 RNA Polymerase

1. pAR1219 plasmid (Sigma-Aldrich, Saint Louis, MO, USA).
2. Electro-competent *Escherichia coli* cells BL21(DE3).
3. 5 ml Ni-NTA HiTrap column (GE Healthcare, Pittsburgh, PA, USA).

4. Äkta purifier (GE Healthcare, Pittsburgh, PA, USA) or similar chromatographic system.
5. 100 W sonicator.
6. LB-agar plates.
7. 2xTY culture media.
8. Luria–Bertani (LB) culture media.
9. Ampicillin.
10. Isopropyl β -D-thiogalactopyranoside (IPTG) 200 mM.
11. Buffer A: 50 mM phosphate buffer pH 8.0, 300 mM NaCl, 10 % glycerol, 0.5 mM DTT, 20 mM imidazole pH 8.0, PMSF at 20 μ g/mL.
12. Buffer B: Buffer A supplemented with 500 mM imidazole pH 8.0.
13. Buffer C: 20 mM phosphate buffer pH 7.7, 100 mM NaCl, 50 % glycerol, 1 mM DTT, 1 mM EDTA.
14. Lysis buffer: 20 mM Tris–HCl pH 8.0, 50 mM Na Phosphate pH 8.0, 20 mM NaCl, lysozyme 100 mg/mL, phenylmethane-sulfonylfluoride (PMSF) 0.1 M, and optionally: benzamidine 0.1 M, pepstatin 1 μ g/mL, leupeptin 2 μ g/mL.
15. Dialysis sacks or cassette, MWCO <50 kDa.
16. MiniProtean[®] protein gel electrophoresis system (Bio-Rad, Hercules, CA, USA).
17. SDS-PAGE gel (10 %) for MiniProtean[®].

2.2 In Vitro

Production:

Preparation of DNA Template for In Vitro Transcription

1. Oligonucleotides for the PCR.
2. PCR tubes.
3. *Pyrococcus furiosus* DNA polymerase Pfu (30 U/ μ L) in 10 \times Pfu buffer.
4. 10 \times Pfu buffer.
5. Orange G loading dye 10 \times : 0.4 % (w/v) orange G, 30 % (v/v) glycerol.
6. Thermocycler.
7. Agarose.
8. SYBR[®] Safe (Life Technologies, Carlsbad, CA, USA).
9. 10 \times TAE Buffer: 48.4 g of Tris-base, 11.4 mL of glacial acetic acid, 3.7 g of EDTA disodium salt, complete to 1 L with water.
10. Agarose gel electrophoresis system.
11. Stock solution containing a mix of each dNTP at 10 mM in water.

2.3 In Vitro

Production: T7 Transcription

1. MgCl₂ 1 M.
2. Polyethylene glycol (PEG) 4000.
3. Stock solution of rNTP mix at 100 mM.

4. In vitro transcription buffer concentrated ten-times (T10 buffer): 400 mM Tris-HCl pH 8.0, 10 mM spermidine, 50 mM DTT, 1 % Triton X-100.
5. EDTA, pH 8.0 0.5 M.
6. Sodium dodecyl sulfate (SDS) 20 %.
7. Tris-HCl, pH 8.8 1 M.
8. Tris-HCl, pH 6.8 1 M.
9. Blue loading buffer 5×: 300 mM Tris-HCl pH 6.8, 1 % (w/v) SDS, 1 % (w/v) bromophenol blue.

2.4 Gel Purification

1. 10× Tris-Borate-EDTA buffer (TBE): 890 mM Tris base, 890 mM boric acid, 20 mM EDTA pH 8.4.
2. Urea, powder.
3. 10 % (w:v) ammonium persulfate (APS) solution.
4. *N,N,N',N'*-tetramethylethylenediamine (TEMED), ≥99 %.
5. RNA loading dye: 1× TBE, 8 M urea, 0.01 % xylene cyanol and 0.01 % bromophenol blue.
6. 40 % acrylamide-bis-acrylamide 19:1 solution.
7. High-voltage power supply for gel migration system.
8. DNA sequencing-like gel migration system.
9. DNA sequencing glass plates (420×330×3.5 mm).
10. Spacers (420×20×2 mm) for DNA sequencing glass plates.
11. Large comb for DNA sequencing gel (100×265×2 mm).
12. Large plastic wrap (width larger than 50 cm).
13. Silica plate for UV shadowing.
14. Cutter blade.
15. Rotation instrument (rotating wheel or similar).
16. Syringe with needle.
17. UV lamp.
18. 0.2 % Methylene blue staining solution.
19. Steel spoon.
20. Sodium acetate pH 5.2 3 M.
21. 0.2 μm RNase-free Stericup-GP Express PLUS, polyethersulfone, 150 mL (Millipore, Billerica, MA, USA).
22. RNase-free plastic spoon.
23. DEAE sepharose resin (GE Healthcare, Pittsburgh, PA, USA).
24. Small empty column with tubing.
25. Ethanol ≥99.8 %.
26. Large binder clips.
27. Sealing tape, 48 mm width.

4. In vitro transcription buffer concentrated ten-times (T10 buffer): 400 mM Tris-HCl pH 8.0, 10 mM spermidine, 50 mM DTT, 1 % Triton X-100.
5. EDTA, pH 8.0 0.5 M.
6. Sodium dodecyl sulfate (SDS) 20 %.
7. Tris-HCl, pH 8.8 1 M.
8. Tris-HCl, pH 6.8 1 M.
9. Blue loading buffer 5×: 300 mM Tris-HCl pH 6.8, 1 % (w/v) SDS, 1 % (w/v) bromophenol blue.

2.4 Gel Purification

1. 10× Tris-Borate-EDTA buffer (TBE): 890 mM Tris base, 890 mM boric acid, 20 mM EDTA pH 8.4.
2. Urea, powder.
3. 10 % (w:v) ammonium persulfate (APS) solution.
4. *N,N,N',N'*-tetramethylethylenediamine (TEMED), ≥99 %.
5. RNA loading dye: 1× TBE, 8 M urea, 0.01 % xylene cyanol and 0.01 % bromophenol blue.
6. 40 % acrylamide-bis-acrylamide 19:1 solution.
7. High-voltage power supply for gel migration system.
8. DNA sequencing-like gel migration system.
9. DNA sequencing glass plates (420×330×3.5 mm).
10. Spacers (420×20×2 mm) for DNA sequencing glass plates.
11. Large comb for DNA sequencing gel (100×265×2 mm).
12. Large plastic wrap (width larger than 50 cm).
13. Silica plate for UV shadowing.
14. Cutter blade.
15. Rotation instrument (rotating wheel or similar).
16. Syringe with needle.
17. UV lamp.
18. 0.2 % Methylene blue staining solution.
19. Steel spoon.
20. Sodium acetate pH 5.2 3 M.
21. 0.2 µm RNase-free Stericup-GP Express PLUS, polyethersulfone, 150 mL (Millipore, Billerica, MA, USA).
22. RNase-free plastic spoon.
23. DEAE sepharose resin (GE Healthcare, Pittsburgh, PA, USA).
24. Small empty column with tubing.
25. Ethanol ≥99.8 %.
26. Large binder clips.
27. Sealing tape, 48 mm width.

12. RNA storage buffer: 50 mM Tris-HCl pH 8.0, 100 mM KCl and 10 mM MgCl₂.
13. Amicon Ultra-15 concentrators, 10000 MWCO (Millipore, Billerica, MA, USA).

2.8 NMR Spectroscopy

1. NMR Buffer: 10 mM KPO₄ pH 6.5, 50 mM KCl, 10 mM MgCl₂.
2. 3 mm NMR tube.
3. D₂O 99.9 %, 15 µL.
4. NMR spectrometer (600 MHz or equivalent).

2.9 Crystallization

1. Cristalquick, 96-well plates (Greiner Bio-One, Frickenhausen, Germany) or equivalent.
2. Natrix sparse-matrix kit (Hampton Research, Aliso Viejo, CA, USA) or Nucleix sparse-matrix kit (Qiagen, Venlo, The Netherlands).

3 Methods

It is important to always use talc-free gloves during sample preparation and manipulation as human skin is rich of RNases. For materials that cannot be exclusively used for RNA preparation or than are not single use, wash them when possible with 1 M NaOH and rinse extensively with bi-distilled water. When adjusting pH of an RNase-free solution, avoid inserting a pH meter electrode into the solution; instead prefer paper pH-indicator strips. Store RNase-free solution at low temperature or frozen when possible, and keep them sterile. The use of bi-distilled water (18.2 MΩ) is sufficient to assure RNase-free condition, provided that the water purification system is well-maintained.

3.1 In Vitro Production: Preparation of T7 RNA Polymerase

1. Transform by electroporation BL21(DE3) competent cells with the pAR1219 plasmid expressing the T7 RNA polymerase (98 kDa). Add 1 mL of LB medium and incubate the cells at 37 °C for 15 min. Spread 10–100 µL of this solution on a LB-agar plate supplemented with 50 µg/mL of ampicillin. Incubate the plate at 37 °C overnight.
2. The next evening, start a pre-culture (5 mL culture in LB supplemented with 50 µg/mL of ampicillin) inoculated with one isolated colony of bacteria containing the pAR1219 plasmid. Incubate overnight at 37 °C.
3. Start a 1 L culture in 2xTY medium supplemented with ampicillin (50 µg/mL) with the necessary volume of the overnight culture to obtain a starting optical density at 665 nm (OD₆₆₅) around 0.05. Incubate the remaining culture at 37 °C with

Chapitre III

Production d'un substrat de TrmK : l'ARNtSer de *B. subtilis* –
Article 2: Dégut et al. (2016)

vigorous shaking. When the OD_{665} reached 0.6, add 1 mM (final concentration) of IPTG to induce the expression of T7 RNA polymerase. Let the bacteria grow for 3–4 h.

4. Pellet the bacteria by centrifugation ($6,000 \times g$ for 30 min).
5. Resuspend the pellet in the lysis buffer. Sonicate the cell suspension with ten short burst of 10 s (45 W) followed by intervals of 30 s for cooling. Keep the suspension on ice as much as possible. Remove cell debris by centrifugation at 4 °C, 30 min at $20,000 \times g$.
6. Purify the T7 RNA polymerase at 4 °C using a chromatographic system: equilibrate the Ni-NTA column with 5 volumes of buffer A. Inject the supernatant at 2 mL/min, wash the column with buffer A until the baseline is recovered to be sure that all the flow-through is gone. Then, elute the T7 RNA polymerase using a 100 mL gradient (from 0 to 100 % of buffer B). Collect 2 mL fractions. Analyze the fractions on a SDS-PAGE gel (10 %). Pool the fractions that contain the T7 RNA polymerase and dialyze at 4 °C against 2 L of buffer C. The volume in the dialysis chamber will be roughly reduced by a factor 3.
7. Recover the solution and aliquot the T7 RNA polymerase by 1 mL in microtubes for long-term storage at -80 °C. Approximately, 7 mL of T7 RNA polymerase (15 mg/mL) can be obtained from 1 L of culture.

**3.2 In Vitro
Production:
Preparation of DNA
Template for In Vitro
Transcription**

1. DNA template for in vitro transcription of tRNA^{Ser} was designed as described in Fig. 2, using two consecutive PCR (Polymerase Chain Reaction) (*see Note 1*). The construct must include the T7 promoter sequence upstream the tRNA gene (in bold in Fig. 2a). The first two oligonucleotides must cover all the sequence of the DNA template and should be designed so that they share complementary sequences between the 3' of oligonucleotide 1 and the 5' of oligonucleotide 2 (Fig. 2b). We recommend at least 20 base pairs between oligonucleotides 1 and 2. The two other oligonucleotides are primers for classic PCR that allows the amplification of the DNA template.
2. For the first PCR (Fig. 2b), mix 40 μ L of 10 \times Pfu buffer, 8 μ L of the dNTP mix, 1 μ L of oligonucleotides 1 and 2 (previously dissolved at 100 μ M in water), 8 μ L of Pfu DNA polymerase and complete the volume to 400 μ L with water. Aliquot the reaction in PCR tubes (50 μ L each). On a thermocycler, start with the first step of denaturation by heating to 95 °C for 1.5 min; then repeat two times the following cycle: 35 s of denaturation at 95 °C, 35 s of annealing at 55 °C, and 30 s of elongation at 72 °C. Carry on the reaction by a final step of elongation at 72 °C for 5 min and then hold the samples at 4 °C.

A T7 DNA template

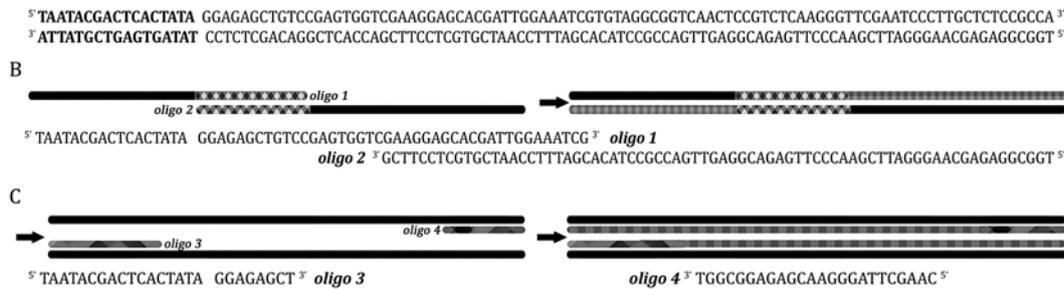


Fig. 2 Design of the DNA template for in vitro T7 transcription. The gene encoding *B. Subtilis* tRNA^{Ser} (GGA) has been assembled by two consecutive PCR. (a) DNA template of *B. Subtilis* tRNA^{Ser}. (b) Oligonucleotide 1 and 2 are needed during the first PCR to generate the double-stranded DNA template (c) Amplification of the DNA template by a second PCR using oligonucleotides 3 and 4

- For the second PCR (Fig. 2c), mix 120 μ L of 10 \times Pfu buffer, 24 μ L of the dNTP mix, 12 μ L of oligonucleotides 3 and 4 (previously dissolved at 100 μ M in water), 1 μ L of the PCR1 product, 24 μ L of Pfu DNA polymerase and complete the volume of reaction to 1.2 mL with water. Aliquot the reaction in PCR tubes (50 μ L each). Repeat 25 times the PCR cycle described in **step 2**. This PCR product can be used immediately as DNA template for in vitro T7 transcription without any further treatment. For long-term conservation, keep it at -20°C .
- Analyze the PCR product on a 2 % agarose gel stained with SYBR[®] Safe. Weight 0.5 g of agarose, add 2.5 mL of 10 \times TAE buffer, and complete to 25 mL with water. Melt the agarose powder until complete dissolution in a microwave at 20 % power. Let cool down the solution until you can hold the vessel with your bare hands and add 2 μ L of SYBR[®] Safe. Pour the solution in an agarose gel mold. Put the comb and wait for gel solidification. Take 10 μ L of your PCR reaction and add 5 μ L of orange G loading dye. Run the gel at 100 V for 15 min in 1 \times TAE buffer. Orange G dye runs like a 50 base-pair DNA fragment, so the PCR product (109 base pairs) should be slightly above.

3.3 In Vitro Production: T7 Transcription

The optimal in vitro transcription conditions depend on the concentrations of the DNA template, rNTP and/or MgCl₂ concentrations (Fig. 3). These concentrations should be adjusted to optimize the yield of the RNA of interest. It is recommended to test small-scale reactions (currently 100 μ L) before large-scale production of RNA (5–20 mL). In practice, we usually vary the concentration of MgCl₂ in small-scale reactions (Table 1) and analyzed the yield of transcript RNA by SDS-PAGE (*see* **Note 2** and Fig. 3). The conditions described in Table 1 are good starting point. Concentrations

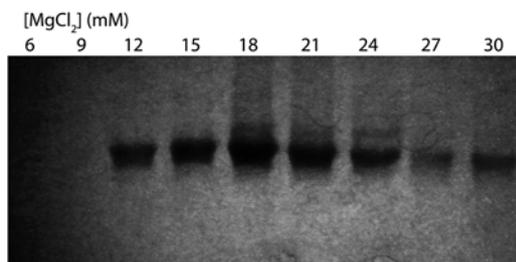


Fig. 3 Sensitivity to the magnesium concentration of the RNA production by in vitro T7 transcription. SDS-PAGE gel at 14 % revealed by UV-shadowing showing the analysis of small-scale assays of Table 1 testing the production of tRNA^{Ser} by T7 transcription

Table 1
Examples of small-scale assays used to define the optimal MgCl₂ concentration

Assays	1	2	3	4	5	6	7	8	9
[MgCl ₂] (mM)	6	9	12	15	18	21	24	27	30
DNA template	10 μL								
T10 buffer	10 μL								
rATP (100 mM)	4 μL								
rUTP (100 mM)	4 μL								
rGTP (100 mM)	4 μL								
rCTP (100 mM)	4 μL								
MgCl ₂ (300 mM)	2 μL	3 μL	4 μL	5 μL	6 μL	7 μL	8 μL	9 μL	10 μL
H ₂ O	58 μL	57 μL	56 μL	55 μL	54 μL	53 μL	52 μL	51 μL	52 μL
T7 RNA polymerase (1/10)	4 μL								

of T7 RNA polymerase and DNA template can also be further optimized. For small RNA (<50 nucleotides), the addition of 5 % (w/v) final of PEG 4000 should be tested.

1. To investigate the impact of MgCl₂ concentration on the yield of RNA, the mixtures detailed in Table 1 should be prepared in microtubes. The T7 RNA polymerase is used diluted tenfold compared to the stock obtained in Subheading 3.1, step 7.
2. Run the test reactions for 3–4 h at 37 °C. The appearance of a white precipitate due to magnesium pyrophosphate is most of the time visible and can be used to monitor the transcription efficiency.
3. Store at –20 °C or analyze immediately by SDS-PAGE. Alternatively, the reaction can be stopped by adding 10 μL of 0.5 M EDTA pH 8.0.

Chapitre III

Production d'un substrat de TrmK : l'ARN^tSer de *B. subtilis* –
Article 2: Dégut et al. (2016)

4. To analyze the products of transcription, prepare a SDS-PAGE gel using the same materials as those used for protein SDS-PAGE analysis. The percentage of gel depends on the length of the expected product, i.e., a 14 % gel for a 92 nucleotides tRNA. For 2 gels running back to back, mix 3.5 mL of acrylamide–bis-acrylamide 19:1, 3.75 mL Tris–HCl pH 8.8, 50 μ L SDS 20 % and adjust the volume to 10 mL with water. Add 140 μ L of APS 10 % and 10 μ L of TEMED just before placing the gel between the glass plates. Leave a sufficient space at the top of the glass plates to pour then the stacking gel. For the stacking gel, mix 0.5 mL of acrylamide–bis-acrylamide 19:1, 630 μ L of Tris–HCl pH 6.8, 50 μ L of SDS 20 % and complete to 5 mL with water. Add just before pouring, 100 μ L of APS 10 % and 5 μ L of TEMED. Put the comb immediately. The stacking gel should take one third of the total volume of the gel.
5. Take an aliquot of each small-scale reaction (typically 5 μ L) and add 5 μ L of blue loading buffer. Do not heat the samples before loading them on the gel. Load your samples on the wells and run the gel for 1 h at 24 mA.
6. The visualization is then done by UV-shadowing. Carefully remove one of the glass plate, leaving the gel on the other one. Put a clear plastic wrap over the gel and pull carefully the gel away from the second glass plate. Put the plastic wrap bearing the gel on a piece of silica plate and then illuminate the gel with a portable UV lamp (254 nm). You should see clear shadows that correspond to the RNA in the gel. The gel corresponding to the assays described in Table 1 and revealed by UV-shadowing is shown in Fig. 3. The best yield in our case is obtained for a concentration of MgCl_2 around 18 mM. This concentration is thus used for large-scale preparation of this tRNA. The yield of RNA is very sensitive to the concentration of MgCl_2 . We recommend carrying out these small-scale assays each time that a new preparation of DNA template is used, even if the protocol of preparation is the same.

3.4 In Vitro

Production: tRNA

Purification Using

Denaturing

Polyacrylamide Gel

For large-scale in vitro transcription reaction (more than 5 mL), sequencing-like gel migration system can be adapted to purify the RNA transcripts. It is often the best method to obtain a homogeneous sample of tRNA ready for crystallization assay. The problem with in vitro transcription is that it generates $n-1$ and/or $n+1$ product in addition to the desired RNA. These species ($n-1$, $n+1$) can be a bottleneck to get diffracting crystals (*see* Note 3).

1. One volume of RNA loading buffer is added to one volume of transcription reaction. For really large transcription (up to 25 mL), in order to minimize the loading volume, a 8 M urea concentration can be possibly obtained by adding the

corresponding urea mass as a powder to the transcription reaction and a 50-fold concentrated xylene cyanol and bromophenol blue solution.

2. The DNA sequencing gel glass plates and spacers are washed and assembled using clamps and tape (at the bottom), in RNase-free conditions. To prepare 300 mL of 12 % polyacrylamide and 8 M urea gel, dissolve 144 g urea in 30 mL 10× TBE and 90 mL 19:1 polyacrylamide–bis-acrylamide. Add just before pouring, 5 mL of APS 10 % and 250 µL of TEMED. Put the large comb for DNA sequencing gel immediately before the gel has polymerized. After polymerization, the well is washed using a syringe with a needle and 1× TBE buffer, and the transcription solution containing the RNA loading buffer is loaded on the gel. It is recommended to check the *in vitro* transcription efficiency on small 8 M urea gel or SDS-PAGE gel before large-scale assays. The appropriate percentage of 19:1 polyacrylamide gel is determined according to the size of the RNA molecule (Table 2). The gel is pre-run for 2 h in 1× TBE running buffer to reach an optimal migration temperature. Set the power to 65 W for one large gel and 80 W for two gels running together. The well is washed with running buffer before loading the transcription reaction. Syringe and needle are used for these two steps. The migration runs for 6–8 h depending on the percentage of polyacrylamide gel.
3. There are two options to visualize the band corresponding to the tRNA molecule:
 - (a) Using the UV-shadowing approach. The gel is removed from the system, protected with a thin plastic wrap, placed on a silica plate and exposed to UV (wavelength close to 254 nm) using a short-wave UV lamp. Silica reflects the UV while the RNA bases absorb UV. A shadow corresponding to the tRNA band can thus be observed on the silica.
 - (b) Using methylene blue. The gel is removed from the system. A small band is cut along the gel, stained with a 0.2 %

Table 2
Percentage of 19:1 polyacrylamide–bis-acrylamide gel according to the size of the RNA

RNA molecule (nucleotides)	19:1 Acrylamide (%)	10 mL gel (mL)	50 mL gel (mL)	300 mL gel (mL)
Less than 50	15	3.75	18.75	112.5
50	12	3	15	90
100	10	2.5	12.5	75
More than 100	8	2	10	60

For each percentage, the volumes of acrylamide–bis-acrylamide 19:1 have been indicated to prepare a 10 mL, 50 mL, or 300 mL gel solution, respectively

Chapitre III

Production d'un substrat de TrmK : l'ARNtSer de *B. subtilis* –
Article 2: Dégut et al. (2016)

methylene blue solution for few seconds and destained with water. In that case, a small quantity of tRNA is lost, compared to the first option.

4. The band corresponding to the tRNA is excised with a blade and placed into a 50 mL Falcon tube. The gel is then crushed with an RNase-free steel spoon. tRNA is eluted by diffusion with 40 mL of 0.3 M sodium acetate pH 5.2 on the rotating wheel at 4 °C for several hours. The time required depends on the size of the RNA, the acrylamide gel concentration and the size of the crushed gel particles. The eluate is filtered on a 0.2 µm RNase-free Stericup-GP Express PLUS. Acrylamide crumbs are recovered with an RNase-free plastic spoon and put back in the same 50 mL Falcon tube for a second elution. This step can be repeated once.
5. The sample (in 0.3 M sodium acetate pH 5.2) is concentrated using DEAE sepharose resin and a small column. A small column apparatus was linked by tubing to a 50 mL syringe that is used as a reservoir. Before use, the homemade system is cleaned with 1 M NaOH and rinsed extensively with water. 1 mL of DEAE sepharose resin is packed into the column. The resin is equilibrated with 0.3 M sodium acetate pH 5.2. The sample is then loaded and the system is washed with 10 mL of 0.3 M sodium acetate pH 5.2. The tRNA is finally eluted with 7 mL of 3 M sodium acetate pH 5.2 and ethanol-precipitated by adding 21 mL of ethanol. The sample is then placed at -20 °C.

3.5 In Vitro Production: tRNA Concentration and Folding

1. After ethanol precipitation, the tRNA is centrifuged at 4 °C and 6,000 × g. The pellet is then washed three times with 70 % ethanol, dried under vacuum, and dissolved in 1 mL of pure water for 1 mL of transcription reaction. For the tRNA folding, the sample is incubated for 1 min at 85 °C and slowly (i.e., 45 min) cooled to room temperature by switching off the heating of water bath. The tRNA can be further concentrated, and buffer exchanged by using an Amicon centrifugal concentration unit. The sample can be stored in RNA storage buffer. Typically, a concentration of 10 mg/mL is a good stock concentration to conduct crystallization assays with this tRNA. Adding 10 mM MgCl₂ in the final buffer can greatly improve tRNA folding and is sometimes necessary to stabilize the tertiary interactions between the D- and T-loops. Keep your sample at -20 °C.
2. The folding of tRNA can be analyzed by NMR spectroscopy (see **Note 4** and Fig. 4). Put the tRNA in the NMR buffer at a concentration around 0.2 mM in a 3 mm NMR tube. In this kind of tube, only 150 µL of sample is needed. Put 135 µL of tRNA and 15 µL of D₂O. Record a 1D proton spectrum using a watergate sequence to suppress the water signal [20]. The imino protons are carried by G and U nucleotides and resonate

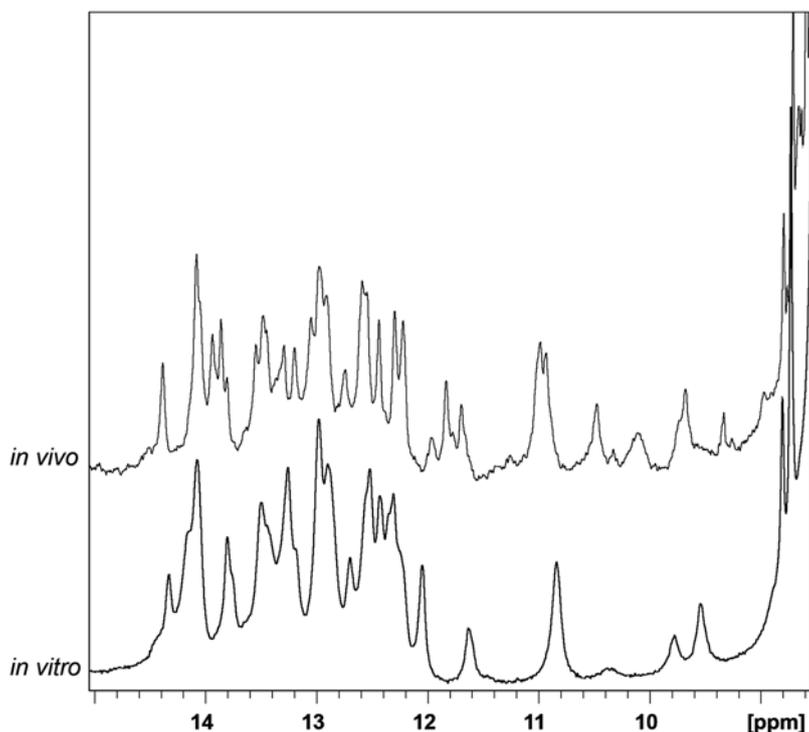


Fig. 4 1D ^1H NMR spectra of tRNA^{Ser} produced by *in vitro* transcription (*lower spectrum*) and *in vivo* in *E. coli* (*top spectrum*). The spectra were recorded at 35 °C on a Bruker DRX 600 MHz NMR spectrometer. The samples were put in a 10 mM KPO_4 pH 6.5 with 50 mM KCl and 10 mM MgCl_2

between 10 and 15 ppm. They constitute the signature of base-pair within a folded RNA. The NMR spectrum of tRNA^{Ser} imino protons produced *in vitro* is consistent with one molecular species and showed about 25 peaks (uncertainty due to overlap), which is consistent with the number of base-pairs expected in the 3D structure of this tRNA. These data support folding of the tRNA into its correct tertiary structure.

3.6 *In Vivo* Production: Cloning a tRNA Gene in the pBSTNAV Vector

tRNA can also be easily produced *in vivo* in *E. coli* using the pBSTNAV vector [16, 17, 19, 21]. The cloning procedure of the tRNA^{Ser} gene in this plasmid and the expression and purification of tRNA^{Ser} are described in the following paragraphs.

1. Phosphorylate oligonucleotides 2, 3, 5 (Fig. 5). In a micro-tube, put 1 μL of each oligonucleotide previously diluted at a concentration of 100 μM in water. Add 10 U of T4 polynucleotide kinase 3'-phosphatase free and 1.5 μL of the buffer recommended by the manufacturer. Complete with water to 15 μL (final volume of the reaction). Incubate at 37 °C for 30 min. Leave 10 min at 65 °C to inactivate the kinase.

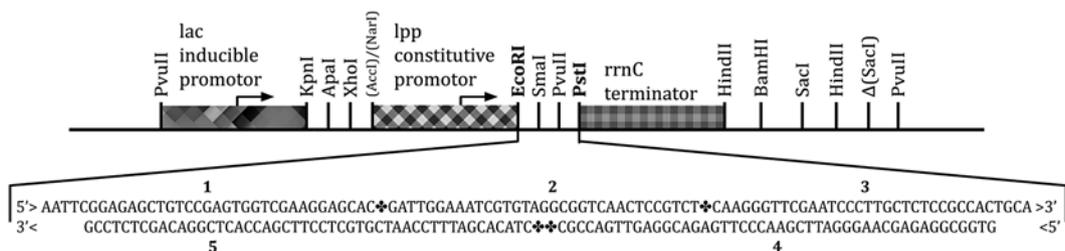


Fig. 5 Schematic representation of the cloning region of pBSTNAV plasmid and sequence of DNA coding for *B. subtilis* tRNA^{Ser} (GGA) inserted between EcoRI and PstI restriction sites. In pBSTNAV, the lpp promoter upstream of the tRNA gene, one of the strongest natural *E. coli* promoters that controls the transcription of the mRNA of the lipoprotein, allows production of high levels of recombinant transcripts. Accumulation of stable recombinant RNA throughout the transcription phase can be achieved by the combined use of a strong promoter and a multicopy plasmid. Yields reach a maximum within the first few hours of the stationary phase

2. In the microtube, add in the 1 μL of oligonucleotides 1 and 4 previously diluted at a concentration of 100 μM in water. To hybridize the oligonucleotides as indicated in Fig. 4 in order to make the insert, put the previous mix 2 min at 95 $^{\circ}\text{C}$, let them cool down slowly to 37 $^{\circ}\text{C}$. Add 10U of T4 DNA ligase in the appropriate buffer that should contain ATP. Complete with water to a volume of reaction of 30 μL . Incubate for 1 h and 30 min at 37 $^{\circ}\text{C}$.
3. Preparation of the pBSTNAV plasmid. To linearize the pBSTNAV plasmid, digest it with PstI and EcoRI restriction enzymes. For 10 μg of plasmid, add 3 μL of PstI-HF and EcoRI-HF, 10 μL of NEB CutSmart buffer and complete the volume of the reaction with water to reach a volume of 100 μL . Incubate 1 h30 at 37 $^{\circ}\text{C}$. Add 250 μL of ethanol and 10 μL of NaCl 5 M and pellet plasmid by centrifuging the mix for 15 min at 4 $^{\circ}\text{C}$ and at 7,000 $\times g$. Remove carefully the ethanol and dry the sample in a vacuum concentrator. Dissolve the linearized plasmid in 30 μL of water.
4. To insert and ligate the oligonucleotide containing the tRNA^{Ser} gene into the linearized pBSTNAV plasmid, mix 2 μL of the linearized plasmid, 15 μL of the insert, 4 μL of the T4 DNA ligase, 3 μL of the manufacturer recommended T4 ligase buffer (X10), and complete the volume to 30 μL . Incubate one night at 14 $^{\circ}\text{C}$. Put the mix for 2 min at 95 $^{\circ}\text{C}$ to inactivate the ligase.
5. Counter-selection by SmaI (*see Note 5*). After the ligation step, add 100 μL ethanol and 3 μL NaCl (5 M) and recover the pellet DNA by centrifuging for 15 min at 4 $^{\circ}\text{C}$ and 7,000 $\times g$. Redissolve the plasmid in 10 μL of water. Add 3.5 μL of SmaI and 1.5 μL of NEB CutSmart buffer. Incubate at 20 $^{\circ}\text{C}$ for 1 h 30 min. Then, add 50 μL of ethanol and 1 μL NaCl 5 M, centrifuge for 15 min at 4 $^{\circ}\text{C}$ and 7,000 $\times g$. Carefully remove

the supernatant and wash by ethanol 70 % (add 50 μ L of cold ethanol 70 % mix and centrifuge for 15 min at 4 °C and 7,000 $\times g$). After the wash, remove as much as possible ethanol and dry the sample in a vacuum concentrator. Dissolve DNA in 15 μ L of water and store at -20 °C.

3.7 In Vivo

Production:

Expression in *E. coli* and Total tRNA

Extraction

1. Transform *E. coli* electro-competent cells (JM101TR strain) by electroporation. The bacteria are grown overnight at 37 °C in LB medium supplemented with 50 μ g/mL of ampicillin (*see Note 6*).
2. Recover the bacteria by centrifugation for 30 min at 6,000 $\times g$.
3. Resuspend the bacteria in 10 mL of buffer D. Add 10 mL of phenol saturated in the same buffer.
4. Mix slowly 1 h at room temperature. Centrifuge for 30 min, 8,000 $\times g$ at 15 °C. Carefully recover the soluble phase.
5. Do a counter-extraction with 2 mL of buffer d. Recover again the soluble phase and add 0.1 volume of NaCl 5 M and 2 volumes of ethanol to the soluble phase that contains total nucleic acids extracted from the bacteria. Centrifuge for 30 min, 8,000 $\times g$ at 4 °C.
6. Resuspend the pellet in 5 mL of NaCl 1 M (*see Note 7*). Centrifuge for 30 min at 4 °C and at 8,000 $\times g$. Recover the supernatant and add 2.5 volume of ethanol. Centrifuge for 30 min, 8,000 $\times g$ at 4 °C. Discard the supernatant and resuspend the pellet in 2 mL of Tris-HCl pH 8.0 2 M and incubate the mix at 37 °C for 2 h to insure a full deacylation of the tRNA molecules (*see Note 8*).
7. Precipitate the total extract of tRNA by adding 2 volumes of ethanol. Centrifuge for 30 min, 8,000 $\times g$ at 4 °C. Redissolve the pellet in RNA storage buffer (or equivalent).

3.8 In Vivo

Production: tRNA Purification Using Chromatography Columns

tRNAs produced in vivo in *E. coli* have the advantages to be correctly folded and matured by the enzymes of the host. Therefore, even if the purification by denaturing polyacrylamide gel is still an option, it could be advantageous to optimize a protocol of purification using anion-exchange and hydrophobic chromatography. This protocol is detailed below.

1. First, wash the ÄktaPurifier (or other FPLC purification system) and the columns with 1 M NaOH and rinse them extensively with bi-distilled water to work in RNase-free conditions. After the total tRNA extraction (*see Subheading 3.7*), load the tRNA sample on a Mono Q HR 10/10 column (8 mL) previously equilibrated with buffer A. Wash by several volume of column with buffer A. Then, separate the tRNAs using a gradient from 35 to 55 % of buffer B with a length of 240 mL and a flow of 4 mL/min. Collect the eluate putting 2 mL/fraction.

Noticeably, the gradient has to be optimized according to the size and the sequence of the tRNA to purify.

- You then need to analyze the content of the fractions to identify the fractions containing the overexpressed tRNA. One simple experiment is to perform gel-shift assay using SDS-PAGE (Fig. 5). First, design a DNA that is complementary to the overexpressed tRNA in a region that is not well-conserved in the other *E. coli* tRNAs. In our case, we select a DNA complementary to the anticodon arm and the large variable loop (Fig. 6). Take 5 μ L of each fraction to analyze and mix with 1 μ L of the complementary DNA (previously dissolved in water at 100 μ M). In a thermocycler, heat the samples at 95 °C for 2 min and cool down to 4 °C. The loading blue is added after this operation, the shift will not appear if it is present when samples are heated.
- Load the samples on a 14 % SDS-PAGE gel (*see* Subheading 3.3, **step 4**). Visualize the RNA present in each fraction by UV-Shadowing. tRNA^{Ser} can be easily discriminated from *E. coli* tRNAs by a shift of its species in presence of the complementary DNA (arrows in Fig. 5.). Indeed, the band of tRNA^{Ser} has disappeared in the presence of the DNA and a band of higher molecular weight has appeared.
- Pool the fractions that contain the tRNA of interest. The next step consists in further purifying the tRNA by hydrophobic chromatography. First, put in your tRNA sample the quantity of ammonium sulfate in powder necessary to have your sample in 1.7 M ammonium sulfate before loading it on the Phenyl

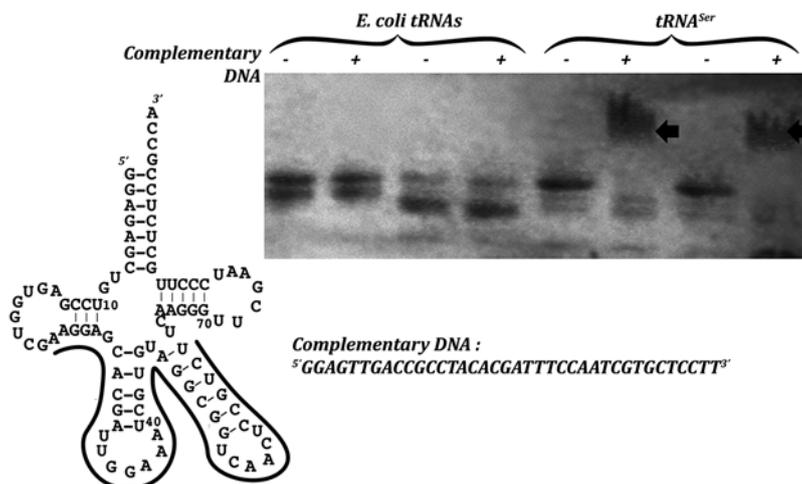


Fig. 6 Gel-shift assays (SDS-PAGE 14 %) realized on fractions obtained after the Mono Q column. This simple experiment allows to discriminate the tRNA of interest from those of *E. coli* using an hybridization test of a complementary DNA to tRNAs present in the different fractions

Superose. Equilibrate the Phenyl Superose with buffer B2. Then, separate the tRNAs using a gradient from 10 to 45 % of buffer A with a length of 240 mL and a flow of 0.7 mL/min. Collect 1 mL/fraction.

5. Analyze the fractions with a 14 % SDS-PAGE gel (*see* Subheading 3.8, step 3). The high concentration of ammonium sulfate (>1 M) prevents the good migration of RNA samples in SDS-PAGE. You should take 5 μ L of each fraction and dilute them in 20 μ L of water to avoid migration problem. Visualize the bands by UV-shadowing. Pool the fractions that contain the tRNA of interest (*see* Note 9).
6. The tRNA of interest has then to be desalted to get rid of the large amount of ammonium sulfate. This can be done by extensive dialysis in RNA storage buffer or by using a desalting column (HiPrep 26/10 desalting column, GE healthcare). Concentrate to the desired concentration using an Amicon centrifugal concentrator. Typically, a concentration of 10 mg/mL is a good storage concentration to conduct crystallization assays with this tRNA. Keep your sample at -20°C .
7. Optionally, you can check the homogeneity of your tRNA sample and its folding by recording a 1D NMR spectrum (Fig. 4). The NMR spectrum of tRNA^{Scr} produced *in vivo* compared to that produced *in vitro* is noticeably different.

3.9 tRNA Crystallization Free or in Complex

1. Prepare the tRNA sample as described in previously.
2. Automated initial crystallization screen can be performed using a robot using 96-well microplates. For initial screening, sparse matrix crystallization screens (Nucleix or Natrix) give good results. Several concentrations of tRNA should be tested (typically 3 mg/mL, 5 mg/mL and 10 mg/mL) (*see* Note 11). The robot mixes the same volume (0.7 μ L in our case) of tRNA sample and of the reservoir solution. Screening of crystallization conditions is carried out with the sitting drop, vapor diffusion technique. Put the crystallization plate in an incubator set at a temperature between 4 and 25 $^{\circ}\text{C}$. We usually tested 4 and 18 $^{\circ}\text{C}$.
3. Check for crystals the next day, in 3 days, and then every week. Crystals may appear after several hours, after several months, or never. It is important to check for crystals on a regular basis as they can disappear. If precipitate is observed in most of the crystallization condition, then proceed to new assays but lower the concentration of tRNA.
4. When crystals are observed (Fig. 7), optimize crystallization conditions by varying the ratio between the volumes of the sample and that of the reservoir solution in the drop, the concentration of tRNA in the drop, the concentration of the precipitant or salts in the reservoir solution, the buffer, the pH,

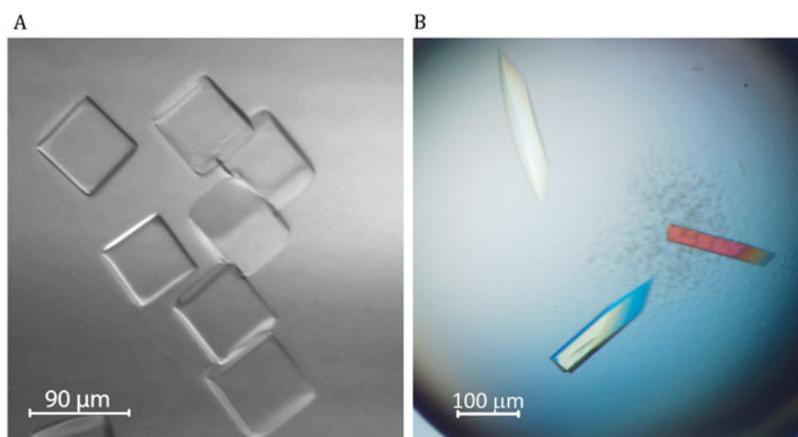


Fig. 7 Examples of crystals obtained for (a) tRNA^{Ser} produced in vivo and (b) tRNA^{Leu} produced in vitro and co-crystallized with LeuRS (see **Note 10**) [18]. The conditions of crystallization are 100 mM HEPES pH 7.5 and 1.5 M sulfate lithium for tRNA^{Ser} and 0.2 M sodium Nitrate, 20 % PEG 3350 for tRNA^{Leu} in complex with LeuRS

various additives, For tRNA^{Ser}, the diffraction quality of crystals was improved by modifying the precipitant and its concentration in 2 μ L drops.

4 Notes

1. In vitro run-off transcription with T7 RNA polymerase [22] is commonly used to prepare milligram quantities of RNA required for crystallization studies. For efficient transcription, the promoter-region of the DNA template should be double-stranded and the first nucleotide in the RNA product should be a guanine. T7 RNA polymerase tends to add one or more random nucleotides (usually adenosine) at the 3'-end of the run-off transcript, leading to problems with 3' homogeneity. For tRNA run-off transcription, the DNA template can be either a purified PCR product containing the tRNA gene and the T7 promoter or a linearized plasmid. In the case of tRNA^{Ser}, the higher yield of production was obtained for a DNA template that is a PCR product. We thus detail this protocol.
2. For the use of a linearized plasmid, the template plasmid must contain a restriction site incorporated at the 3'-end of the desired RNA sequence. For tRNA, the commonly used restriction enzyme is BstNI that allows to generate a CCA sequence at the 3'-end of the transcript tRNA [23]. The vector is usually a pUC18 or pUC19. When we used a PUC18 vector carrying the tRNA^{Ser} gene linearized with BstNI restriction enzyme, other large species of transcript products were observed. This feature reduced the quantity of tRNA transcript.

3. Although SDS-PAGE is commonly restricted to protein electrophoresis, this technique is well suited for RNA production analysis [16, 17, 19, 21]. Gel migration is much faster, bands are well-defined and the migration is not sensitive to salts present in the samples like in urea PAGE. Furthermore, these gels are not denaturing for RNA and thus hybridized RNA or RNA/DNA complexes can be detected if they are quite stable (i.e., if they involved a large number of base pairs, more than 10). We used this feature to easily detect the fractions containing the overexpressed tRNA during chromatographic steps of purification (Fig. 6).
4. Homogeneous samples are crucial for the success of crystallization. Crystals containing tRNA molecules are very often packed in such a way that two tRNA molecules form a pseudo-continuous helix with their 3'-ends [16, 17, 19, 21]. To facilitate such packing, it is important to ensure that 3'-end of tRNA are homogeneous. Unfortunately, run-off transcription with T7 RNA polymerase usually results in product RNA that are heterogeneous at its 3'-end. The gel purification of the product of in vitro transcription enables us to eliminate unincorporated NTP, products of aborted transcription, template DNA and most importantly serves to produce RNA with homogeneous ends. A nucleotide-level resolution on a preparative gel can be obtained. In case of heterogeneity at the 3'-end of the tRNA, we recommend to avoid overloading of the sample. Several bands can be observed and it could be interesting to cut all major bands to use them in crystallization trials, of course separately!
5. The refolding of the tRNA prior to crystallization assays is very important, especially for large molecules. The ability of magnesium to cut RNA at high temperatures has to be taken into account. Usually, the high temperature denaturation of RNA is performed without magnesium and it is added at a later step when the sample is cooled down at room temperature.
6. The counter selection by SmaI is essential to avoid false positive clones, i.e., clones transformed by the pBSTNAV vector that does not carry the insert sequence encompassing tRNA^{Ser} gene.
7. Since high yields of recombinant tRNA are required, it is quite important that fast growth rates are achieved during expression. Expression in rich medium at 37 °C should thus be favored, as it will support the fastest growth rates. For large-scale expression, use 2xTY medium supplemented with 0.4 % (w/v) glucose in a fermentor equipped with pH and dissolved oxygen (pO₂) control. Glucose contributes to increasing the growth rate, but requires real-time neutralization of acids produced by fermentation and hence must be done into a fermentor. Optimal aeration is also a key factor, most importantly at high

Chapitre III

cell densities. For medium scale expression, 2xTY cultures (without glucose) in flat culture flasks (Fernbach) are a simpler yet quite efficient alternative. With the strong lpp promoter on a multicopy plasmid, stable recombinant tRNA will accumulate during all their transcription phase. Yield will therefore be maximum in the first few hours of the stationary phase. This leads to a very straightforward protocol: liter-sized culture flasks are typically inoculated at low density with freshly transformed bacteria and allowed to grow overnight (15 h) at 37 °C under vigorous shaking. Putting only 500 mL of 2xTY medium in liter-sized culture flasks allows a better oxygenation of cells and thus provides a simple way to improve yield.

8. Only the smallest nucleic acids like tRNAs and 5S RNA will be soluble in 1 M NaCl. This step is not obligatory but it enables to get rid of the largest nucleic acids like chromosomes for instance.
9. This step enables the nucleophilic deacylation of the 3'-end of tRNA.
10. 10–20 mg of pure tRNA per liter of culture can be routinely obtained. The yield and the purity will depend on the tRNA sequence and notably on the posttranscriptional modifications carried by the overexpressed tRNA. Complex modifications, notably in the anticodon, can be incorporated at a substoichiometric level [16]. This can be a source of heterogeneity preventing crystallization.
11. For the crystallogensis, a parameter to consider in order to prepare the protein-tRNA sample, is the ratio between the two partners. Most of the time, a little excess of nucleic acids is required. For example, in the case of *E. coli* LeuRS-tRNA^{Leu} complex, the best crystals were obtained with a 1:1.2 ratio.

Acknowledgements

C. D. and C. T. thank Marjorie Catala for her help in optimizing in vitro T7 transcription on tRNA.

References

1. Grosjean H, De Henau S, Doi T, Yamane A, Ohtsuka E, Ikehara M, Beauchemin N, Nicoghosian K, Cedergren R (1987) The in vivo stability, maturation and aminoacylation of anticodon-substituted *Escherichia coli* initiator methionine tRNAs. *Eur J Biochem* 166: 325–332
2. Motorin Y, Helm M (2010) tRNA stabilization by modified nucleotides. *Biochemistry* 49:4934–4944
3. Phizicky EM, Hopper AK (2010) tRNA biology charges to the front. *Genes Dev* 24: 1832–1860
4. 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
5. Suddath FL, Quingley GJ, McPherson A, Sneden D, Kim JJ, Kim SH, Rich A (1974) Three-dimensional structure of yeast

Chapitre III

Production d'un substrat de TrmK : l'ARN^tSer de *B. subtilis* –
Article 2 : Dégut et al. (2016)

- phenylalanine transfer RNA at 3.0 angstroms resolution. *Nature* 248:20–24
6. Rould MA, Perona JJ, Soll D, Steitz TA (1989) Structure of *E. coli* glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. *Science* 246:1135–1142
 7. Selmer M, Dunham CM, Murphy FV 4th, Weixlbaumer A, Perty S, Kelley AC, Weir JR, Ramakrishnan V (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313:1935–1942
 8. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FV 4th, Weir JR, Ramakrishnan V (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* 326:688–694
 9. Jenner L, Demeshkina N, Yusupova G, Yusupov M (2010) Structural rearrangements of the ribosome at the tRNA proofreading step. *Nat Struct Mol Biol* 17:1072–1078
 10. Demeshkina N, Jenner L, Yusupova G, Yusupov M (2010) Interactions of the ribosome with mRNA and tRNA. *Curr Opin Struct Biol* 20:325–332
 11. Wegrzyn G, Wegrzyn A (2008) Is tRNA only a translation factor or also a regulator of other processes? *J Appl Genet* 49:115–122
 12. Li Y, Zhou H (2009) tRNAs as regulators in gene expression. *Sci China C Life Sci* 52:245–252
 13. Green NJ, Grundy FJ, Henkin TM (2010) The T box mechanism: tRNA as a regulatory molecule. *FEBS Lett* 584:318–324
 14. Zhang J, Ferre-D'Amare AR (2013) Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* 500:363–366
 15. Giege R (2008) Toward a more complete view of tRNA biology. *Nat Struct Mol Biol* 15:1007–1014
 16. Tisne C, Rigourd M, Marquet R, Ehresmann C, Dardel F (2000) NMR and biochemical characterization of recombinant human tRNA(Lys)₃ expressed in *Escherichia coli*: identification of posttranscriptional nucleotide modifications required for efficient initiation of HIV-1 reverse transcription. *RNA* 6:1403–1412
 17. Barraud P, Schmitt E, Mechulam Y, Dardel F, Tisne C (2008) A unique conformation of the anticodon stem-loop is associated with the capacity of tRNA^{fMet} to initiate protein synthesis. *Nucleic Acids Res* 36:4894–4901
 18. Palencia A, Crépin T, Vu MT, Lincecum TL Jr, Martinis SA, Cusack S (2012) Structural dynamics of the aminoacylation and proof-reading functional cycle of bacterial leucyl-tRNA synthetase. *Nat Struct Mol Biol* 19:677–684
 19. Ponchon L, Catala M, Seijo B, El Khouri M, Dardel F, Nonin-Lecomte S, Tisné C (2013) Co-expression of RNA-protein complexes in *Escherichia coli* and applications to RNA biology. *Nucleic Acids Res* 41:e150
 20. Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* 2:661–665
 21. Meinnel T, Mechulam Y, Fayat G (1988) Fast purification of a functional elongator tRNA^{fMet} expressed from a synthetic gene in vivo. *Nucleic Acids Res* 16:8095–8096
 22. Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189:113–130
 23. Sampson JR, Uhlenbeck OC (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc Natl Acad Sci U S A* 85:1033–1037

Chapitre III

Production d'un substrat de TrmK : l'ARN^tSer de *B. subtilis* –
Article 2 : Dégut et al. (2016)

V.3 Cristaux de l'ARNt^{Ser} produit *in vivo*

Tous les ARNt que j'ai produits au cours de ma thèse ont été utilisés pour des tests de cristallisation. Seul l'ARNt^{Ser} produit *in vivo* a donné des cristaux (**Figure V-1**). J'ai obtenu avec cet ARNt un très grand nombre de conditions de cristallisation différentes. En effet sur les 2 kits testés à deux température (4°C et 18°C), j'ai obtenu près de 50 conditions différentes. Cependant, je n'ai pu, pour aucune condition, obtenir de diffraction dépassant les 18Å de résolution. La très grande conditions de cristallisation a rendu l'optimisation assez difficile, aucun dénominateur commun n'étant évident. De plus, les cristaux avaient tous une morphologie parfaite (cubes bien définis, lisses, relativement solides...). Aucun indice visuel ne permettait de guider l'optimisation. A ce jour, je n'ai pas réussi à obtenir de cristaux diffractant avec une résolution suffisante pour résoudre la structure de l'ARNt^{Ser} de *B. subtilis*.

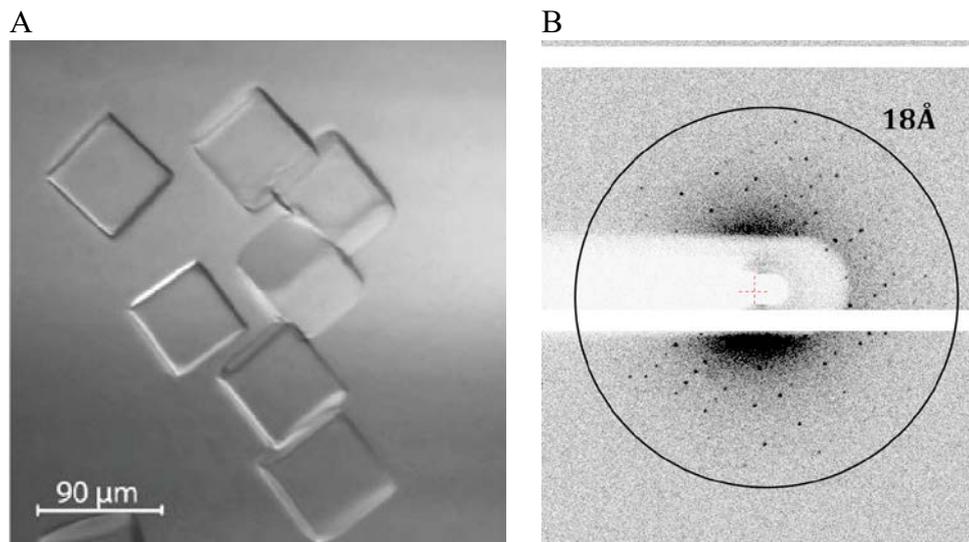


Figure V-2 A. Forme cubique des cristaux de l'ARNt^{Ser}. Les cristaux sont lisses et parfaitement géométriques. B) L'un des meilleurs clichés de diffraction des cristaux de l'ARNt^{Ser} obtenus. Le pouvoir diffractant des cristaux est très faible (< 18 Å).

VI. Attribution du squelette peptidique de TrmK

L'attribution séquentielle du squelette peptidique de TrmK était indispensable à l'analyse des empreintes RMN du complexe entre l'enzyme et son substrat. Les travaux concernant cette attribution sont présentés dans un article qui doit être soumis au journal *Biomolecular NMR Assignments* (paragraphe VI.3 en page 135).

VI.1 Etude de TrmK par RMN

TrmK est une protéine de 26kDa et son étude devait donc être réalisable grâce à un double marquage $^{13}\text{C}/^{15}\text{N}$ et des techniques de RMN classiques. Les spectres 2D ($^1\text{H},^{15}\text{N}$)-TROSY à 600 MHz ont néanmoins des pics plutôt larges (40 Hz) et une forte superposition dans la partie centrale du spectre (**Figure VI-1**). Ceci est dû en particulier au tampon qui contient 2% de glycérol, la viscosité augmentant fortement la largeur de raie. Par ailleurs, la stabilité de la protéine à haute concentration nécessite une grande quantité de sel (500 mM NaCl). Ces fortes concentrations de sels nuisent également au rapport signal/bruit des expériences RMN enregistrées, notamment sur les sondes cryogéniques.

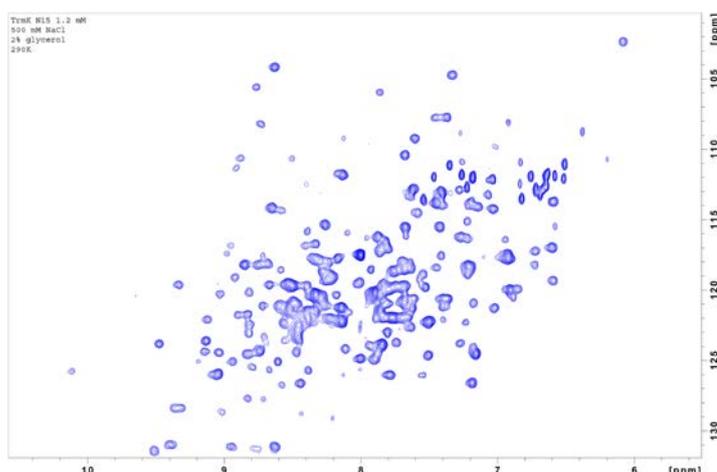


Figure VI-1 Spectre ($^1\text{H},^{15}\text{N}$)-TROSY de la protéine TrmK à 1.2 mM à 600 MHz. La largeur de raie est relativement élevée (40 Hz) et le centre du spectre comporte beaucoup de pics superposés.

Dans ces conditions, l'attribution des signaux RMN du squelette peptidique de la protéine était impossible. Nous avons décidé de produire la protéine triplement marquée $^{13}\text{C}/^{15}\text{N}/^2\text{H}$ et d'enregistrer les spectres sur les spectromètres 800 MHz et 950 MHz du TGIR-RMN à l'ICSN à Gif-sur-Yvette. Le milieu de culture utilisé est un milieu riche entièrement deutéré permettant d'obtenir un niveau de deutération d'environ 97%. Lors de la production de la protéine dans ce milieu, les acides aminés incorporés sont entièrement deutérés, y compris sur les hydrogènes amides de la chaîne principale. Cependant, ceux-ci sont échangeables avec les hydrogènes du solvant. Lors de la purification, ces deutérium échangeables sont remplacés par des protons, permettant ainsi l'étude de la protéine grâce à cette reprotonation partielle. Il arrive pour les protéines de grande taille que certains deutérium amide ne s'échange pas, car trop peu en contact avec le solvant, ceci n'a pas été le cas pour TrmK.

La durée de vie du signal RMN décroît avec l'augmentation de la taille des protéines, ce qui dégrade les expériences RMN d'attribution, comportant de nombreux transferts. Cette relaxation du signal est en grande partie due aux couplages dipolaires entre protons et hétéroatomes (carbones et azotes). Le remplacement de tous les protons de la protéine par du deutérium diminue cette relaxation, augmentant ainsi la durée de vie du signal. Les raies des signaux sont plus fines et le signal sur bruit est meilleur.

L'utilisation conjointe des très hauts champs, de la deutération de la protéine et de l'acquisition des données en échantillonnage non uniforme (NUS), m'a permis d'enregistrer des spectres de très bonne qualité.

VI.2 Principe de l'attribution séquentielle

Le spectre ($^{15}\text{N},^1\text{H}$)-HSQC (Heteronuclear Single Quantum Coherence) d'une protéine est un spectre RMN de corrélation proton-azote. Chaque pic sur ce type de spectre correspond au groupement amide de la chaîne principale d'un acide aminé à l'exception des prolines qui ne sont pas observables dans ce type de spectres. On

retrouve également des pics pour les chaînes latérales portant des groupements N-H. La position de ce pic sur le plan 2D est dictée par le type d'acide aminé auquel le groupement amide appartient et par l'environnement chimique de ce groupement. Le travail d'attribution séquentielle consiste à trouver pour chaque acide aminé de la protéine, le pic du spectre HSQC lui correspondant. Pour cela, des expériences triples résonances utilisant les noyaux ^1H , ^{15}N et ^{13}C sont utilisées. Ces expériences comportent typiquement trois dimensions et donnent des informations sur l'acide aminé précédent dans la séquence l'acide aminé considéré, ce qui permet d'attribuer de proche en proche la totalité de la séquence (**Figure VI-2**). La majorité des expériences que j'ai enregistrées sont de type TROSY (Transverse Relaxation Optimized Spectroscopy), expériences qui donnent des largeurs de raie plus faible et une meilleure sensibilité, particulièrement à haut champs (décrites dans [295]).

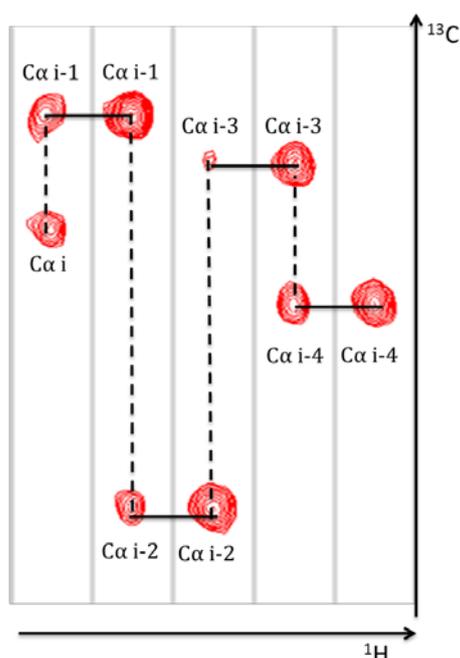
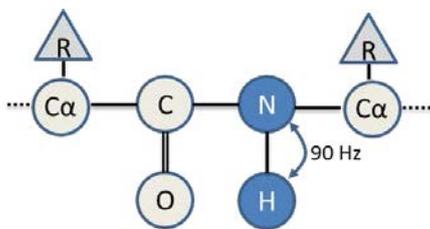


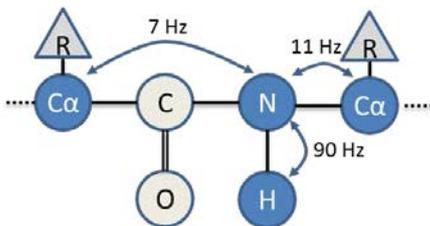
Figure VI-2 : Exemple d'attribution sur une expérience de type 3D HNCA. Chaque colonne (bande ou "strip") correspond à un pic N/H de la 2D HSQC. Dans chacune de ces bandes, on trouve un pic intense et un pic plus faible. Le pic intense correspond au $\text{C}\alpha$ du résidu étudié (i), le plus faible est celui du $\text{C}\alpha$ précédent ($i-1$). En faisant correspondre les bandes, on attribue la séquence de proche en proche.

Les expériences RMN que j'ai utilisées pour l'attribution des signaux RMN de TrmK sont très classiques –en dehors peut être de l'(H)N(COCA)NH. J'en propose ici une explication succincte qui je l'espère permettra aux non-spécialistes de comprendre le travail d'attribution résumé dans l'article 3. Les flèches sur les schémas

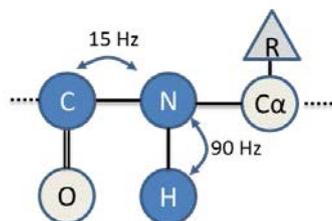
permettent de visualiser les transferts de magnétisation le long de la chaîne peptidique. Les atomes marqués en bleu sont ceux dont le noyau sera observé, c'est-à-dire, ceux qui donneront lieu à un pic sur le spectre. Les constantes de couplage scalaire entre les différents noyaux sont données sur les schémas. Plus les constantes de couplages sont faibles, plus les transferts de magnétisation d'un noyau à l'autre seront longs. L'allongement du temps de transfert de la magnétisation conduit à une diminution du signal du fait de la relaxation du signal pendant cette durée. Dans le cadre d'expériences sur une protéine deutérée, la magnétisation commence toujours sur le proton amide du résidu étudié puisque ce sont les seuls protons disponibles. Elle est transférée de proche en proche le long des liaisons covalentes, puis ramenée à ce même proton pour être observée.



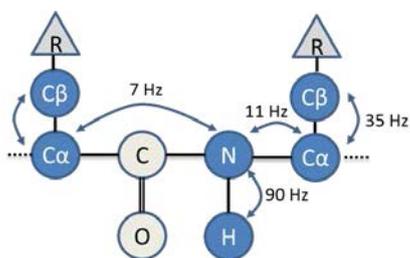
L'expérience **2D-HSQC** permet l'observation des groupements amides des protéines et iminos des acides nucléiques. C'est sur ce type de spectre que l'on pourra observer l'empreinte de partenaire.



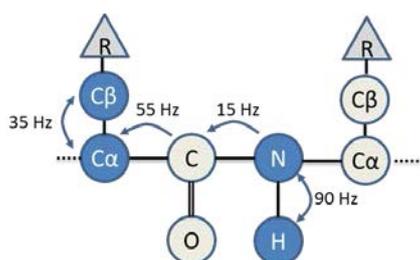
L'expérience **3D HNCA** est la plus sensible des expériences 3D (avec l'HNCOC) car elle requiert le moins de transfert. Dans cette expérience, on observera des corrélations entre le groupement NH du résidu considéré et son Cα ainsi que le Cα du résidu précédent dans la séquence de la protéine (**Figure VI-2**).



L'expérience 3D **HNCO** est très sensible, elle donne la corrélation entre le groupement NH du résidu considéré et le déplacement chimique du CO (carbone du carbonyl) du résidu précédent. Elle est importante pour les calculs structuraux car le déplacement chimique du CO donne des informations sur les angles dièdres de la liaison peptidique, mais cette expérience n'apporte aucune information séquentielle.



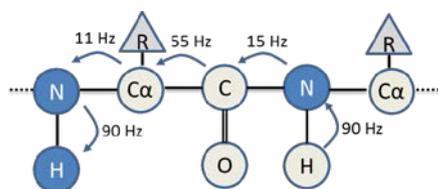
L'expérience 3D **HNCACB** donne des pics de corrélation du NH avec le C α et le C β (carbone β) du résidu considéré ainsi qu'avec les C α et C β du résidu précédent. Cette expérience est moins sensible que l'HNCA car avec un transfert supplémentaire, le temps s'allonge et le signal décroît du fait de la relaxation.



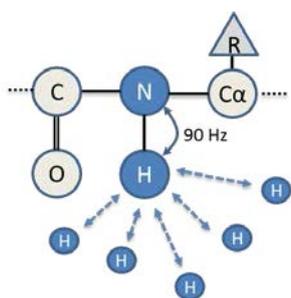
L'expérience 3D **HN(CO)CACB** permet de corréler le NH du résidu considéré aux carbones C α et C β du résidu précédent. Elle peut sembler redondante avec l'expérience précédente, mais elle est beaucoup plus sensible, du fait des constantes de couplages plus favorables en passant par le CO). Elle est de ce fait indispensable pour des protéines où les transferts sont parfois difficiles (comme c'est le cas pour TrmK).

Chapitre III

Attribution du squelette peptidique de TrmK – Principe de l'attribution séquentielle



Enfin l'expérience **(H)N(COCA)NH** est assez peu sensible sur les protéines structurées, mais reste efficace dans le cadre de la deutération à 100%. Cette expérience corrèle le NH du résidu considéré à l'azote du résidu suivant. Elle a été cruciale dans l'attribution de TrmK. Le schéma du transfert de l'aimantation est ici différent, le point de départ et d'arrivée sont deux les protons de deux groupements amides amide différents.



L'expérience 3D ^{15}N -NOESY-HSQC [296] ne repose pas uniquement sur les transferts par couplage scalaire. Elle renseigne sur les protons proches du proton amide dans l'espace (5~6 Å). On saura ainsi facilement si le résidu en question est impliqué dans une hélice ou un feuillet. Cette expérience donne notamment des possibilités de vérifications à longue distance de l'attribution.

L'utilisation conjointe des trois expériences HNCA, HNCACB, HN(CO)CACB permet de définir pour chaque pic de corrélation HN présent dans le spectre 2D-HSQC, un système de spin regroupant les résonances des atomes H_i , N_i , $\text{C}\alpha_i$, $\text{C}\beta_i$, et $\text{C}\alpha_{i-1}$, $\text{C}\beta_{i-1}$. Le travail consiste ensuite à faire correspondre les systèmes de spin entre eux afin d'obtenir une séquence de plusieurs systèmes. Certains acides aminés ont des résonances particulières qui permettent de les reconnaître de façon certaine. Par exemple, le $\text{C}\beta$ des alanines a un déplacement chimique autour de 20 ppm qui le différencie des $\text{C}\beta$ des autres acides aminés que l'on trouve entre 30 et 40 ppm. C'est le cas également des sérines et thréonines qui ont, au contraire, des déplacements

chimiques de leur C β entre 60 et 70 ppm. Les glycines sont également facilement identifiables par l'absence de C β . Une séquence comportant plusieurs systèmes de spins avec ce type de particularités peut souvent être positionné sans ambiguïté dans la séquence protéique, et constitue un point de départ pour l'attribution séquentielle.

J'ai également enregistré sur notre spectromètre (600 MHz) une expérience (H)N(COCA)NH, plutôt utilisée pour les protéines désordonnées [297–299] ; mais qui a donnée de très bon résultat sur les échantillons de TrmK deutérée. Cette expérience permet d'obtenir directement la résonance azote du résidu suivant le résidu considéré (N $_{i+1}$) et offre un chemin d'attribution et de validation complètement indépendant des expériences classiques impliquant les carbones α et β (C α et C β). Cette indépendance permet de discriminer facilement les solutions lorsque plusieurs semblent possibles. Cette expérience a été cruciale dans l'attribution des signaux de TrmK résonnant dans des régions très denses des spectres.

Mon travail d'attribution des signaux RMN de la chaîne peptidique de TrmK m'a permis d'attribuer 96% des groupements amides des résidus de TrmK ainsi que 97% des résonances des C α , 96% des C β et 92% des CO. Ce travail d'attribution fait l'objet d'une publication qui sera soumise à Biomolecular NMR Assignments qui accompagnera le dépôt des attributions à la Biological Magnetic Resonance Data Bank.

Backbone resonance assignments of the m¹A22 tRNA methyltransferase TrmK from *Bacillus subtilis*

Clément Dégut^{a,b}, Pierre Barraud^{a,b,*}, Valéry Larue^{a,b} & Carine Tisné^{a,b,*}

^aCentre national de la recherche scientifique, Laboratoire de cristallographie et RMN biologiques, FR-75006 Paris, France

^bUniversité Paris Descartes, Sorbonne Paris Cité, Laboratoire de cristallographie et RMN biologiques, FR-75006 Paris, France

Abstract

RNA modification is a post-transcriptional process by which certain nucleotides are altered after their initial incorporation into an RNA chain. Transfer RNAs (tRNAs) is the most heavily modified class of RNA molecules. These modifications expand the chemical and functional diversity of tRNAs and enhance their structural stability. To date, more than 100 modifications have been identified, the majority of which are specific from one domain of life. However, few modifications are extensively present in the three domains of life. Among those, the m¹A nucleotide, which consists in the methylation at position 1 of the adenine aromatic ring, is found in tRNAs and ribosomal RNAs. In tRNAs, the m¹A modification occurs at position 9, 14, 22, 57 and 58. The enzyme TrmK catalyzes the m¹A formation at position 22. Here we report the backbone ¹H, ¹⁵N and ¹³C chemical shift assignments of TrmK from *Bacillus subtilis* obtained by heteronuclear multidimensional NMR spectroscopy as well as its secondary structure in solution as predicted by TALOS+. These assignments of TrmK pave the way for interaction studies with its tRNA substrates.

To cite this article: C. Dégut *et al.* **Biomolecular NMR Assignments (2016), in press.**

The final publication is available at: <http://link.springer.com/article/10.1007/s12104-016-9680-0>

Key words: backbone resonance assignment, heteronuclear NMR, unlabeled, methyltransferase, TrmK, tRNA, m¹A

1. Biological context

The biosynthesis of transfer RNAs (tRNAs) is a complex process composed of several steps leading to the formation of mature tRNAs with correct structures and functionalities. Apart from the processing of their extremities, the most salient property of the maturation process consists in the post-transcriptional incorporation of a large number of chemical modifications by the so-called modifica-

tion enzymes. Among the modification enzymes, the methyltransferases (MTases) are the most frequent and diverse (Hori 2014). They catalyze the transfer of a methyl group from a methyl donor, mostly the S-adenosyl-L-methionine (SAM), towards different positions of the nucleotides. The m¹A modification, which consists in the incorporation of a methyl group at position 1 of adenines, occurs on nucleotides 9, 14, 22 and 58 of tRNAs. This modification brings a positive charge on the adenine aromatic ring. We focused our work on the methyltransferase TrmK from *B. subtilis* that catalyzes the methylation of adenine 22 of tRNA^{Ser} and tRNA^{Tyr} (Roovers *et al.* 2008). *B. subtilis* TrmK belongs to

* Corresponding authors

Email addresses: pierre.barraud@cncrs.fr (Pierre Barraud), carine.tisne@cncrs.fr (Carine Tisné).

the COG2384 (Cluster of orthologous groups). The members of this family are found in Gram-negative and Gram-positive bacteria. Their sequences are well-conserved in many bacterial pathogens (*L. Monocytogenes*, *V. Cholerae*, *S. Pneumoniae*...). Since TrmK is essential for cell viability in *S. Pneumoniae* and since no homologues are found in humans, it was proposed to be a good target for the discovery of novel antibiotics (Thanassi *et al.* 2002). We solved the X-ray structure of *B. subtilis* TrmK (Dégut *et al.*, manuscript in preparation) and used NMR chemical shift mapping to get insight on the protein-RNA recognition mode. We report here the backbone chemical shift assignments of TrmK, assignments that were necessary to interpret the NMR chemical shift mapping with its tRNA substrate.

2. Methods and experiments

2.1. Protein expression and purification

Recombinant *B. subtilis* TrmK was expressed and purified with a protocol adapted from Roovers *et al.* (Roovers *et al.* 2008). Since wild-type TrmK is prone to aggregation through cysteine oxidation, we produced a protein variant with cysteine to serine mutations (*i.e.* C35S and C152S), which abolished protein aggregation and retained full enzymatic activity (Dégut *et al.*, manuscript in preparation). Mutagenesis was performed by use of the Quickchange site-directed mutagenesis kit (Stratagen). The presence of the desired mutations in *trmK* was checked by sequencing. This variant was overexpressed in BL21(DE3) *E. coli* cells, in rich labeled media or minimum media (see below) supplemented with kanamycin at 30 $\mu\text{g}\cdot\text{mL}^{-1}$. The cells were grown at 37 °C to OD600 ~0.6, cooled down at 18 °C and induced by adding isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cells were harvested 24 h after induction by centrifugation and frozen at -80 °C until further use. The frozen cells were suspended in 1/50 of the culture volume of a 50 mM Tris-HEPES buffer pH 8.2 containing 500 mM NaCl, 5% glycerol and 1 mM of phenylmethanesulfonyl fluoride (PMSF). The suspension was sonicated and the lysate was centrifuged for 30 min at 45'000 g. The resulting supernatant was loaded to a 5 mL Nickel Sepharose column (HisTrap, GE Healthcare) previously equilibrated with a 50 mM Tris-HCl buffer pH 8.0 containing 500 mM NaCl and 5% glycerol (equi-

libration buffer). The resin was then washed with 30 mL of buffer and the protein was eluted with a gradient of the equilibration buffer supplemented with 500 mM imidazole pH 8.0. The N-terminal His₆-tag of TrmK was removed by thrombin cleavage (25 U thrombin/mg of protein) performed overnight at 4 °C. PMSF at 250 μM , and EDTA at 1 mM were then added to the protein sample. The sample was concentrated with Amicon 10'000 MWCO (Millipore) and injected on a size exclusion chromatography column (Superdex-75 26/60, GE Healthcare) equilibrated with a 50 mM sodium phosphate buffer pH 7.0 containing 500 mM NaCl and 2% glycerol.

Doubly labeled (¹⁵N/¹³C) and triply labeled (²H/¹⁵N/¹³C) TrmK samples were obtained by growing the cells in Spectra-9CN and Spectra-9DCN media, respectively (Spectra Stable Isotopes, Inc.). Singly labeled (¹⁵N) TrmK samples were obtained by growing the cells in M9 minimum media supplemented with ¹⁵NH₄Cl. To achieve high yield of protein in the fully deuterated medium, a double selection protocol of the strain was performed as previously described (Sivashanmugam *et al.* 2009). Besides, specifically unlabeled samples were produced in M9 minimum media supplemented with ¹⁵NH₄Cl and with non-labeled histidine, arginine, or lysine at a final concentration of 1 mM, 1 hour before induction (Rasia *et al.* 2012). NMR samples of TrmK at ~0.7 mM were prepared in a 50 mM sodium phosphate buffer pH 7.0, 500 mM NaCl, 2% glycerol, 1 mM EDTA, 250 μM PMSF and 10% D₂O. The sample was put in a 3 mm-diameter NMR tube.

2.2. NMR experiments

All NMR spectra were recorded at 15 °C on Bruker 600, 800, or 950 MHz spectrometers equipped with cryogenic probes. Backbone assignment was performed using the following standard 3D NMR experiments (Salzmann *et al.* 1998): TROSY-HNCA, TROSY-HNCACB, TROSY-HN(CO)CACB and [¹H-¹⁵N] NOESY-HSQC. A 3D (H)N(COCA)NH experiment was also measured (Bracken *et al.* 1997). Internal DSS standard was used for direct referencing of the ¹H chemical shifts, and indirect referencing of ¹⁵N and ¹³C shifts (Wishart *et al.* 1995). Data processing was carried out with Topspin 3.2 for standard acquisition, and with MddNMR (Orek-

hov and Jaravine 2011) for non-uniform sampling acquisition. Analysis of spectrum and backbone assignment was performed with the Sparky suite (Goddard and Kneller). Table 1 summarizes the NMR experiments used for the assignment.

Table 1

NMR experiment recorded for backbone chemical shift assignment of TrmK

Spectra	Spectrometer	² H	Duration	NUS
HNCA	800 MHz	Y	27 h	Y (65%)
HNCACB	800 MHz	Y	86 h	Y (62%)
HN(CO)CACB	600 MHz	Y	90 h	N
(H)N(COCA)NH	600 MHz	Y	89 h	N
HNCO	950 MHz	Y	24 h	Y (35%)
¹⁵ N-HSQC-NOESY	950 MHz	N	49 h	N

²H: indicates whether TrmK was deuterated (Y: yes, N: no). Duration: indicates the experimental time in hours. NUS: indicates whether non-uniform sampling was used for the acquisition of data (Y: yes, N: no). The percentage in parenthesis indicates the amount of measured data using sparse sampling compared to conventional acquisitions.

3. Resonance assignments and data deposition

Before tackling the assignment of TrmK NMR signals, we needed to cope with problems of precipitation and proteolysis of TrmK. Indeed, TrmK precipitated during concentrations steps, needed high concentrations of salt (500 mM NaCl) and was subjected to intensive proteolysis within a few days after purification. Numerous conditions were tested using dialysis buttons. They allowed us to test the stability of TrmK in various conditions of buffer and temperature while working at high concentrations of TrmK (15 mg/mL) and low volumes (50 μ L). As a result, TrmK does not precipitate at high concentration (> 15 mg/mL) in a phosphate buffer whereas we observed precipitation more rapidly in a Tris-HCl buffer. A high concentration of salt (400 to 500 mM NaCl) limited the precipitation of TrmK and its proteolysis. Anti-proteases (PMSF, Pepstatine A, Leupeptine, EDTA) were tested. PMSF and EDTA revealed to be efficient to stop proteolysis. PMSF was added in the lysis buffer and at different stages of purification and in the NMR tube. Lastly, glycerol was necessary to prevent precipitation when concentrating TrmK.

5% of glycerol was an optimal concentration during purification steps and for storage of the protein. For the NMR studies, 2% of glycerol was used as the best compromise between the stability of the protein and the linewidth of NMR signals.

The protein sample in this study consists in 238 residues after the cleavage of the N-terminal His-tag and contains 5 prolines. NMR assignments were based on 3D heteronuclear NMR experiments performed on ²H/¹⁵N/¹³C-labeled TrmK. Interestingly, the use of the (H)N(COCA)NH experiment (Bracken *et al.* 1997), which correlates each NH group to the nitrogen of the amide group of the following residue in the sequence of the protein, revealed to be very efficient for the achievement and the validation of the assignments. This experiment is not sensitive and is commonly used for unfolded protein. However, with a fully deuterated protein, this experiment turned out to be very helpful in the assignment procedure. Besides, we used several specifically unlabeled samples to confirm the assignment or to provide starting point for the assignment (Fig. 1). For instance, Fig.1 shows the unlabeled of all the lysines of TrmK. In this sample, all the amide groups of TrmK are ¹⁵N-labeled except those of lysines. Consequently, the lysines are not observable in the red spectrum of Fig. 1. The superimposition with a reference TROSY spectrum (in black, Fig. 1) allows one to easily localize the peaks originating from lysine amide groups.

Backbone amide ¹H-¹⁵N resonance assignment of TrmK was achieved for 222 of 233 non-proline residues, corresponding to 96% of completeness (Fig. 2). Amide groups from residues M1, S28, H30, A31, L39, N40, H41, K42, T113, E116 and R117 could not be assigned the majority of which are found in loop regions. In addition, 98% of C α , 98% of C β and 92% of CO were assigned. The backbone chemical shift assignments of TrmK were deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under accession number 26744.

An analysis of the chemical shifts of HN, N, CO, C α and C β atoms was conducted with the TALOS+ webserver (Shen *et al.* 2009). Fig. 3 compares the prediction of TrmK secondary structure obtained from TALOS+ with the secondary structure observed in the crystal structure of TrmK (Dégut *et al.*, manuscript in preparation). The only light dif-

ferences lie in the prediction from the NMR data of short α -helical turns (2 or three residues) around residue 68 and 165 that are not observed in the crystal structure. Therefore, both data are in very good agreement. The TrmK fold in solution and in the crystal is thus identical. Subsequently, the NMR footprint of binding of the tRNA on TrmK will be mapped on the X-ray structure of TrmK.

Acknowledgments

Financial support from the TGIR-RMN-THC FR3050 CNRS for conducting the research is gratefully acknowledged. We thank Nelly Morellet for helpful discussions and technical advice.

4. References

Bracken C, Palmer AG, Cavanagh J (1997) (H)N(COCA)NH and HN(COCA)NH experiments for ^1H - ^{15}N backbone assignments in $^{13}\text{C}/^{15}\text{N}$ -labeled proteins *J Biomol NMR* 9:94-100.

Goddard TD, Kneller DG SPARKY 3 University of California, San Francisco.

Hori H (2014) Methylated nucleosides in tRNA and tRNA methyltransferases *Front Genet* 5:144.

Orekhov VY, Jaravine VA (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition *Prog Nucl Magn Reson Spectrosc* 59:271-292.

Rasia RM, Brutscher B, Plevin MJ (2012) Selective isotopic unlabeled proteins using metabolic precursors: application to NMR assignment of intrinsically disordered proteins *Chembiochem* 13:732-739.

Roovers M, Kaminska KH, Tkaczuk KL, Gigot D, Droogmans L, Bujnicki JM (2008) The YqfN protein of *Bacillus subtilis* is the tRNA: m $^1\text{A}22$ methyltransferase (TrmK) *Nucleic Acids Res* 36:3252-3262.

Salzmann M, Pervushin K, Wider G, Senn H, Wuthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins *Proc Natl Acad Sci U S A* 95:13585-13590.

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts *J Biomol NMR* 44:213-223.

Sivashanmugam A, Murray V, Cui C, Zhang Y, Wang J, Li Q (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli* *Protein Sci* 18:936-948.

Thanassi JA, Hartman-Neumann SL, Dougherty TJ, Dougherty BA, Pucci MJ (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae* *Nucleic Acids Res* 30:3152-3162.

Wishart DS et al. (1995) ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR *J Biomol NMR* 6:135-140.

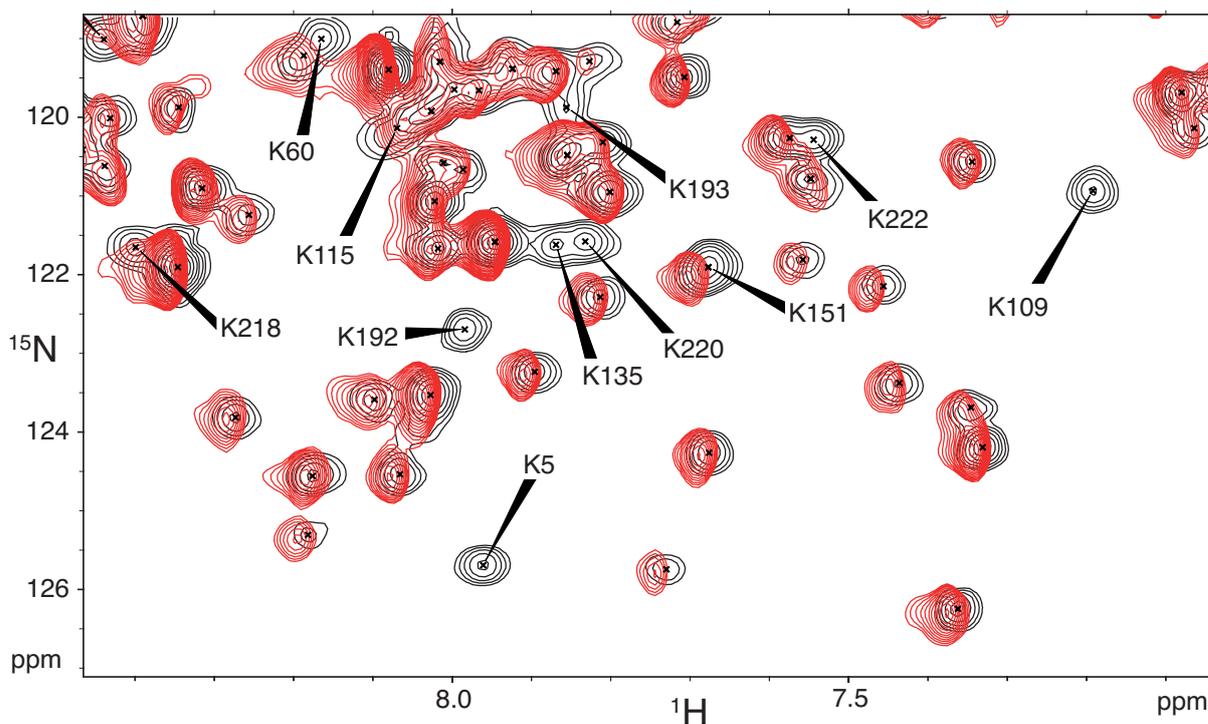


Figure 1. Superimposition of two ^1H - ^{15}N TROSY spectra of TrmK in black for TrmK uniformly ^{15}N -labeled and in red for TrmK ^{15}N -labeled except on lysines. Lysines are thus not observable on the red spectrum confirming immediately the assignment of amide group for each lysine of the protein.

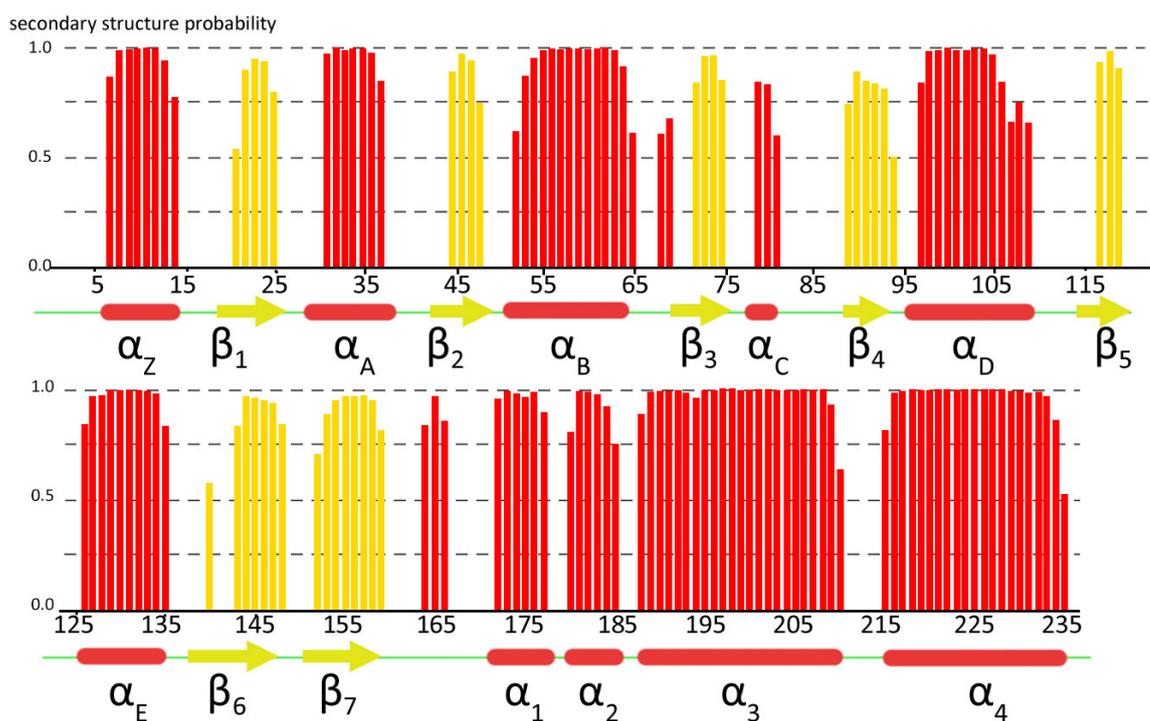


Figure 3. Prediction of TrmK secondary structure based on its backbone NMR chemical shifts using TALOS+ (Shen *et al.* 2009). The secondary structure prediction obtained with TALOS+ is shown as red bars for α -helices and yellow ones for β -strands, the height of the bars represent the probability of the secondary structure assigned by the software. Secondary structure of the TrmK X-ray structure is reported under the graph with yellow arrows for β -strands and red cylinders for α -helices.

Chapitre III

Attribution du squelette peptidique de TrmK –
Article 3 : Dégut et al. (2015)

VII. Etudes de l'interaction de la méthyltransférase TrmK avec son substrat ARNt

VII.1 Problématique et résumé de l'article 4

L'article 4 rassemble les résultats de nos recherches sur l'enzyme TrmK. Tout d'abord, des études menées par l'équipe de Louis Droogmans (Université libre de Bruxelles) montrent que TrmK est active sous forme de monomère, et des expériences avec différents ARNt ou fragments d'ARNt nous permettent de mettre en évidence des déterminants de séquence des ARNt substrats de TrmK. La structure cristallographique de TrmK de *B. subtilis* a été résolue et est décrite en détail. L'étude par RMN du complexe TrmK/ARNt^{Ser} nous a permis de déterminer la surface d'interaction de TrmK avec son substrat. Il est particulièrement remarquable que le domaine C_{ter} ne montre aucune interaction avec l'ARNt. Nous utilisons aussi les données de l'empreinte RMN entre TrmK et l'ARNt pour guider le docking d'une adénine dans le site actif de TrmK. L'ensemble des données met en évidence quatre résidus clés importants pour maintenir l'adénine dans le site actif et pour la catalyse.

Study of the bacterial m¹A₂₂ tRNA methyltransferase TrmK: Insights into tRNA recognition and catalytic mechanism.

Clément Dégut¹, Martine Roovers², Franck Brachet¹, Pierre Barraud¹, Valéry Larue¹, Joël Caillet³, Louis Droogmans⁴, Carine Tisné^{1*}

(1) Laboratoire de Cristallographie et RMN biologiques, CNRS, Université Paris Descartes, Paris Sorbonne Cité, 4 avenue de l'Observatoire, 75006 Paris, France.

(2) Institut de Recherches Microbiologiques Jean-Marie Wiame, B-1070 Bruxelles, Belgium.

(3) CNRS FRE3630, Université Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique, Paris, France.

(4) Laboratoire de Microbiologie, Université libre de Bruxelles (ULB), 6041 Gosselies, Belgium

* : To whom correspondence should be addressed. E-mail address: carine.tisne@parisdescartes.fr, Tel:+33 1 53 73 15 74, Fax:+33 1 53 73 99 25.

ABSTRACT

The enzyme TrmK of *B. subtilis* catalyzes the formation of m¹A at position 22 of tRNA^{Ser} and tRNA^{Tyr}. We solved the crystal structure of *B. subtilis* TrmK and refined it to 1.98 Å. The N-terminal catalytic domain of TrmK harbours the Rossman-fold typical of Class-I methyltransferase and thus contains the S-adenosyl-L-methionine (SAM) binding site. The C-terminal domain shows a new fold organized with a coiled-coil of two large helices preceded by two small α -helices. NMR chemical shift mapping shows that the co-factor SAM is essential for the binding of the tRNA substrate. It also shows that the C-terminal

domain does not bind the tRNA substrate. In addition, we identify the adenine binding pocket in TrmK and a major binding site of the tRNA located at the head of TrmK, above the SAM binding site. We finally used molecular docking to position a N9-methyladenine in the active site of TrmK. Interestingly, we identify four essential conserved residues, R9, Q121, N123 and G95 that are involved in the binding of the substrate adenine. A network of interactions between these residues, the SAM cofactor and the adenine positions the N1 atom of the substrate adenine directly near the SAM methyl group, ready to accept the methyl transfer.

INTRODUCTION

Transfer RNAs (tRNAs) contain a large number of modified nucleosides incorporated post-transcriptionally by a variety of enzymes (1). Among nucleoside modifications, methylations are the most frequent and diverse. They are catalyzed by methyltransferases (MTases), in great majority using S-adenosyl-L-methionine (SAM) as methyl donor.

The m¹A modification, which consists in the incorporation of a methyl group at position 1 of adenine, can be found at positions 9, 14, 22 and 58 of tRNAs. It is also formed at position 57 as an intermediate in the biosynthesis of 1-methylinosine (1). m¹A₉ occurs in mammalian mitochondrial tRNAs and in Archaea (2). In mitochondrial tRNA^{Lys}, m¹A₉ prevents the formation of an alternative structure of this tRNA by avoiding the base-pairing between A₉ and U₆₄ (3). The m¹A₁₄ is rare and thus far has been reported only in cytoplasmic tRNA^{Phe} from several mammals (4). The gene encoding this enzyme is still unknown. The m¹A₅₈ is found in tRNAs from organisms in the three domains of life. The genes and corresponding enzymes, Trm6 in eukaryota and TrmI in procaryota, have been identified and characterized in yeast and humans, in the bacteria *T. thermophilus* (5-9) and *M. tuberculosis* (10) and in the archaeon *P. abyssi* (11,12). TrmI is a methyltransferase of class I, *i.e* its catalytic domain at the C-terminal part of TrmI is composed of seven β -strands surrounded by six α -helices and contains the SAM-binding pocket. Its N-terminal domain is globally a β -structure composed of six β -strands and one small α -helix. We solved the crystal structures of the *T. thermophilus* TrmI and of two mutants of its catalytic center (6,9). We showed that TrmI is active as a tetramer, remains tetrameric upon tRNA binding and binds up to two tRNA per tetramer. In addition, it was recently shown that the

sequence of tRNA have an importance for TrmI activity and determinants of recognition between TrmI and its substrate tRNA at the level of the T-arms are now known (8). Besides, we identified three conserved residues (Y78, D170, Y194) important for the methylation reaction and we showed that they fashion an optimal shape of the catalytic pocket in order to position the adenine and the methyl donor in a reactive position. Notably, D170 acts as a latch that holds the co-factor in the catalytic pocket and locks the adenine in a reactive position. D170 is thus a central residue for the methyl transfer reaction, and two mechanisms, with or without deprotonation of the N₆ exocyclic amino group, are in accordance, with the available information.

The m¹A₂₂ modification is less documented. Earlier studies showed that a SAM-dependent m¹A₂₂ MTase activity was present in *B. subtilis* extracts (13,14). More recently, it was shown that the *yqfN* gene in *B. subtilis* encodes for the enzyme responsible of m¹A₂₂ tRNA methylation, named TrmK (15). The *trmK* mutant showed no detectable phenotype, neither at the growth or at the sporulation levels. However, the cumulation of several mutations affecting nucleoside modifications is often required to observe a phenotype (16,17). TrmK belongs to the COG2384 protein family. TrmK has orthologs not only in Gram-positive, but also in Gram-negative bacteria without any equivalent found to date in eucaryotes. Interestingly, TrmK is very well conserved in a number of pathogenic bacteria (*Vibrio cholerae*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Streptococcus pneumoniae*...) with high percentage of sequence identity (> 40%). In *Streptococcus pneumoniae*, it was shown that TrmK ortholog (Sp1610) is essential for the bacteria survival (18). Three X-ray structures of proteins from this COG have been solved and deposited in the protein data bank (PDB code 3LEC, 3GNL, 3KR9)(19), but their functions and their interactions with their substrates were not studied.

In this article, we report the X-ray structure of TrmK from *B. subtilis* which has been shown to be a m¹A₂₂ tRNA methyltransferase (15). We show that TrmK only methylates full-length tRNA, either with large or small variable loop, and is not active on mini-substrates. We also demonstrate that the presence of others modifications decrease the efficiency of TrmK methyl transfer. With NMR chemical shift mapping, we show that prior binding of the SAM cofactor is required for TrmK to bind its tRNA substrate. These NMR experiments also put in light two essential residues for the binding of the target adenine, namely R9 and Q121. We finally perform a molecular docking to position a N⁹-methyladenine in the active site of TrmK. Interestingly, we identify four conserved residues, R9, Q121, N123 and G95

that are likely involved in the binding of the substrate adenine. A network of interactions between these residues, the SAM cofactor and the adenine positions the N1 atom of the substrate adenine directly near the SAM methyl group, ready to accept the methyl transfer.

MATERIALS AND METHODS

Cloning, Expression and purification of TrmK for NMR and X-ray studies

Recombinant *B. subtilis* TrmK was expressed and purified as previously described (15). The *trmK* gene was amplified by PCR and cloned into the PCRII blunt vector and then transferred into the pET28b expression vector, allowing T7 expression of a N-terminal His-tagged recombinant protein. For structural studies, a variant of TrmK mutated on cysteines by serines (C35S and C152S) was used. Mutagenesis was performed by use of the Quikchange site-directed mutagenesis kit (Stratagen), using the oligonucleotides outlined in Table S1. The presence of the desired mutation in *trmK* was checked by sequencing. This variant was overexpressed in *E. coli* (BL21(DE3) strain), the induction was performed by adding 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) after having grown the bacteria to an optical density of 0.6 at 600 nm. Cells were harvested after incubation at 18°C during 24 h, pelleted and frozen at -80°C until further use. The frozen cells were suspended in 20 ml of a 50 mM Tris/HEPES buffer pH 8.2 containing 500 mM NaCl, 5% glycerol and 1 mM of phenylmethanesulfonylfluoride (PMSF). The suspension was sonicated and the lysate was centrifuged for 30 min at 15 000 g. The resulting supernatant was applied to a 5 ml Nickel Sepharose column (HisTrap, GE Healthcare) previously equilibrated in a 50 mM Tris.Cl buffer pH 8.0 containing 500 mM NaCl and 5% glycerol (equilibration buffer). The resin was then washed with 30 ml of buffer and the protein was eluted with a gradient of the equilibration buffer supplemented with 500 mM imidazole pH 8.0. The N-terminal His₆-tag of TrmK was removed by thrombin cleavage (25 U thrombin / mg of protein) performed overnight at 4°C. 250 μ M PMSF, 1 mM EDTA were then added to the protein sample. For X-ray studies, an additional step of hydrophobic chromatography was performed. The amount of ammonium sulfate necessary to have the protein in 1 M ammonium sulfate were added to the protein sample before applying it to a Phenyl Sepharose column (HiLoad 26/10 SP Sepharose, GE Healthcare). TrmK was eluted by a 1 to 0 M ammonium sulfate reverse gradient. The sample was then concentrated (Millipore Amicon, Molecular weight cutoff 10 kDa) and injected on a size exclusion chromatography column (Superdex-75 26/60, GE

Healthcare) equilibrated in a 50 mM Tris.Cl buffer pH 8.0 containing 500 mM NaCl and 2% glycerol.

For NMR experiments, the (C35S, C152S) TrmK variant was purified as previously described (20). Briefly, for the assignment procedure of NMR signals, samples of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled TrmK at 0.7 mM were prepared in a 50 mM sodium phosphate buffer pH 7.0, 500 mM NaCl and 2% glycerol. NMR chemical shift mappings were performed with samples of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled TrmK concentrated around 0.2 mM in a 50 mM Tris.Cl buffer pH 8.0 containing 100 mM NaCl, 10 mM MgCl_2 and 2% glycerol.

Production of *B. subtilis* tRNA^{Ser}

B. subtilis tRNA^{Ser} was produced *in vitro* by T7 transcription and also *in vivo* in *E. coli* as described previously (21).

Fragments of tRNA^{Ser} were purchased from Dharmacon (GE Healthcare).

Analysis of TrmK oligomerisation and subsequent MTase activity

Pure TrmK (5 mg) was loaded on a Superose P12 column equilibrated in a 50 mM Tris.Cl buffer pH 8.0 supplemented with 10 mM MgCl_2 and 500 mM NaCl at a flow rate of 24 mL/h. The same sample in the presence of 10 mM DTT was also similarly analysed. The calibration of the column was assessed by loading in the same buffer a mixture of ovalbumine and bovine serum albumine (BSA). The MTase activity of each peak was determined by measuring the amount of ^{14}C transferred to total *E. coli* tRNA using methyl- ^{14}C SAM as methyl donor. The reaction mixture was 200 μL : 70 μg of total *E. coli* tRNA, 4 μL of methyl- ^{14}C SAM (130 dpm/pmol), the fraction containing TrmK oligomer and H_2O . The mixtures were incubated 1 hour at 37°C. The reaction was stopped by phenol extraction, and tRNA was TCA precipitated and captured on a GC/C Glass microfiber filter for scintillation counting.

Characterization of the MTase activity of TrmK using ^{32}P -radiolabeled *B. subtilis* tRNA^{Ser} variants

Radioactive (^{32}P) *in vitro* transcripts were obtained as previously described (15) using MvaI digested pUC-tser plasmid or variants as template. [α - ^{32}P]-ATP was purchased from ICN Biomedicals (Irvine, USA) and T7 RNA polymerase from Roche Diagnostics. Radioactive

transcripts were purified by 10% polyacrylamide gel electrophoresis. The transcripts (300000 cpm) were incubated in presence of 0.1, 0.5, 1 and 5 μg of TrmK and SAM (0.2 mM). They were then digested with nuclease P1 and the resulting nucleotides were analyzed by 2D-TLC on cellulose plates as described in (2). The plates were then submitted to autoradiography.

Steady-state kinetic analysis of TrmK

[Methyl- ^3H]-SAM (555 GBq/mmol, 20.5 MBq/mL, Perkin Elmer) was mixed with non-radioactive SAM (Sigma) to achieve a specific radioactivity of about 500 cpm/pmol. The methylation kinetic assays were performed in 50 mM Tris.Cl pH 8.0 with 5 mM MgCl_2 at 37 °C. Aliquots (300 μL) were removed after different incubation times and transferred into 500 mL of 5 % (w/v) trichloroacetic acid (TCA) at room temperature for 10 min in order to quench the reaction and to precipitate the tRNA. The precipitates were collected by filtration using GF/C filters (Whatman). The filters were washed with ethanol, dried, and the radioactivity was measured by liquid scintillation counting for 2 min, resulting in a counting error below 4 %. Data were corrected by subtracting the background radioactivity determined from a control without enzyme.

For the determination of the K_M and V_{max} of TrmK for tRNA^{Ser}, the reaction mixtures contained TrmK (100 nM), tRNA^{Ser} (10 to 2000 nM) and ^3H -AdoMet (90 nM). Initial rates (v_i) for each substrate concentration (tRNA^{Ser}) were determined from the slope of linear fits of time course data points (5, 10, 15, 20, 30 min). Enzyme parameters were obtained by non-linear least square fitting using equation (1) of Michaelis-Menten kinetics. Confidence limits on the parameters were estimated by Monte-Carlo sampling using the MC-Fit program (22).

$$(1) v_i = \frac{V_{max}}{2E_0} \left(K_M + S_0 + E_0 - \sqrt{(K_M + S_0 + E_0)^2 - 4S_0E_0} \right)$$

Crystallisation and X-ray structure resolution of TrmK

Crystallisation of TrmK was performed at 4°C using the sitting-drop vapor-diffusion method. Protein samples were prepared at 3 mg/ml in 50 mM Tris.Cl buffer (pH 8.0) containing 500 mM NaCl and 2% glycerol. Drops of 1 μL were prepared by a Cybi-Disk robot mixing equal volume of protein and reservoir (100 μL). Crystals of TrmK were

obtained in 0.2 M ammonium acetate, 0.1 M sodium acetate trihydrate pH 5, 30% w/v polyethylene glycol 4000 and 8% 2-Methyl-2,4-pentanediol after microseeding from needle clusters. Crystals were harvested and flash-frozen in liquid nitrogen. Diffraction data were collected at beamline ID14-1 of the European Synchrotron Radiation Facility (Grenoble, France).

X-ray diffraction data were processed using XDS and scaled with SCALA. The structure was solved by molecular replacement with PHASER using the structures of sp1610 (*Streptococcus pneumoniae* ortholog of TrmK, PDB code 3KR9), of LMOF2365_1472 (*Listeria Monocytogens*, PDB code 3GNL) and of SAG1203 (*Streptococcus agalactiae*, PDB code 3LEC). In the resulting model, non-conserved residues were mutated into alanine and model building was first performed with ARP/WARP using the warpNtrace automated procedure. Restrained refinements of the structure were performed with the program phenix.refine in the Phenix suite (23). Model and map visualizations for manual reconstructions were performed with the program Coot (24). TrmK structure has been deposited to the Protein Data Bank under the code XXXX. (*Accession codes will be provided at the proof stage*).

Normal mode analysis of TrmK was conducted with the program ProDy (25) and visualized using VMD (26).

Isothermal titration microcalorimetry (ITC)

ITC experiments were carried out at 18°C in a Microcal ITC200 microcalorimeter with TrmK extensively dialysed against 50 mM Tris.Cl pH 8.0, 100 mM NaCl, 10 mM MgCl₂, 2% glycerol. 24 injections of 1.6 µL of S-adenosyl-L-homocysteine (SAH) at 800 µM, dissolved in the buffer of TrmK dialysis, were made into the cell containing TrmK (58 µM). Injections were spaced by 180 s. Data were analysed by using the one-site binding model in Origin 7 (Microcal) software.

NMR chemical shift mapping

NMR spectra were recorded at 18°C on Bruker spectrometers (either 950 MHz, 800 Mhz or 600 MHz) equipped with TCI-5 mm cryoprobes.

Backbone assignments were performed as previously described (20). Chemical shift mappings of the interaction between SAH and TrmK and tRNA^{Ser} and TrmK previously bound to SAH were obtained using ²H/¹³C/¹⁵N-labeled TrmK (0.14 mM) mixed with 1.4

equivalent of SAH and 2 equivalent of tRNA^{Ser} in a 3 mm NMR tube. The NMR chemical shift mapping was conducted in a 50 mM Tris.Cl buffer pH 8.0 containing 100 mM NaCl, 10 mM MgCl₂ and 2 % glycerol. For each mixture, one TROSY experiment (27) was recorded. TOPSPIN, mddNMR (28) and Sparky softwares were used to process and analyze NMR data. Chemical shift differences $\Delta(H,N)$ were derived from ¹H and ¹⁵N shift differences : $\Delta(H,N) = \sqrt{[(\Delta^{15}N W_N)^2 + (\Delta^1H W_H)^2]}$, where $\Delta = \delta_{\text{complex}} - \delta_{\text{free}}$ (difference of chemical shift between the complex and the free states) and $W_H=1$ and $W_N=1/5$.

Molecular docking

Docking of the N9-methyl adenine into the active site of TrmK was performed using HADDOCK 2.2 (29,30). HADDOCK is particularly well-suited to deal with data from NMR chemical shift mapping through the possibility to introduce ambiguous restraints (*i.e.* restraint that are applied between residues and not between atoms). Two ambiguous restraints between the adenine and neighbours R9 and Q121, put in light by the NMR chemical shift mapping, were set to 2 Å. One distance restraint set to 2 Å was applied between the N1 of adenine and the C ϵ of the methyl group of SAM co-factor. Loops made by residues 6 to 10, 93 to 99, 120 to 128, and 147 to 152 were defined as semi-flexible regions during the docking procedure.

RESULTS

TrmK is active as a monomer.

The analysis by size exclusion chromatography of TrmK (Figure 1A, red line) shows that oligomers of TrmK are present in solution. Three species can be observed, one (peak 3) corresponding to a monomer of TrmK, the second (peak 2) is a dimer of TrmK whereas the third peak (1) corresponds to aggregates of more than 300 kDa. The oligomerisation state of TrmK is dramatically different in presence of a reducing agent (DTT) (Figure 1A, blue line). Indeed, only the peak corresponding to the monomer of TrmK could be observed. This result indicates that oligomers resulting from disulfide bond between monomers of TrmK are present in non-reducing conditions. The *B. subtilis* TrmK contains two cystein residues, one at position 35 and the second one at position 152. Interestingly, the measurement of MTase activity of the different oligomers shows that the monomer of TrmK is the active species (Figure 1A, green bars). This result is reinforced by the fact that the activity of

TrmK is dramatically increased in presence of DTT (Figure 1C). High concentration of DTT (10 mM) are, necessary to observe only the monomer of TrmK in solution. No crystals of TrmK could be obtained either in the presence or absence of DTT. To get rid of the presence of this high concentration of reducing agent in crystallogenesis trials and to get a sample as homogeneous as possible, we decided, as commonly performed, to mutate the residues Cys35 and Cys152 to serine residues. The mutation of only one cysteine was not enough to prevent the formation of oligomers (data not shown), the double-mutant of TrmK (C35S, C152S) was thus cloned, produced and purified as described in Materials and Methods. The MTase activity of this mutant was the same as the activity of the WT protein. In this paper, all the structural studies were conducted with the double-mutant TrmK on cysteines 35 and 152. The measurements of MTase activity were conducted on the wild-type TrmK in presence of 10 mM DTT.

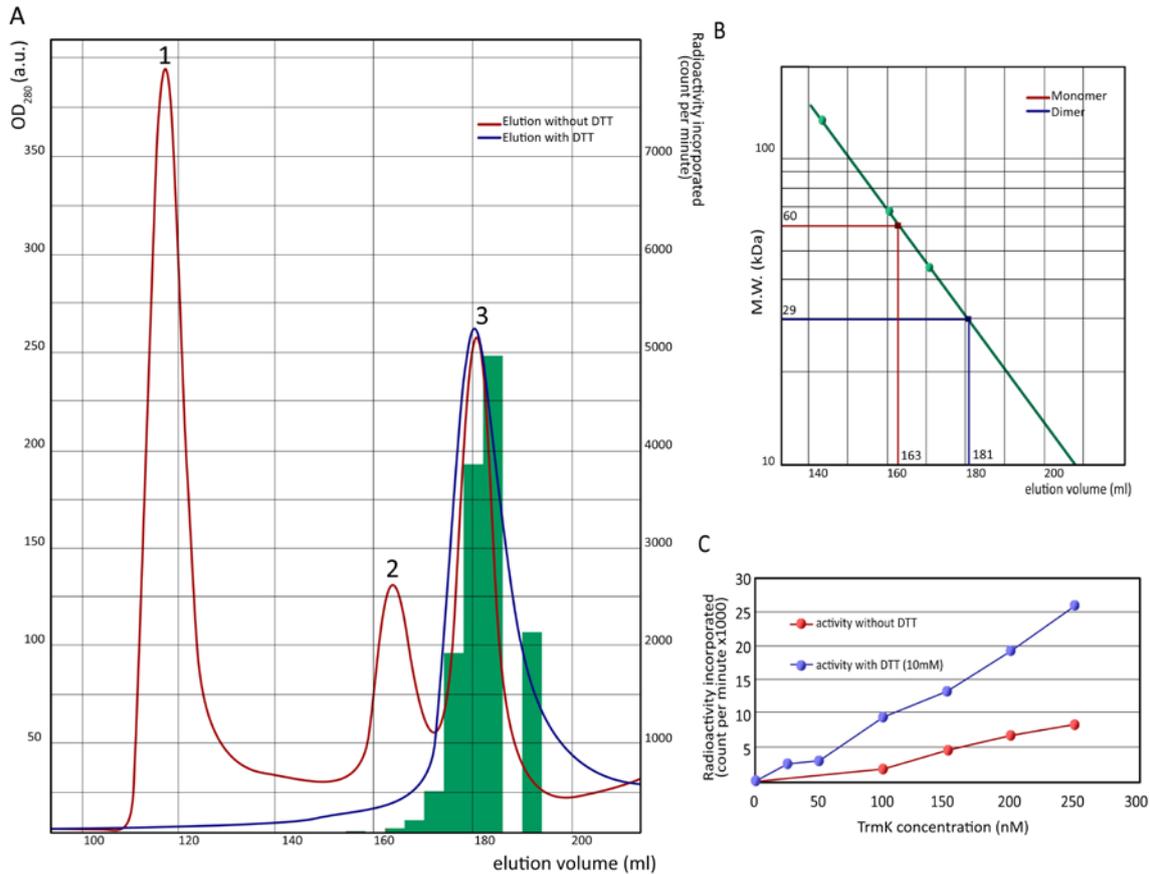


Figure 1. Analysis of TrmK oligomerisation and relative MTase activity.

(A) Size exclusion chromatography (Superose 12) of *B. subtilis* TrmK, in a buffer with (in blue) or without (in red) reducing agent (10 mM DTT). The radioactivity incorporated in the MTase assays by each fraction of the profile is indicated with green bars. (B) The molecular weights (log) are plotted against the elution volume of the Superose 12 calibrated using weight standards (ovalbumine and bovine serum albumine). Measured molecular weights of TrmK give 30 and 60 kDa for peaks eluted at 180 and 163 mL respectively. (C) MTase activity of TrmK with and without DTT measured on *E. coli* total tRNA extract (90 μ g) as a substrate and methyl- 3 H SAM as methyl donor.

Determinants of recognition of tRNA substrates by TrmK

Among organisms for which tRNA sequences are available (1), the m^1A_{22} is found in tRNA of *Bacillus subtilis* (tRNA^{Ser}, tRNA^{Tyr}), of *Geobacillus stearothermophilus* (tRNA^{Leu} and tRNA^{Tyr}), of *Mycoplasma caprilicum* (tRNA^{Gln}, tRNA^{Cys}, tRNA^{Glu}, tRNA^{His}, tRNA^{Leu}, tRNA^{Ser}, tRNA^{Tyr}) and of *Mycoplasma micoides* (tRNA^{Ser}). All these tRNA present a purine at position 13, preventing a Watson-Crick (WC) base-pairing with adenine 22 (Supplementary Figure S1). In addition, the nucleotide at position 21 is always a purine and in *B. subtilis* tRNAs, it is an adenine. In *B. subtilis*, the variable loop of the tRNAs bearing

an m¹A₂₂ is an extra-arm composed of 13 to 20 nucleotides depending on the tRNA. For the other organisms, the length of the variable loops varies from 4 to 18 nucleotides. Regarding the modification patterns of these tRNAs, in *B. subtilis*, they all bear a ribothymidine (T) at position 54 and a pseudouridine (Ψ) at position 55. Two other modifications can also be present, namely 4-thiouridine (s⁴U) at position 8 and a dihydrouridine (D) in the D-loop.

We investigated these sequence singularities by mutating some nucleotides in *B. subtilis* tRNA^{Ser} (GGA) and measuring MTase activity of TrmK on the tRNA variants (Figure 2B). First, the guanosine at position 13 was mutated into a uridine to form a WC base-pair U13-A22. This variant revealed to be a very poor substrate of TrmK (Figure 2B). Secondly, the adenosine at position 21 was mutated into a guanosine. As observed for the previous mutation, this variant is very slightly modified by TrmK. These results indicate that the specificity of TrmK is guided by the presence of a purine in position 13 and of an adenine at position 21. On the contrary, the deletion of the extra-arm of tRNA^{Ser} and its replacement by a short variable loop of sequence 5'AGGUC3' causes an increase of the MTase activity of TrmK. Therefore, the presence of a large variable loop is not a determinant for the methyl transfer by TrmK, even in *B. subtilis*.

We were also interested to decipher if a mini-substrate mimicking the D-loop of tRNA^{Ser} could be a substrate of TrmK. Two different constructs containing the D-arm and a piece of the anticodon arm were tested (Figure 2C). None of these hairpins are substrate of TrmK.

Finally, we investigated whether other modifications can influence the activity of TrmK. We overexpressed and purified the *B. subtilis* tRNA^{Ser} in *E. coli* as previously described (21) which enabled to obtain this tRNA bearing the modifications engineered by *E. coli* enzymes like T₅₄ and Ψ₅₅ (Figure 2D). We then compared the catalytic efficiency of TrmK towards the *B. subtilis* tRNA^{Ser} with or without modifications (Figure 2E). The tRNA^{Ser} produced *in vivo* in *E. coli* is less efficiently modified than the naked tRNA. Indeed, the catalytic efficiency V_{\max}/K_m of TrmK only reaches 44% for the *in vivo* tRNA^{Ser} compared to that obtained for the naked tRNA^{Ser} prepared *in vitro* (100%). This result implies that the others modifications present in the *in vivo* tRNA^{Ser} decelerate the methylation by TrmK. In particular four modifications in the D or in the T-loop can be responsible for this slow down: G_{m18}, D₁₉, T₅₄ and Ψ₅₅. These modifications, especially Ψ₅₅, that confers local structural rigidity could slow down the rate of formation of m¹A₂₂, like previously demonstrated for several modifications in tRNA of *T. thermophilus* (31).

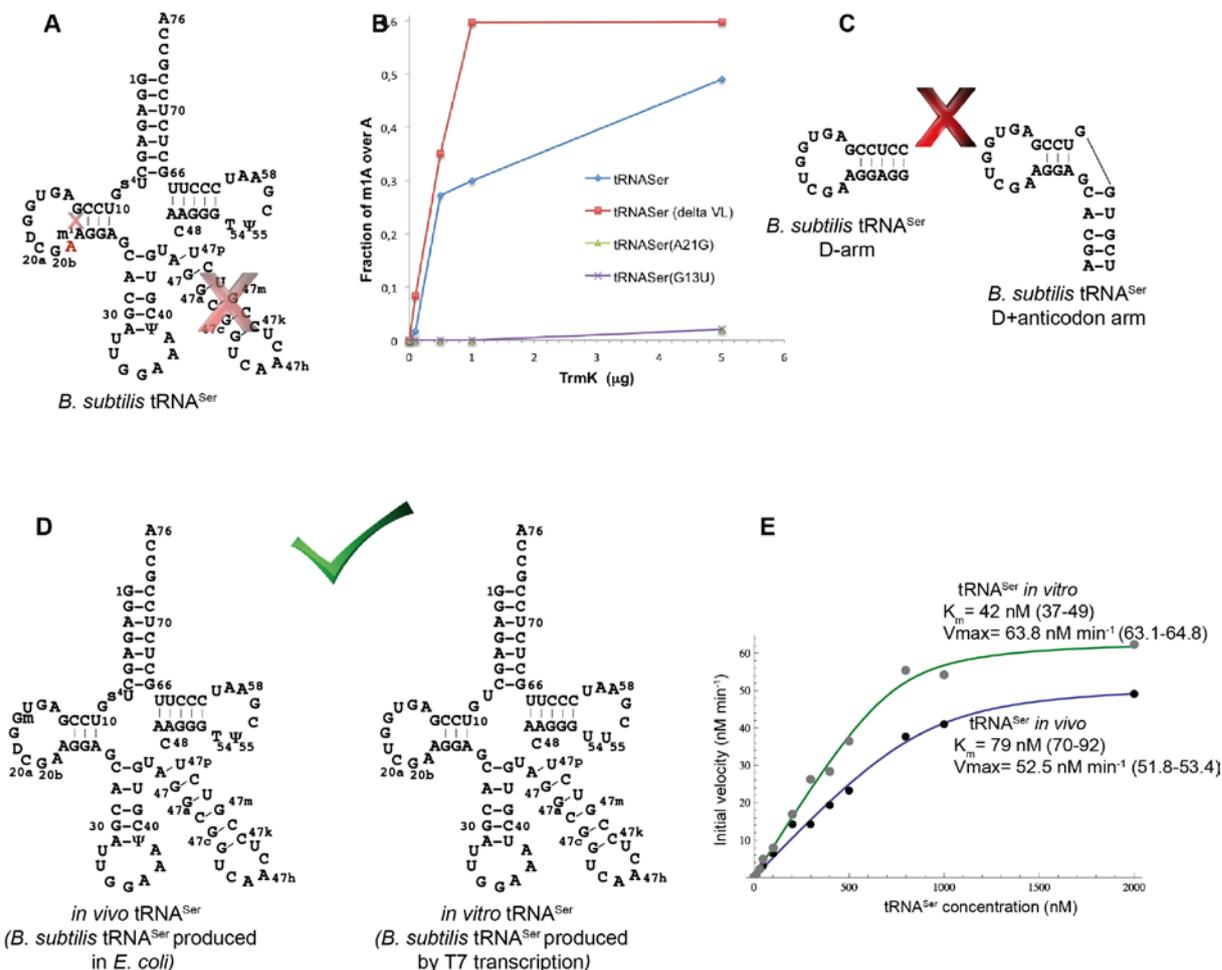


Figure 2: Determinants of tRNA^{Ser} sequence for TrmK Mtase activity.

A) Sequence of *B. subtilis* tRNA^{Ser} (GGA), the determinants are summarized on the sequence, the nucleotide at position 21 should be an adenine, the long extra-arm is not required and the A₂₂ should be unpaired., B) Fraction on methylation at position 22 incorporated by TrmK for variants of tRNA^{Ser}., C) Mimics of D-arm of *B. subtilis* tRNA^{Ser} are not TrmK substrates. D) Sequences of *B. subtilis* tRNA^{Ser} with modifications incorporated by *E. coli* and naked when produced by T7 transcription. E) Kinetic and enzymatic parameters of TrmK for tRNA^{Ser} according whether it bears or not post-transcriptional modifications.

Optimization of purification of TrmK for structural studies

Before tackling the determination of the crystal structure and NMR chemical shift mapping, we needed to cope with problems of precipitation and proteolysis of TrmK. Indeed, TrmK precipitated during concentrations steps, needed high concentrations of salt (500 mM NaCl) and was subjected to intensive proteolysis within a few days after purification. Numerous conditions were tested using dialysis buttons (see Materials and Methods and

Supplementary Figure S2). They allowed us to test the stability of TrmK in various conditions of buffer and temperature while working at high concentrations of TrmK (15 mg/ml) and low volumes (50 μ L). As a result, TrmK does not precipitate at high concentration (> 15 mg/ml) in a phosphate buffer whereas we observed precipitation in a Tris.Cl buffer. A high concentration of salt (400 to 500 mM NaCl) limited the precipitation of TrmK and its proteolysis (Supplementary Figure S2). Anti-proteases (PMSF, Pepstatine A, Leupeptine, EDTA) were tested. PMSF and EDTA revealed to be efficient to stop proteolysis (Supplementary Figure S3). PMSF was added in the lysis buffer and at different stages of purification and in the NMR tube. For crystallization purpose, we noted that it was important not to add PMSF in TrmK buffer before crystallisation assays, otherwise TrmK does not crystallize anymore. EDTA was added after thrombin cleavage of the Histidine-tag. Lastly, glycerol was necessary to prevent precipitation when concentrating TrmK. 5% of glycerol was an optimal concentration during purification steps and for stocking the protein. For the NMR studies, 2% of glycerol was used as the best compromise between the stability of the protein and the linewidth of NMR signals.

Structure determination and overall structure of TrmK

We solved the crystal structure of TrmK and refined it to 1.98 Å with a final R_{free} factor of 24 % (Table 1). The structure was solved by molecular replacement using the structure of three orthologs of TrmK from the COG2384 that present sequence identity around 40% with TrmK (Supplementary Figure S4). Four molecules are present in the asymmetric unit for a solvent percentage of 38.6 % (Supplementary Figure S5A). TrmK has an L-shape structure (Figure 3) that allows two molecules to face together in the crystal linked by a two-fold symmetry axis. The analysis of the contacts made between these two molecules by PISA server (32) stated that too few contacts (Complexation Significance Score equal to 0) are made to form the dimer in solution, which is consistent with our data described in the first paragraph. This dimer is related to another one in the asymmetric unit by a rotation of 180° around another symmetry axis.

Table 1. Data collection (Beamline ID 14eh1 at ESRF synchrotron), processing and refinement statistics.

Numbers in parentheses correspond to the highest resolution shell. CC factors are defined in (33).

A. Data collection statistics		B. Refinement statistics	
Wavelength (Å)	0.933		
Resolution range (Å)	46.24-1.98 (2.05-1.98)	Number of reflections	59409 (5790)
Space group	P 1 2 ₁ 1	R _{work} (%)	21.44 (38.86)
Unit-cell parameters (Å)		R _{free} (%)	23.95 (42.93)
a, b, c	81.56, 51.98, 108.13	CC(work)	94.7% (41.6%)
α, β, γ	90.0, 110.6, 90.0	CC(free)	95.4% (29.3%)
No. of measured reflections	223089 (20775)	Number of atoms	7539
No. of unique reflections	59508 (5793)	Protein	7232
Multiplicity	3.7 (3.6)	Ligands	4
Data completeness (%)	100.0 (98.0)	Waters	303
Mean I/ σ (I)	11.73 (0.57)	Average isotropic B-factors (Å ²)	41.06
Wilson B-factor	34.75	RMSD (bounds) (Å)	0.009
R _{merge}	21.44 (220)	RMSD (angles) (°)	1.12
CC1/2 (%)	98% (9.6%)	Ramachandran	
		Favored (%)	99
		Allowed (%)	1.2

The four molecules in the asymmetric unit (ASU) exhibit very low conformational changes (rmsd 0.25 Å for the C α atoms). The superimposition of TrmK structure with those of its orthologs shows rmsd around 1.5 Å of difference for C α atoms, which is not surprising regarding their sequence identity with TrmK. TrmK structure is composed of two domains (Figure 3A). At the N-terminus of TrmK, the catalytic domain that harbors the binding site of the SAM co-factor adopts a typical Rossman fold of class I methyltransferase consisting in a central twisted β -sheet surrounded by six α -helices. In the β -sheet, the five first β -strands are parallel whereas the two last ones are antiparallel (Figure 3A). The C-terminus domain of TrmK is structured by four α -helices, the two largest ones forming a coiled-coil. The two domains are not free from each other and indeed, numerous contacts maintained them consolidated. The normal mode analysis of TrmK structure confirms this feature and indeed only the end of the two large helices involved in the coiled-coil exhibits flexibility

(Figure 3D). The catalytic core exhibits a high level of sequence conservation among the protein of the COG2384 whereas the C-terminal domain shows poor sequence homology (Supplementary Figure S4). The DALI search (34) for structural similarity on the PDB with the catalytic domain of TrmK finds, apart from the proteins of the COG2384, a ribosomal protein L11 methyltransferase (PDB code 3GRZ, Z-score 19, rmsd 2.1 Å on 147 C α), Trm14: a m²G₆ tRNA methyltransferase (PDB code 3TM4, Z score 17.4, rmsd 2.4 Å on 152 C α), TrmI (PDB code 2PWY, Z-score 15.8, rmsd 2.4 Å on 151 C α) and Trm5, a m¹G₃₇ methyltransferase (PDB code 2ZZM, Z-score 14.5, rmsd 2.9 Å on 147 C α). No obvious information for catalysis or substrate recognition could be drawn from these similarities. This result just confirms that TrmK is a class-I methyltransferase. The DALI search for the C-terminus domain of TrmK finds motifs in protein with low Z-scores (less than 6) which suggests that this domain presents a novel fold. However, two proteins in the list retained our attention: the seryl-tRNA synthetase (PDB code 1SRV, Z-score 6.0, rmsd 1.1 Å on 50 C α) and the ribosomal protein L19 (PDB code 1S72, Z-score 6.1, rmsd 2.0 Å on 53 C α) because, in that cases, the coiled-coil is used to bind RNA partners.

Although S-adenosyl-L-methionine (SAM) or S-adenosyl-L-homocysteine (SAH) was present in crystallisation trials, crystals only contained free TrmK. Probably, the formation of the dimer in the ASU prevents the co-crystallisation with the SAM co-factor. In the case of the *S. pneumoniae* ortholog of TrmK, the protein was crystallized with and without SAM (PDB codes 3KR9 and 3KU1) and both structures are identical (rmsd between apo and ligand-bound is 0.54 Å). Therefore, binding of SAM is not likely to induce large conformational changes in TrmK. We thus positioned the SAM into the catalytic pocket of TrmK in a straightforward manner assuming that the structure with and without SAM is the same as it was observed for its ortholog from *S. pneumoniae* (19). The SAM co-factor is bound at the center of the catalytic domain in a pocket with residues harbouring negative electrostatic potentials and extremely conserved among the COG2384 (Figure 3B and 3C). Interestingly, the methyl group is pointing toward a small region that is positively charged and very well conserved, probably important for the tRNA binding.

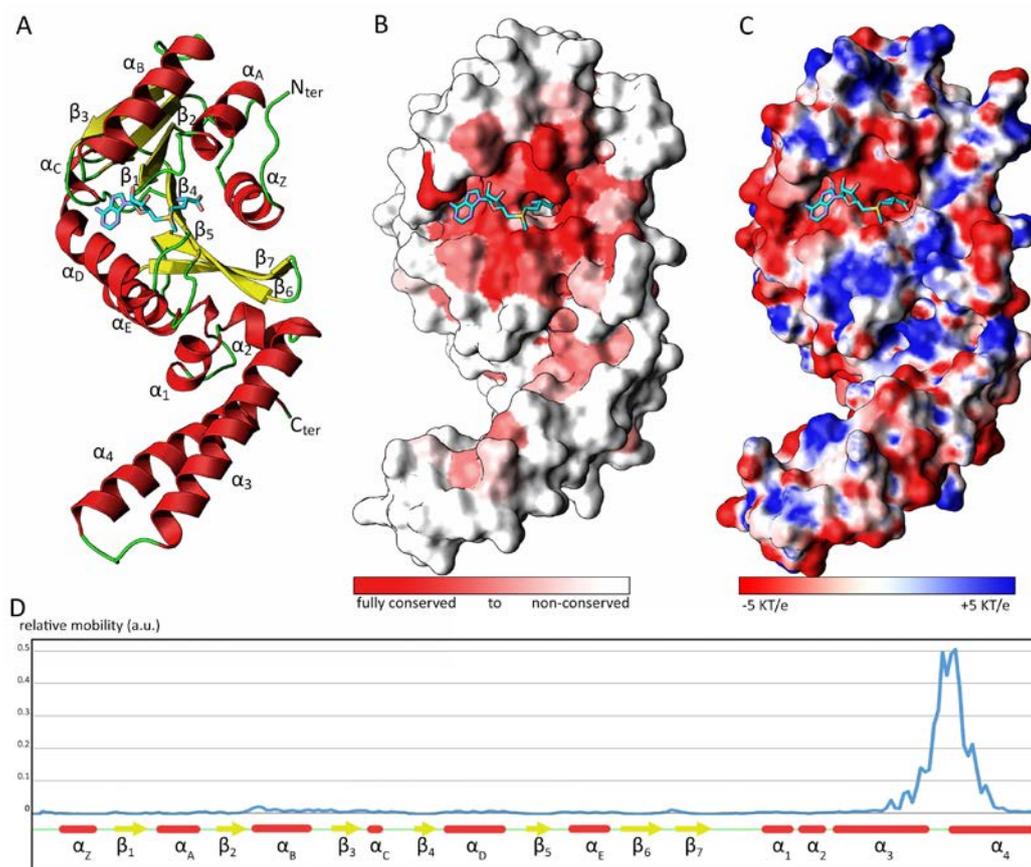


Figure 3: Crystal structure of *B. subtilis* TrmK.

SAM co-factor is represented in sticks. It was placed in the catalytic center of TrmK by superposing the TrmK structure with that of its ortholog from *S. pneumoniae* (PDB code 3KU1).

A) Ribbon representation of TrmK, α -helices are coloured in red whereas β -strands are in yellow. B) Location on the structure of TrmK of conserved residues, the sequence of *B. subtilis* TrmK was aligned with those of the representative members of the COG2384 family, C) Representation of electrostatic surface potentials of TrmK, Positive charges are in blue whereas negative ones are in red with the maximum color saturation corresponding to $-5kT/e$ (red) and $+5kT/e$ (blue). The figure was prepared with the APBS PyMOL plug-in (35) and pdb2pqr webserver (36). D) Normal mode analysis of TrmK showing the mobility of the last part of the C-terminal domain. The mean mobility of residues in the six first non-trivial computed normal modes is plotted as a function of the TrmK residue number.

The SAM co-factor binding site

The SAM binding cleft is formed by conserved residues that made a lot of contacts with TrmK. Not less than 24 residues are involved in SAM binding. Indeed, the adenine moiety is involved in numerous hydrophobic contacts with the side chains of residues I26, G49, E50, I51, G77, G79, M96, L100, I104 of TrmK. In addition, the N6 amino group of adenine makes a hydrogen bond with D78 side chain. The ribose ring makes hydrophobic contacts with residues G27, E50, P55, and M96. The methyl group is in contact with G95 and N123. The SAM carboxylate group makes a hydrogen bond with side chain of R9 and is in

hydrophobic contact with L10 and A94. Finally, the SAM ammonium group forms one hydrogen bond with the carbonyl of G27 main chain and hydrophobic contacts with D25, D29, S28 and A94. The electronic density is extremely well-defined over all the structure of the protein and particularly in the SAM binding site, even in its absence. Only two residues, K151 and Y153, near the active site seems to be mobile as the electronic density is not very well defined and thus the conformation of these residue are not clear (Supplementary Figure S5B).

Point mutation of residues that seem to be crucial for the binding of SAM were generated, namely R9A, D78A and M96A. R9 is strickly conserved, D78 can also be an asparagine, and M96 is also very well conserved (Supplementary Figure S4). These mutations provoke dramatic loss of TrmK MTase activity. Indeed, no residual activity could be measured for TrmK R9A, and only a small one for TrmK M96A. The TrmK D78A mutant only reaches to incorporate 1/10 of the quantity of methyl incorporated by the wild-type protein in the same conditions (see Supplementary Figure S8). These results validate, if it was necessary, the location of SAM co-factor in TrmK structure and the particular involvement of these residues.

The interaction between TrmK and SAM was studied by ITC. As a result, the interaction is enthalpy-driven ($\Delta H = 17$ kcal/mol) with a single binding site as expected and a dissociation constant of 17 μM (Figure 4A).

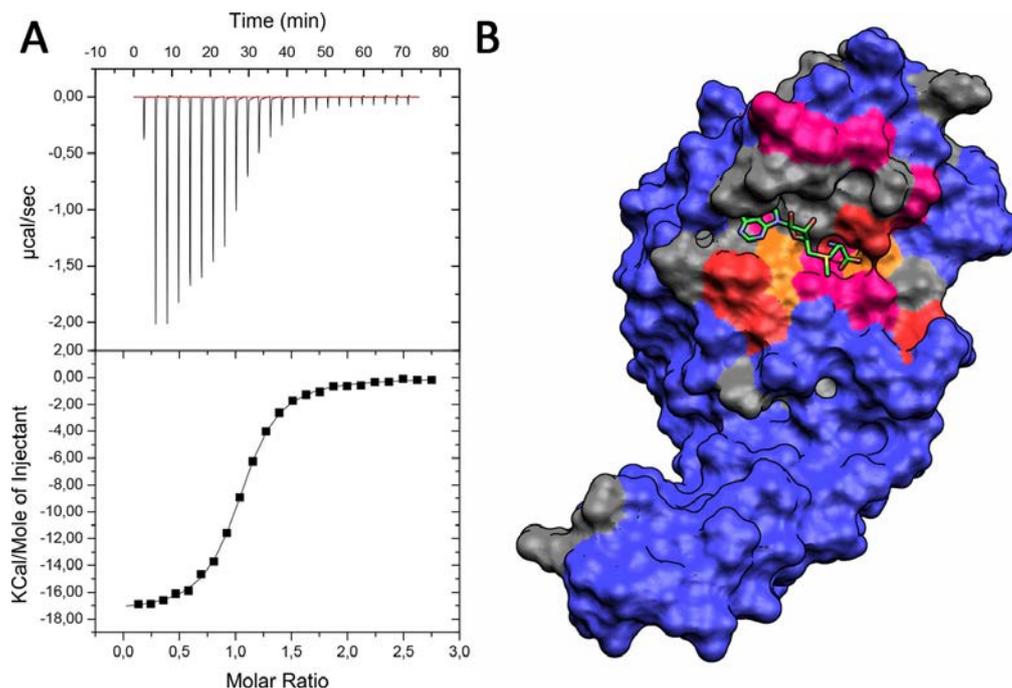


Figure 4: Interaction of TrmK with SAH.

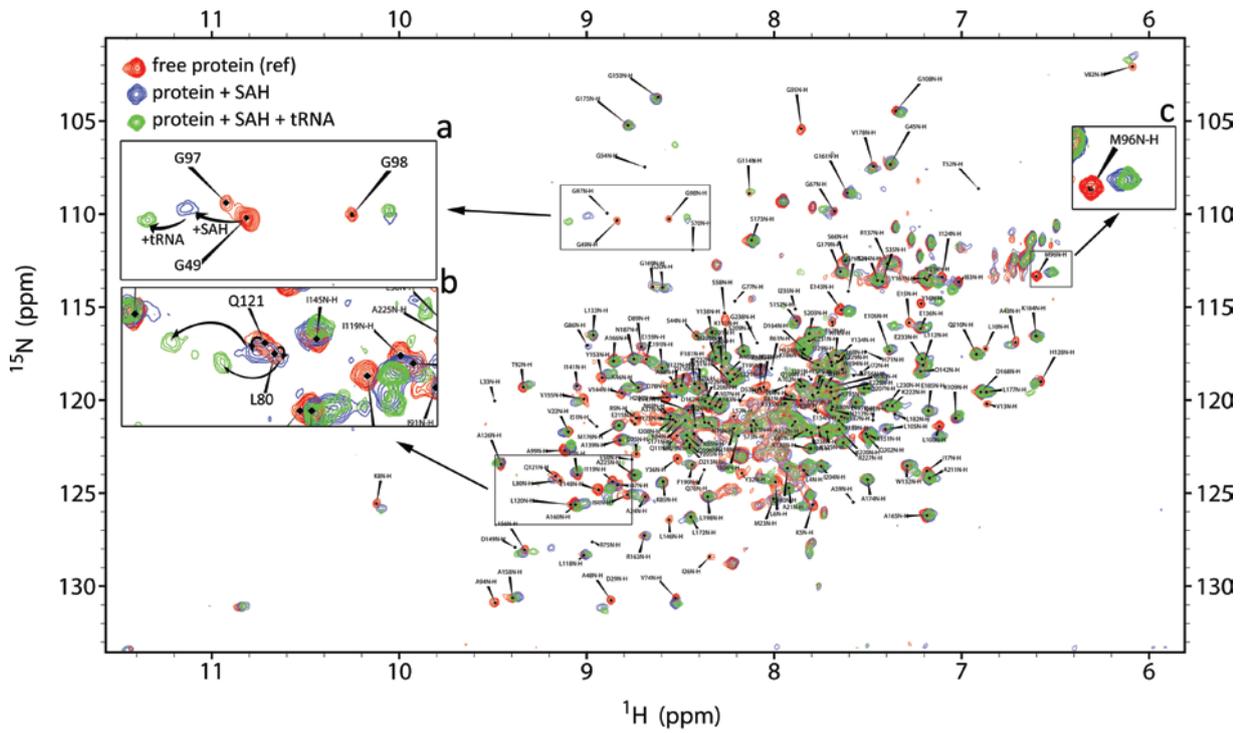
A) Isothermal titration of TrmK by SAH, The K_d value is $16.8 \mu\text{M}$. B) Chemical shift mapping of the SAH reported onto the molecular surface of TrmK, SAM is depicted as green sticks. Pink residues are those that disappeared upon addition of SAH, residues that experience chemical shift variations superior to 0.6 ppm are in orange and in red if the variation is superior to 0.9 ppm .

As described in the next paragraph, we used NMR chemical shift mapping to bring information on the binding site of substrate tRNA on TrmK. Using ^2H - ^{15}N - ^{13}C -labeled TrmK sample, we were able to assign 96% of peaks in the TROSY NMR experiment to amide groups of TrmK residues (20). At this point, it is interesting to notice that we could not observe any TrmK chemical shift variations upon addition of tRNA if the SAH cofactor was not present. Therefore, we needed to first analyse chemical shift variations upon SAH binding to TrmK and before addition of tRNA. So, we performed NMR chemical shift mapping by first recording (^1H - ^{15}N) TROSY experiment on ^2H - ^{15}N - ^{13}C -labeled TrmK bound to unlabeled SAH (Figure 5A). As expected from the previous analysis of contacts between TrmK and SAH, SAH binding causes numerous chemical shift variations and some peak disappearances (Supplementary Figure S6). Figure 4B summarizes the NMR chemical shift mapping and shows that the NMR mapping in solution is consistent with the positioning of SAH in the TrmK X-ray structure. An important network of interactions with the SAH cofactor is organized to position it in the active site (Figure S7) in a wide pocket at the center of the catalytic domain.

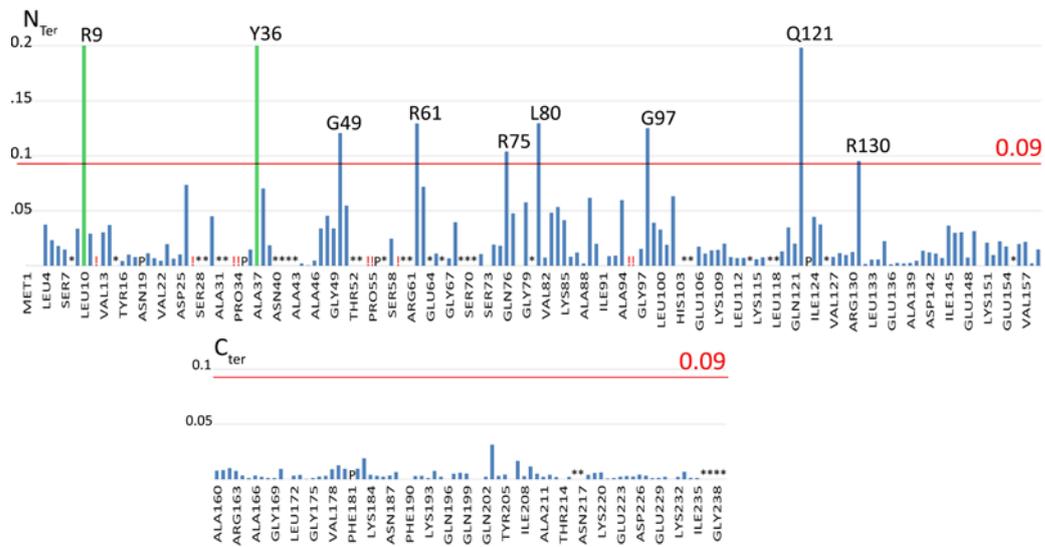
Deciphering TrmK interaction with its tRNA substrate by NMR

It was very difficult to obtain data on the complex between TrmK and tRNA^{Ser} either by biochemical or biophysical techniques for diverse reasons. First, TrmK stability requires high salt concentration. Secondly, we anticipated that the dissociation constant is in the micromolar range since the complex is dissociated by size exclusion chromatography, is in intermediate-fast exchange regime at the NMR chemical shift time-scale and no band shift could be observed either in presence or absence of SAM co-factor. Finally, we did not succeed to obtain diffracting crystals of TrmK/SAM/tRNA^{Ser} complexes. We thus turned to NMR spectroscopy to decipher TrmK interaction with tRNA in solution. We performed NMR chemical shift mapping by recording (¹H-¹⁵N) TROSY experiment on ²H-¹⁵N-¹³C-labeled TrmK bound to unlabeled SAM and unlabeled tRNA^{Ser} (Figure 5A). The residues that experience large variations of NMR chemical shifts belong to the C-terminal domain of the protein. No significant NMR chemical shift variations nor disappearance of peaks could be observed with the C-terminal domain and particularly its coiled-coil (Figure 5B). It is quite surprising as it is commonly accepted that domains of modifying enzymes outside of the catalytic center are RNA-binding domain. In the binding of TrmK to tRNA^{Ser}, the C-terminal domain is undoubtedly not involved in the binding.

A)



B)



Chapitre III

Etudes de l'interaction de la méthyltransférase TrmK avec son substrat ARNt –
Article 4 Dégut et al (2015)

C)

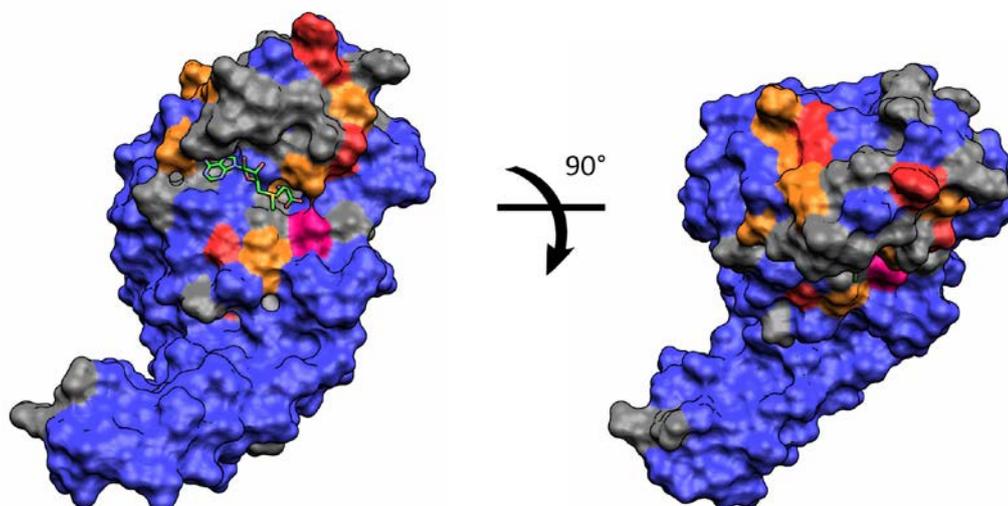


Figure 5: NMR chemical shift mapping of the binding of tRNA^{Ser} on TrmK.

A) Superposition of three 2D ¹H-¹⁵N TROSY experiments, in red TrmK alone, in blue TrmK and the SAH at a 1:1 ration and in green TrmK/SAH/tRNA^{Ser} at a ratio B) Chemical shift variations of TrmK amide groups upon SAH and tRNA^{Ser} binding. Chemical shift variations are in ppm and the number of TrmK residues are in abscissa. Residues that disappeared upon addition of tRNA are represented in green. Their chemical shift variations are arbitrary put to 0.2 ppm to easily identify them in the graph. P indicates Proline residues and the indicates residues for which their amide NMR signals have already disappeared upon SAH addition. C) NMR chemical shift mapping of tRNA^{Ser} binding reported onto the molecular surface of TrmK. Pink residues are those that disappeared upon addition of SAH, residues that experience chemical shift variations superior to 0.4 ppm are in orange and in red if the variation is superior to 0.9 ppm.

The RNA binding surface is mainly located on the same face of TrmK around the co-factor. Residues important for the binding and most probably in direct contact with RNA are those experiencing the largest variations, namely R9, Y36, Q121, L80, G97, R61, G49, R75 and R130 (residues in pink and red on Figure 5C). Interestingly, R9 and Q121 are conserved over the COG2384 and located in the near proximity of the methyl group of SAM.

Near R9 and Q121, we found N123 and G97 that show significant chemical shift variation upon tRNA^{Ser} binding. These residues are conserved and contribute with R9 to make a surface of interaction with tRNA below the methyl of SAM that is globally positively charged (Figure 5 and 3C). A second zone of interaction with the tRNA is evidenced on the TrmK structure above the SAM cavity. This zone corresponds to residues D29, R61, Q62, R75, Q76 and V82 that are contacted by one part of the tRNA in order to position the A22 in the active site. This zone also corresponds to a positively charged surface. It is interesting to notice that a number of residues that we could not assign or follow during the titration because they are in crowded region of the spectrum (residues in grey, Figure 5) are also

probably involved in the tRNA binding given their positions on TrmK surface. These residues are H30, A31, E50, D53, G54, I51, T52 and K65.

Modeling of an adenine in the active site of TrmK

To investigate the positioning of the target A₂₂ into TrmK catalytic pocket, we performed a molecular docking guided by the NMR data to locate a N9-methyladenine in the catalytic pocket. Only the base was docked because we have no data on the conformation of the ribose. We used the program HADDOCK which offers the possibility to take into account NMR chemical shift mapping, mutagenesis or residue conservation data through ambiguous interaction restraints that drive the docking procedure. Among residues that were put in light by the NMR chemical shift mapping (Figure 5B), R9, Q121 and N123 are in the near proximity of the SAM methyl group in the X-ray structure. They are very well conserved. R9 and Q121 exhibit the highest perturbations observed upon tRNA binding whereas the chemical shift perturbation of N123 is more moderate. Therefore, two ambiguous intermolecular distance restraints were applied one between R9 and the N9-methyladenine and the second one between Q121 and the N9-methyladenine. In addition, a distance restraint between the N1 atom of N9-methyladenine and the methyl of SAM was applied. In addition, we leave flexible the residues around K171 that was not well-defined in the density map, so that these residues can move.

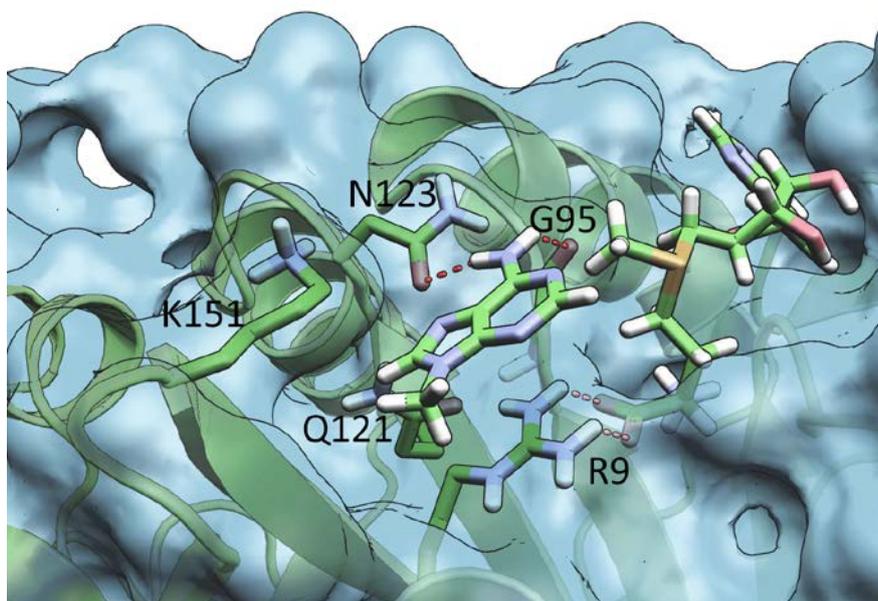


Figure 6: Docking of an adenine in the catalytic pocket of TrmK.

The adenine, SAM and some residues of interest are depicted as sticks. The exocyclic amino group of the adenine is involved in hydrogen bonds with the side chain of N123 and carbonyl group of G95. Residue R9 is hydrogen bonded with the SAM carboxylate group and in cation- π interaction with the adenine.

The model resulting from the docking is presented in Figure 6. R9 is involved in cation- π interaction with the aromatic group of adenine. Q121 δ -NH₂ is hydrogen bonded with the adenine N7. The exocyclic amino group of the adenine is involved in two hydrogen bonds, the first one with the side-chain carbonyl group of N123 and the second one with the main-chain carbonyl group of G95. This network of interaction positions the adenine N1 proximally to the SAM methyl group.

It is interesting to note that the surface of the protein does not define a cavity where the adenine comes to fit in. All these interactions are thus important to position the N1 atom of adenine 22 close to the SAM methyl group.

DISCUSSION

The crystallization of RNA/enzyme complexes remains a highly challenging task. Such structures brought key information about enzyme/RNA recognition and highlight the details of catalytic steps of the reactions. In spite of solving the crystal structure of TrmK/tRNA complex, we showed in this work that the complementary use of X-ray crystallography and NMR is invaluable to bring information on RNA recognition mode.

Figure 7 shows the superimposition of the X-ray structures of *T. thermophilus* TrmI with *B. subtilis* TrmK. The former catalyzes the methylation of m¹A₅₈ whereas the later is responsible for the methylation at position 22 of tRNA. Even, if they catalyze the same modification in tRNA, their structure, organization and recognition of the target adenine are substantially different. First, the catalytic domain is encompassed in the N-terminal domain for TrmK and in the C-terminal domain for TrmI. We showed, here, that TrmK is active as a monomer whereas TrmI is organized as a tetramer. TrmK has thus one active site whereas TrmI can bind 2 tRNA simultaneously. In addition, the extra-domain outside of the methyltransferase domain are not related either by their sequence or their structure. The superposition of the Rossman fold implies that these domains are oriented in opposite direction (Figure 6).

These two m¹A MTases recognized also differently their tRNA substrates. TrmI can methylate small hairpin mimicking the T-arm of TrmI substrate whereas we showed that TrmK only modifies full-length tRNA, even if a small variable loop is however tolerated. We also showed that the presence of post-transcriptional modifications, provokes a deceleration of the incorporation of m¹A₂₂ whereas it was shown that the other modifications accelerate the methylation by TrmI (8).

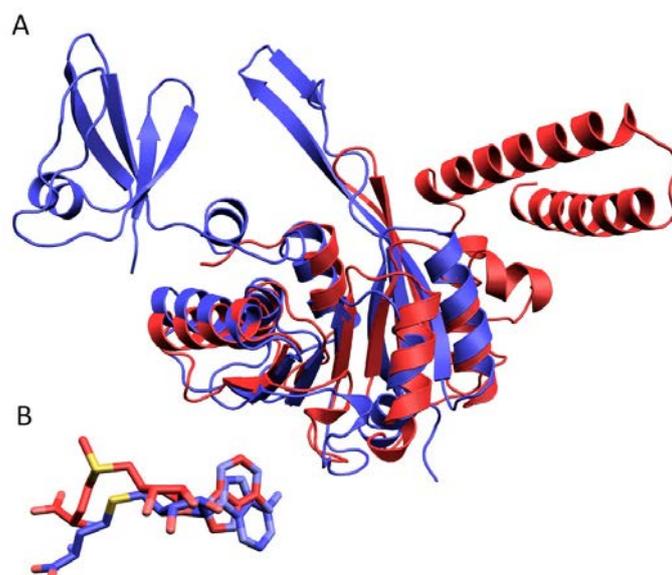


Figure 7: Structural superimposition of *B. subtilis* TrmK with *T. thermophilus* TrmI. A) Superimposition of TrmK (in blue) with *T. thermophilus* TrmI (in red, PDB code 2PWY). Both protein are superposed by their Rossman-fold catalytic domain (RMSD 2.6 Å). B) Comparison of the SAM conformation into the catalytic pocket of TrmK and TrmI. The SAM is in extent conformation in both active site,

In TrmI, the substrate adenine snugly fits into the active site pocket whereas, in TrmK, we could not obviously identify an adenine binding pocket. It is also important to notice that the SAM co-factor is required to observe TrmK-tRNA interaction, which was not the case for TrmI. Finally, comparison of their active sites revealed no obvious similarities. We identified three conserved residues (D170, Y78, Y194) in TrmI that are essential to position the substrate adenine in the catalytic pocket of TrmI. For TrmK, the combination of residue conservations, NMR chemical shift mapping together with molecular docking was essential to identify key residues (R9, Q121, N123, G95) and their role in adenine binding. R9 is probably the most important residue as it binds SAM and adenine. For TrmI, we proposed that the catalysis of m¹A by TrmI does not necessitate a deprotonation step that could be done by D170 but can be transferred directly given that they are oriented correctly. We think that the same situation is true for TrmK. We did not identify a residue that could be a base catalyst. In that case, we propose that a network of interactions between SAM, adenine and four residues of TrmK are sufficient to catalyse the methyl transfer.

The role of the C-terminal domain in TrmK remains a mystery. Here, we showed that the C-terminal domain is not involved in tRNA binding and that its structure represents a new fold that is not encountered in other methyltransferases. It resembles however to domains of the seryl-tRNA synthetase and to the ribosomal protein L29 (Figure 8). In these proteins, the

coiled-coil is used to bind RNA. It could be envisaged that the m¹A₂₂ function of TrmK is not its sole function, and that its C-terminal domain is used to interact with an other partner methylated by TrmK. Further studies will be necessary to investigate this point.

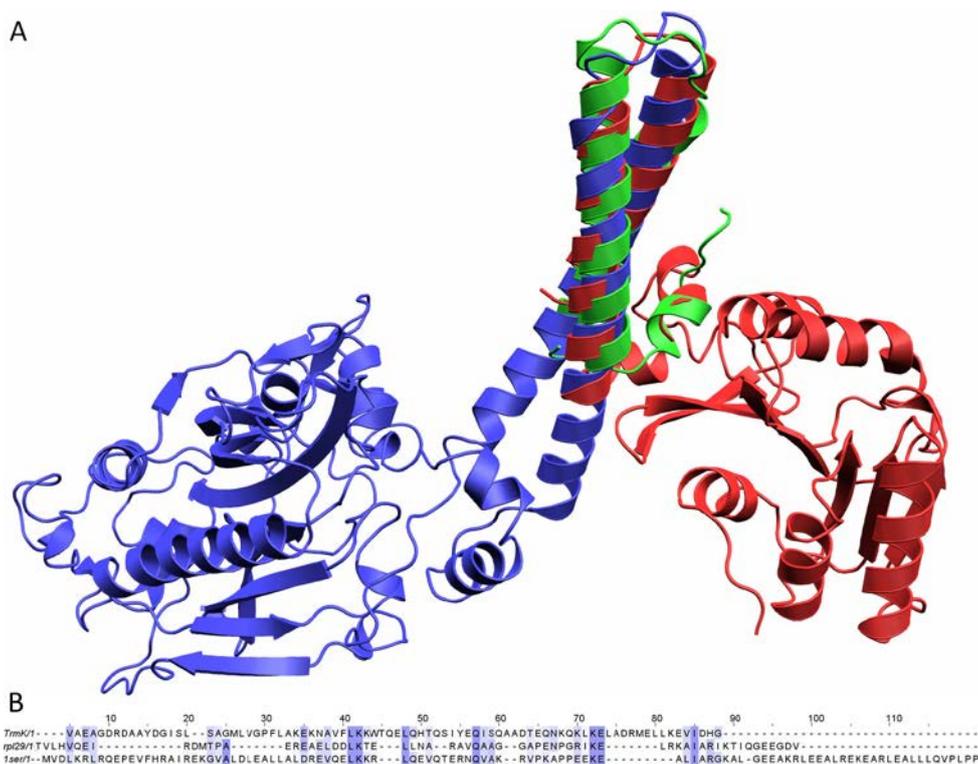


Figure 8: Comparison of C-terminal domain of TrmK with the Seryl-synthetase of X (PDB code) and the ribosomal protein L29 (PDB code).

A) Structural superposition of the C-terminal domain of TrmK (in green), L29 in red and the seryl-synthetase in blue. B) Sequence alignment of the domains in coiled-coil.

ACKNOWLEDGMENTS

Financial support from the TGIR-RMN-THC FR3050 CNRS for conducting the research is gratefully acknowledged. We thank the TGE ICSN at Gif sur Yvette for NMR measurement time and Nelly Morellet for helpful discussion and technical advice. The X-ray data were collected on the ID14-eh1 beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). We are grateful to Local Contact at ESRF for providing assistance in using the beamlines. We also thank Jessica Berlier, Marie Morel and Stéphane Ndio for their participations to this work.

REFERENCES

1. Machnicka, M.A., Olchowik, A., Grosjean, H. and Bujnicki, J.M. (2014) Distribution and frequencies of post-transcriptional modifications in tRNAs. *RNA Biol*, **11**, 1619-1629.
2. Kempnaers, M., Roovers, M., Oudjama, Y., Tkaczuk, K.L., Bujnicki, J.M. and Droogmans, L. (2010) New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Res*, **38**, 6533-6543.
3. Helm, M. and Attardi, G. (2004) Nuclear control of cloverleaf structure of human mitochondrial tRNA(Lys). *J Mol Biol*, **337**, 545-560.
4. Sprinzl, M. and Vassilenko, K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*, **33**, D139-140.
5. Droogmans, L., Roovers, M., Bujnicki, J.M., Tricot, C., Hartsch, T., Stalon, V. and Grosjean, H. (2003) Cloning and characterization of tRNA (m1A58) methyltransferase (TrmI) from *Thermus thermophilus* HB27, a protein required for cell growth at extreme temperatures. *Nucleic Acids Res*, **31**, 2148-2156.
6. Barraud, P., Golinelli-Pimpaneau, B., Atmanene, C., Sanglier, S., Van Dorsselaer, A., Droogmans, L., Dardel, F. and Tisne, C. (2008) Crystal structure of *Thermus thermophilus* tRNA m1A58 methyltransferase and biophysical characterization of its interaction with tRNA. *J Mol Biol*, **377**, 535-550.
7. Guelorget, A., Barraud, P., Tisne, C. and Golinelli-Pimpaneau, B. (2011) Structural comparison of tRNA m(1)A58 methyltransferases revealed different molecular strategies to maintain their oligomeric architecture under extreme conditions. *BMC Struct Biol*, **11**, 48.
8. Takuma, H., Ushio, N., Minoji, M., Kazayama, A., Shigi, N., Hirata, A., Tomikawa, C., Ochi, A. and Hori, H. (2015) Substrate tRNA recognition mechanism of eubacterial tRNA (m1A58) methyltransferase (TrmI). *J Biol Chem*, **290**, 5912-5925.
9. Degut, C., Ponchon, L., Folly-Klan, M., Barraud, P. and Tisne, C. (2015) The m1A modification in eubacterial tRNA: An overview of tRNA recognition and mechanism of catalysis by TrmI. *Biophys Chem*.
10. Varshney, U., Ramesh, V., Madabushi, A., Gaur, R., Subramanya, H.S. and RajBhandary, U.L. (2004) *Mycobacterium tuberculosis* Rv2118c codes for a single-component homotetrameric m1A58 tRNA methyltransferase. *Nucleic Acids Res*, **32**, 1018-1027.
11. Roovers, M., Wouters, J., Bujnicki, J.M., Tricot, C., Stalon, V., Grosjean, H. and Droogmans, L. (2004) A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase. *Nucleic Acids Res*, **32**, 465-476.
12. Guelorget, A., Roovers, M., Guérineau, V., Barbey, C., Li, X. and Golinelli-Pimpaneau, B. (2010) Insights into the hyperthermostability and unusual region-specificity of archaeal *Pyrococcus abyssi* tRNA m1A57/58 methyltransferase. *Nucleic Acids Res*, **38**, 6206-6218.
13. Raettig, R., Kersten, H., Weissenbach, J. and Dirheimer, G. (1977) Methylation of an adenosine in the D-loop of specific transfer RNAs from yeast by a procaryotic tRNA (adenine-1) methyltransferase. *Nucleic Acids Res*, **4**, 1769-1782.
14. Kersten, H., Raettig, R., Weissenbach, J. and Dirheimer, G. (1978) Recognition of individual procaryotic and eucaryotic transfer-ribonucleic acids by B subtilis

- adenine-1-methyltransferase specific for the dihydrouridine loop. *Nucleic Acids Res*, **5**, 3033-3042.
15. Roovers, M., Kaminska, K.H., Tkaczuk, K.L., Gigot, D., Droogmans, L. and Bujnicki, J.M. (2008) The YqfN protein of *Bacillus subtilis* is the tRNA: m1A22 methyltransferase (TrmK). *Nucleic Acids Res*, **36**, 3252-3262.
 16. Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S.L., Hughes, T.R., Grayhack, E.J. and Phizicky, E.M. (2006) Rapid tRNA decay can result from lack of nonessential modifications. *Mol Cell*, **21**, 87-96.
 17. Kotelawala, L., Grayhack, E.J. and Phizicky, E.M. (2008) Identification of yeast tRNA Um(44) 2'-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA(Ser) species. *Rna*, **14**, 158-169.
 18. Thanassi, J.A., Hartman-Neumann, S.L., Dougherty, T.J., Dougherty, B.A. and Pucci, M.J. (2002) Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic Acids Res*, **30**, 3152-3162.
 19. Ta, H.M. and Kim, K.K. (2010) Crystal structure of *Streptococcus pneumoniae* Sp1610, a putative tRNA methyltransferase, in complex with S-adenosyl-L-methionine. *Protein Sci*, **19**, 617-624.
 20. Dégut, C., Barraud, P., Larue, V. and Tisné, C. (2015) Backbone resonance assignments of the m1A22 tRNA methyltransferase TrmK from *Bacillus subtilis*. *Biomol NMR Assign*, **in preparation**.
 21. Degut, C., Monod, A., Brachet, F., Crepin, T. and Tisne, C. (2016) In Vitro/In Vivo Production of tRNA for X-Ray Studies. *Methods Mol Biol*, **1320**, 37-57.
 22. Dardel, F. (1994) MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Comput Appl Biosci*, **10**, 273-275.
 23. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213-221.
 24. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501.
 25. Bakan, A., Meireles, L.M. and Bahar, I. (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics*, **27**, 1575-1577.
 26. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph*, **14**, 33-38, 27-38.
 27. Pervushin, K.V., Wider, G. and Wuthrich, K. (1998) Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [15N,1H]-TROSY. *J Biomol NMR*, **12**, 345-348.
 28. Orekhov, V.Y. and Jaravine, V.A. (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog Nucl Magn Reson Spectrosc*, **59**, 271-292.
 29. Dominguez, C., Boelens, R. and Bonvin, A.M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc*, **125**, 1731-1737.
 30. de Vries, S.J., van Dijk, A.D., Krzeminski, M., van Dijk, M., Thureau, A., Hsu, V., Wassenaar, T. and Bonvin, A.M. (2007) HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins*, **69**, 726-733.

31. Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., Iwashita, C. and Hori, H. (2011) Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res*, **39**, 2304-2318.
32. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, **372**, 774-797.
33. Karplus, P.A. and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science*, **336**, 1030-1033.
34. Holm, L. and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res*, **26**, 316-319.
35. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*, **98**, 10037-10041.
36. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*, **32**, W665-667.

Supplementary materials

Table S1: Sequences of oligonucleotides used to produce the plasmid expressing TrmK (C35S, C152S)

Name	Sequence	Use
JC360	GAAGAGGACGGAAAGTCATATGAGGTGCTGGTG	Mutagenesis trmK_C35S
JC360b	CACCAGCACCTCATATGACTTTCCGTCTCTTC	Mutagenesis trmK_C35S
JC362	GATTCAGCACCGCATAACTAGGAAGGTAGGCGTG	Mutagenesis trmK_C152S
JC362b	GATTCAGCACCGCATAACTAGGAAGGTAGGCGTG	Mutagenesis trmK_C152S

	Acceptor arm	D arm	D loop	D arm	Ac arm	Anticodon loop	Ac arm	Variable loop	T arm	T loop	T arm	Acceptor arm		
	1	8	10	14	22	27	32	39	44	49	53	61	66	74
<i>Bacillus s. 1423 Ser 5GA</i>	GGAGGAA	5ACCCA	AGUCUGGC	DGA	GGGA	UCGGU	CU	GA	AACCGAC	AGGGUG-UCAA--AGCCCG-CGGGGT	PCGAAU	CCCUC	JUCCUC	CGCCA
<i>Bacillus s. 1423 Ser GCU</i>	GGAGAAG	UACUCA	AGU--GGC	DGA	GAG	CGCCC	UGGUC	AGGGUGU	GUCGC-GUAA--GCGGC-CGAGGG	TPCAAU	CCCUC	JUCCUC	CGCCA	
<i>Bacillus s. 1423 Ser GGA</i>	GGAGAGC	5GUCC	GAGU--GG	DGA	GAG	CGCCC	UGGUC	AGGGUGU	GUCGC-GUAA--GCGGC-CGAGGG	TPCAAU	CCCUC	JUCCUC	CGCCA	
<i>Bacillus s. 1423 Tyr QUA</i>	GGAGGGG	SAGCGA	AGU--GGC	UAA	CGC	CGGAC	UUA	APCCGCU	CCC--UCA--GGGU	CGCAG	TPCGAAU	CUGCC	CCCCU	CCACCA
<i>Bacillus s. 1423 Tyr QUA</i>	GGAGGGG	SAGCGA	AGU--GGC	UAA	CGC	CGGAC	UUA	APCCGCU	CCC--UCA--GGGU	CGCAG	TPCGAAU	CUGCC	CCCCU	CCACCA
<i>Geobacillus 1422 Leu BAA</i>	GC CGAUG	GGCGGAAU	D--GGCA	-G	CGC	CGGAC	UUA	APCCGCU	CCC--UCA--GGGU	CGCAG	TPCGAAU	CUGCC	CCCCU	CCACCA
<i>Geobacillus 1422 Tyr QUA</i>	GGAGGGG	SAGCGA	AGU--GGC	UAA	CGC	CGGAC	UUA	APCCGCU	CCC--UCA--GGGU	CGCAG	TPCGAAU	CUGCC	CCCCU	CCACCA
<i>Mycoplasma 2095 Cys GCA</i>	GGCAAC	ANGGCA	AGC--GGC	D--A	GGC	AGGGU	CUGCA	CACCCUGA	-----U-CAUCGGU	PCGAAU	CCGAU	JGUUGC	UCCA	
<i>Mycoplasma 2095 Gln IUG</i>	UGGGCU	ANAGCCA	AGC--GGC	D--A	GGC	AGGGU	CUGCA	CACCCUGA	-----U-CAUCGGU	PCGAAU	CCGAU	JGUUGC	UCCA	
<i>Mycoplasma 2095 Glu IUC</i>	GGCCUGU	UGGUGA	AGC--GGC	D--A	CAC	ACGGU	UUA	UCCGUGA	-----CACACGGU	PCGAAU	CCCGU	ACAGGC	UACCA	
<i>Mycoplasma 2095 His GUK</i>	GC CGUAG	UGGUGA	AGU--GGC	D--A	CAC	ACGGU	UUA	UCCGUGA	-----CACACGGU	PCGAAU	CCCGU	ACAGGC	UACCA	
<i>Mycoplasma 2095 Leu JAA</i>	CCCCAAG	NGGCGAAU	A--GGD	A--G	CGC	AUUGA	CU	AA--APCCAAC	GGGC--UUAAU--AUCUGU	GCCGGU	PCAAU	CCGGC	CUUGGG	GACCA
<i>Mycoplasma 2095 Leu BAA</i>	GC CGUUU	UGGCGAAU	D--GGCA	-G	CGC	AUUGA	CU	AA--APCCAAC	GGGC--UUAAU--AUCUGU	GCCGGU	PCAAU	CCGGC	CUUGGG	GACCA
<i>Mycoplasma 2095 Leu UAG</i>	GGGGGA	UNGGCGAAU	D--GGCA	-G	CGC	AUUGA	CU	AA--APCCAAC	GGGC--UUAAU--AUCUGU	GCCGGU	PCAAU	CCGGC	CUUGGG	GACCA
<i>Mycoplasma 2095 Ser GCU</i>	GGGUUA	ANACUCA	AGU--GGC	D--A	GAG	SACCC	CUGCU	AGGGUGU	AGGUCGGUCU--CCGGC--CGAGGG	UPCGAGU	CCCUC	JUAACC	CGCCA	
<i>Mycoplasma 2095 Ser UGA</i>	GGAAGA	ANACCCA	AGUCCGGC	DGA	GGGA	UCGGU	CU	AA--APCCAAC	GGGC--UUAAU--AUCUGU	GCCGGU	PCAAU	CCGGC	CUUGGG	GACCA
<i>Mycoplasma 2095 Tyr GUA</i>	SGAGGGG	UABCGA	AGU--GGC	D--A	CGC	CGGUC	UUA	APCCACUU	CC--UUAC--GGU	CGGGU	PCGAAU	CCCUC	CCCCU	CCACCA

Figure S1: Alignment of tRNA sequences for which m¹A is found at position 22 [1].

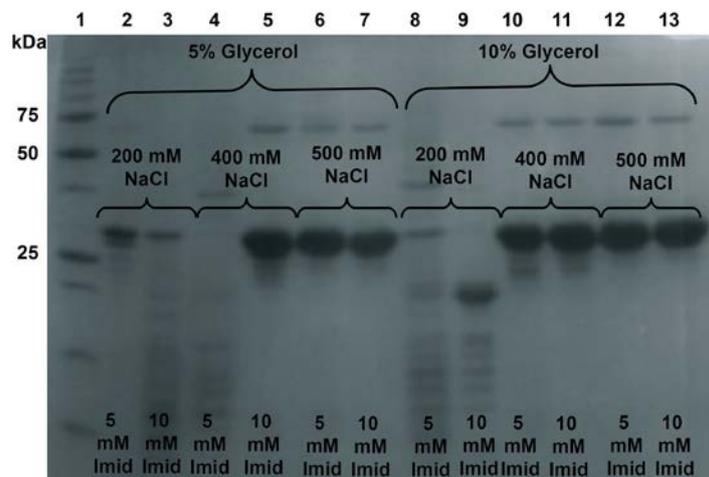


Figure S2: SDS-PAGE gel showing the impact of glycerol, NaCl and imidazole concentrations on TrmK proteolysis. Samples of TrmK in 50 mM Tris.Cl pH 8.0, 500 mM NaCl, 300 mM imidazole concentrated at 5 mg/ml were put into dialysis buttons to test the stability of TrmK regarding the buffer composition. Dialysis buttons were placed into Linbro plates of 24 wells filled with 2 ml of buffer. As shown in this figure, two concentrations of glycerol (5% and 10%), 3 concentrations of NaCl (200, 400 and 500 mM) and 2 concentrations of imidazole (5 or 10 mM) were tested. Lane 1: molecular weight marker, Lane 2 to 13: analysis of the degradation state of TrmK depending on their buffer conditions after 2 weeks at 18°C. As a result, a high concentration of NaCl is necessary to minimize the proteolysis together with the presence of glycerol. Imidazole is dispensable.

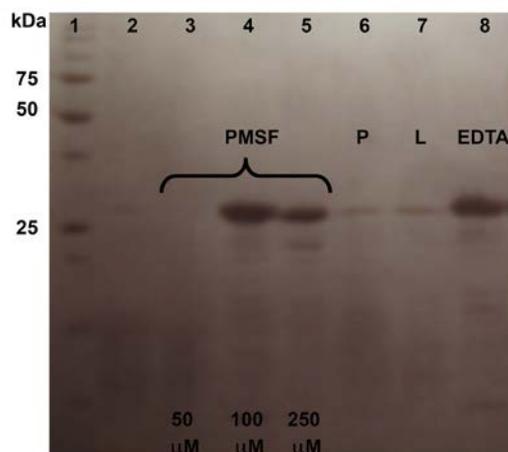


Figure S3: SDS-PAGE gel showing the impact of antiproteases on the degradation of TrmK. Lane 1: molecular weight marker, Lane 2: Reference sample of TrmK without addition of any antiproteases (12 mg/mL in 25 mM phosphate buffer pH 7.0, 500 mM NaCl, 1% glycerol); Lanes 3, 4 and 5: after addition of PMSF 50, 100 and 250 μM in the reference sample; Lane 6: after addition of Pepstatine A (1 $\mu\text{g}\cdot\text{mL}^{-1}$) in the reference sample, Lane 7: after addition of Leupeptine (1 $\mu\text{g}\cdot\text{mL}^{-1}$) in the reference sample, Lane 8: after addition of EDTA (1 mM) in the reference sample. The PMSF at concentration superior to 100 μM and EDTA at 1 mM stopped the proteolysis of TrmK.

<i>BsTrmK</i>		α_z	β_1	α_A	β_2	α_B	
<i>BsTrmK</i>	1MNELKLSKRRLQTVAEYTPNGA..V	ADIGSDHAYLP	CYAVLNHKKASGAIAGEITD	GFFLSAKROVEKSGLN		
EL03696	1MNELKLSKRRLQTVAEYITKGA..V	ADIGSDHAYLP	CYCVLNRLASAAIAGEITD	GFFLSAKQOV	EKLQLS	
<i>Estratos</i>	1MNMNLSKRRLKRVADFPKEA..VF	ADIGSDHAYLP	CYAILHHKALKAIAGEITD	GPLQSAKQOV	HRIELD	
<i>B_Amyl</i>	1MKLSKRRLQSVAEYTPPGA..LV	ADIGSDHAYLP	CYAVLNHLASRAIAGEITD	GFFLSAKQOV	EKLDSL	
<i>Bsp</i>	1	MIVKEKEPEKSCDIVNEMKLSKRRLQSVAEYTPPGA..LV	ADIGSDHAYLP	CYAVLNHLASRAIAGEITD	GFFLSAKQOV	EKLDSL	
<i>Anoxyflavith</i>	1MNVNELQSKRRLQTVASFPKGS..V	ADIGSDHAYLP	CYAYLHGKITKAIAGEVAD	GFFRSAKQOV	EKSGLS	
<i>Geosp</i>	1MNEFRLSKRRLQTVASFPKGA..K	ADIGSDHAYLP	CYAYLHGKITKAIAGEVAD	GFFRSAKQOV	ERTGLK	
<i>Banthracis</i>	1MNEVKLSKRRLQTVASFPVGS..T	ADIGSDHAYLP	CYTIINNIATKAIAGEVVD	GFFRSAQATV	AESGLQ	
<i>Bthuringiensis</i>	1MNEVKLSKRRLQTVASFPVGS..T	ADIGSDHAYLP	CYTIINNIATKAIAGEVVD	GFFRSAQATV	AESGLQ	
<i>Epumilus</i>	1MNMNLSKRRLKRVADFPDEA..VF	ADIGSDHAYLP	CYAILHHKAVKAIAGEITE	GPLQSAKQOV	HRIELD	
<i>Geobthermogluco</i>	1MNEFRLSKRRLQTVASFPKGA..T	ADIGSDHAYLP	CYAYLHGKITKAIAGEVAD	GFFRSAKQOV	ERTGLQ	
<i>Vcolerae</i>	1MKLGKRLQLVQVQKRDYD.H	ADIGSDHAYLP	IAALQKHPSTVHFVDIV	PSLIDKVT	LIRYFPA	
<i>Bhalodurans</i>	1MNVQQLSERLQTVASFPAGA..R	ADIGSDHAYLP	CYLCVKGVEFAIAGEVNE	GFYQSAVEQ	QKVNLI	
<i>Bcereus</i>	1MNEVKLSKRRLQTVASFPVGS..T	ADIGSDHAYLP	CYTIINNIATKAIAGEVVD	GFFRSAQATV	AESGLQ	
<i>Cclostridium</i>	1MELISNRRLKISSEVMDRCE..I	ADIGSDHAYLP	IYLLKNNLCRRFIASVDNA	GPLKKAANNI	NKYKLE	
<i>Dradiodurans</i>	1MTLPALDARLEAVLHLIRAIV..H	ADIGSDHAYLP	LRVLRGGVSRCAVLELNP	GPLALARRMV	ARARLT	
<i>Llactis</i>	1MNEIKLSKRRLKAVANVYDNGA..R	ADIGSDHAYLP	TYLQKLVDFAVAGEVVK	GFFEIAKNHV	SEADLS	
<i>Mgenitalium</i>	1MKKRIITIANVQSFNPKLVYD	ADIGSDHAYLP	SYLTKNQNLITVNSDI	SKNALLSNYQV	KNNNNI	
<i>Mpneumoniae</i>	1MKRRIITIANVQSQKPKAFYD	ADIGSDHAYLP	ARELVQNNP	LVGVNSDI	SANALGAKGL	KDHTNM
<i>Saureus</i>	1MISLNRLITVSRFLKQG..T	ADIGSDHAYLP	IYAILQNHLCFCVAGEV	GFYQAAVNV	AAHNLV	
<i>Spneumoniae</i>	1MISKRRLQTVASFPVGS..T	ADIGSDHAYLP	IELVERGQKSAIAGEVVE	GFYQSAVNV	EAHGLK	
<i>Spycogenes</i>	1MDSQLSNRLQTVASFPKGV..L	ADIGSDHAYLP	IFLVTNQISAIAGEVVR	GFYLSAKNV	TSRGLA	
<i>Saustralis</i>	1MEFTKLSNRRLQTVASFPAGA..R	ADIGSDHAYLP	IALIQEGKFAIAGEVVE	GFYQSAQNV	ADNGLE	
<i>Uparvum</i>	1MNPKTRIAFIADLINDAQN..V	ADIGSDHAYLP	KILLVNNKASHVTNIEV	NOGPELNGN	NLTKNLL	
<i>Linnocua</i>	1MNEEQLSKRRLQTVASFPKGA..R	ADIGSDHAYLP	CFAILKNNIASFVAGEV	NOGPELNGN	RTSGLT	
<i>Lmonocytogenes</i>	1MSLNEEQLSKRRLQTVASFPKGA..R	ADIGSDHAYLP	CFAILKNNIASFVAGEV	NOGPELNGN	RTSGLT	
<i>Gadiacens</i>	1MIELGERLQTVASFPKDS..K	ADIGSDHAYLP	VYLLQNNQITSAIAGEV	NOGPELNGN	RTSGLT	
<i>Lgastrius</i>	1MDALHLSQRRLQTVASFPVGS..T	ADIGSDHAYLP	AYLLNQRTEFAIAGEV	NOGPELNGN	RTSGLT	
<i>Faenibacillus</i>	1MNVNLSKRRLQTVASFPKGA..R	ADIGSDHAYLP	VYLLQNNQITSAIAGEV	NOGPELNGN	RTSGLT	
<i>Lentibacillus</i>	1MNNNLSKRRLQTVASFPKGA..R	ADIGSDHAYLP	CYVCLHDEHARAIAGEV	NOGPELNGN	RTSGLT	
<i>Emegaterium</i>	1MNEHLNLSRRLQTVASFPKGA..R	ADIGSDHAYLP	CFAYLNGYISGAVAGEITE	NOGPELNGN	RTSGLT	
<i>Ecasseliflavus</i>	1MNEHALSRLQTVASFPVGS..T	ADIGSDHAYLP	VALMLQNKITYAVAGEV	NOGPELNGN	RTSGLT	

<i>BsTrmK</i>		β_3	α_c	β_4	α_d	η_1	β_5	α_e	β_6							
<i>BsTrmK</i>	70	SHITSVRQ	...GDL	EVTKKGE	ADAIT	AGMCGALIAH	ILEAG	..KDKL	TGKE	RLILQPN	IHAVH	IREW	ILYKER	VALID	EVIL	
EL03696	70	SLITSVRK	...GDL	EVIEKGE	ADVAIT	AGMGGSLIAH	ILKTG	..KDKL	SGKE	RLVLQPN	IHAHH	IREW	ILYQEG	VALINE	EVIL	
<i>Estratos</i>	70	EQITSVRK	...GDL	EVIEKGE	VNAVIT	AGMGGALIAS	ILHEG	..KDKL	TGQE	RLILQPN	IHAHH	IREW	ILYQEG	VALINE	EVIL	
<i>B_Amyl</i>	67	TLITSVRK	...GDL	SVIEKGE	AGAIT	AGMCGALIAH	ILESG	..KDKL	TGRE	RLILQPN	IHAVH	IREW	ILYQEG	VALINE	EVIL	
<i>Bsp</i>	84	TLITSVRK	...GDL	SVIEKGE	AGAIT	AGMCGALIAH	ILESG	..KDKL	TGRE	RLILQPN	IHAVH	IREW	ILYQEG	VALINE	EVIL	
<i>Anoxyflavith</i>	72	HVITSVRK	...GDL	QVVPNE	VDCIT	AGMGGTLISN	ILENG	..KEK	LENVK	RLILQPN	VGAHI	VRQW	LLNNEW	LIDERIL		
<i>Geosp</i>	70	DVITSVRK	...GDL	EVIEKGE	VDCIT	AGMGGTLIAN	ILEAG	..KDKL	LLGVK	RLILQPN	VGAHI	VRQW	LLNNEW	LIDERIL		
<i>Banthracis</i>	70	DKVDVRK	...GNGL	AVITPGE	VDDIT	VAGMGGALIRD	ILESG	..KEK	LEGVT	RLILQPN	IAAHH	IREW	FIENGW	LIEHEKIV		
<i>Bthuringiensis</i>	70	DKVDVRK	...GNGL	AVITPGE	VDDIT	VAGMGGALIRD	ILESG	..KEK	LEGVT	RLILQPN	IAAHH	IREW	FIENGW	LIEHEKIV		
<i>Epumilus</i>	70	DQITSVRK	...GNGL	EVIEKGE	VNAVIT	AGMGGALIAS	ILNDG	..KDKL	TGHE	RLILQPN	IHAHH	IREW	ILYQEG	VALINE	EVIL	
<i>Geobthermogluco</i>	70	DVITSVRK	...GDL	EVIEKGE	VDCIT	AGMGGTLIAN	ILEAG	..KDKL	SGVK	RLILQPN	VGAHI	VRQW	LLNNEW	LIDERIL		
<i>Vcolerae</i>	68	TDSPRWRTYCLDVRDLPEANAGS	HLVIT	AGVGGDLMT	TEFAELA	KRHPSMTF	DLILCP	VHHTYT	LRQI	IALNAE	LKSERLV					
<i>Bhalodurans</i>	70	HVITSVRK	...GDL	AVIRADQ	DDTVT	AGMGGTLIS	ILESG	..KRN	LQTVR	RLILQPN	VHGKA	IRQW	LSNDG	NLQAEAIL		
<i>Bcereus</i>	70	DKVDVRK	...GNGL	AVITPGE	VDDIT	VAGMGGALIRD	ILESG	..KEK	LEGVT	RLILQPN	IAAHH	IREW	FIENGW	LIEHEKIV		
<i>Cclostridium</i>	67	DKIECRLL	...GS	LVNKLKEE	VDDVV	AGMGGNLRD	ILEER	..LDI	FKKLK	YAVLQPV	QNP	PEI	REYL	IKRGN	ILIEEQLC	
<i>Dradiodurans</i>	69	EQIDVRQ	...GDL	FAPIQPGE	LGSASVT	GMGAYTIR	ILDRAG	..ERL	PPAL	VLQPN	DSP	LW	LRTW	ARENSYH	LRAERLL	
<i>Llactis</i>	70	DRITVRL	...ANG	LGAIENTDK	IDTIV	AGMGGILIS	ILEAG	..KEK	LGHVK	RLILQPN	NHEES	LRQW	LVNHFV	IKKEBIL		
<i>Mgenitalium</i>	67	HFFVSDG	...FNN	LPELNINP	..KIG	VAGLGGK	IKINIS	QK	...ENF	IN	RFVQP	CSN	IL	ERSF	LSNSW	IVNTELV
<i>Mpneumoniae</i>	67	HFITSDG	...FDL	LPNLDLP	..KAV	AGLGGK	IKIMQA	...NNF	LK	HFVQP	CSN	IL	ERSF	LSNSW	IVNTELV	
<i>Saureus</i>	67	ERITVRL	...GDL	SVIHPEDEV	DNIT	AGMGGPLIAK	ILKDG	..OD	KLSQH	RLILQPN	IQTEN	LRQI	QQLN	YV	IDTETLV	
<i>Spneumoniae</i>	66	EKIQVRL	...ANG	LAFAEETD	QVSVIT	AGMGGRLIAR	ILEEG	..LG	KLANVE	RLILQPN	NNREDD	LRIN	IQDHGF	QIVAESIL		
<i>Spycogenes</i>	69	EHIQVRL	...ANG	LAFAEETD	QVSVIT	AGMGGRLIAR	ILEAG	..KEK	LQIE	RLVLQPN	NNREDD	LRAN	LSVNAF	KIVAEITM		
<i>Saustralis</i>	70	DKIEVRL	...ANG	LAFAEETD	QVSVIT	AGMGGRLIAR	ILEAG	..LEK	LQIE	RLVLQPN	NNREDD	LRAN	LSVNAF	KIVAEITM		
<i>Uparvum</i>	67	KRTTNIL	...ND	GKGLILEQ	TFDYCV	AGMCA	TSIEILE	QN	...KN	QIKNY	ILQPN	TQ	SYK	IRKYL	NNHSY	QIKVHIF
<i>Linnocua</i>	70	DQIDVRK	...GNGL	EVIEKDD	IDTIV	AGMGGTLIR	ILEEG	..AEK	LQIE	RLVLQPN	IAAW	LR	WSE	EKN	NLIS	SEAIL
<i>Lmonocytogenes</i>	72	EQIDVRK	...GNGL	AVIEKDD	IDTIV	AGMGGTLIR	ILEEG	..AAK	LQIE	RLVLQPN	IAAW	LR	WSE	EKN	NLIS	SEAIL
<i>Gadiacens</i>	68	ESITVRF	...GDL	QIILSKED	ITAVT	AGMGGELIS	IRILE	AG	FC	SGHLN	CKE	RLILQPN	VAEHF	VRE	IMNHSY	RITHETVV
<i>Lgastrius</i>	70	GSLEARL	...AD	GLVATVADQ	DDIT	AGMGGSLIS	ILDAH	..PEK	LVGVK	RLILQPN	VGESE	LRTW	LVQHHF	EIMTEBIL		
<i>Faenibacillus</i>	68	KQIEVRL	...GDL	SVILHDE	ADTIV	AGMGGSLIS	ILDAH	..AH	RL	SVK	VMVLP	QPN	VGESE	VRE	WLNHSY	GLADTEIL
<i>Lentibacillus</i>	71	DRVVRF	...GDL	QIILN	EREVTV	AGMGGPLIK	ITLID	GAN	..R	DGVE	QIIA	QPN	IDERR	WLNHSY	GLADTEIL	
<i>Emegaterium</i>	70	DVIDVRK	...GNGL	EVISPNE	VDCIT	AGMGGTLIT	MLEEG	..KAK	LEGVK	RLVLQPN	IGSHH	IR	WLNHSY	GLADTEIL		
<i>Ecasseliflavus</i>	70	EKITVRL	...AD	GLDAIQADK	IDTIV	AGMGGTLIR	ILEEG	..SR	RN	LAGVK	RLVLQPN	VGERT	LR	WLNHSY	GLADTEIL	

Chapitre III

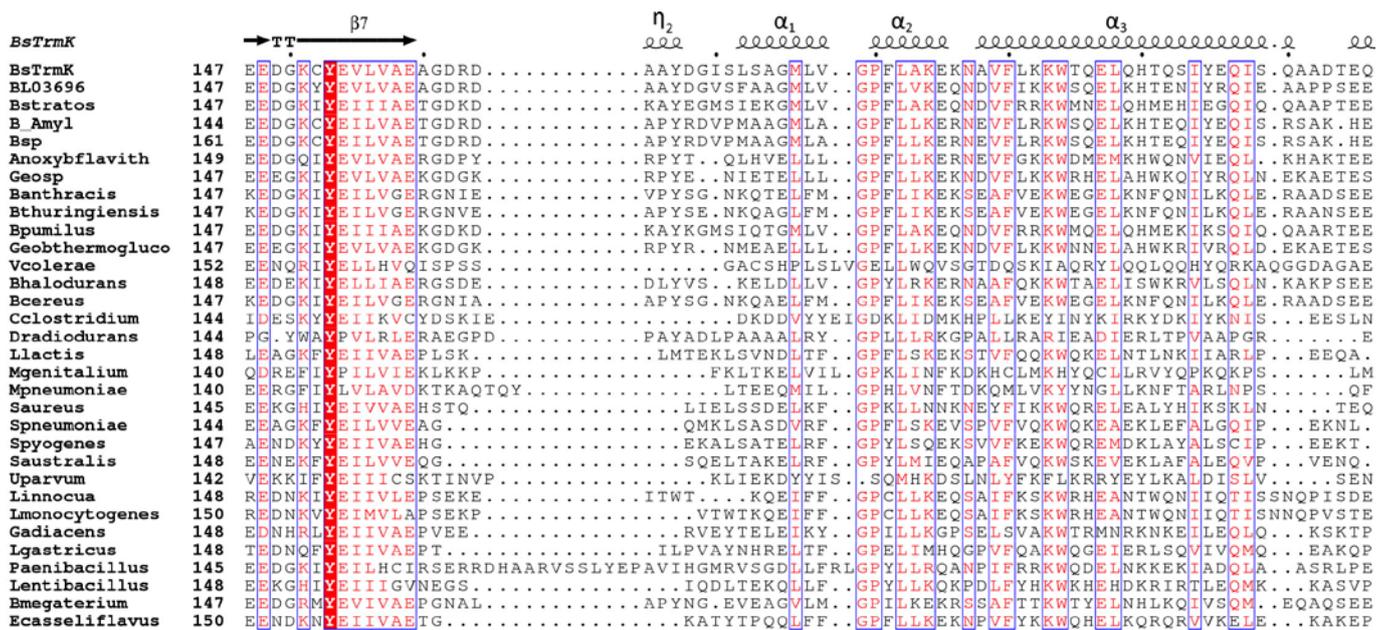
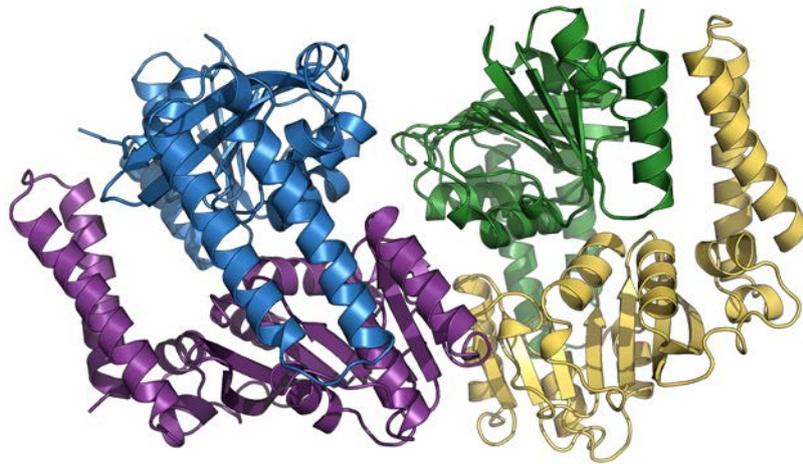


Figure S4: Multiple sequence alignment of *B. subtilis* TrmK and its homologues in the COG2834 family. The sequence of *B. subtilis* TrmK is aligned with those of *B. subtilis*, *B. lichenformis*, *B. stratosphericus*, *B. amyloliquefaciens*, *Bacillus* sp., *A. flavithermus*, *Geobacillus* sp., *B. anthracis*, *B. thuringiensis*, *B. pumilus*, *G. thermoglucosidasius*, *V. colerae*, *B. halodurans*, *B. cereus*, *C. acetobutylicum*, *D. radiodurans*, *L. lactis*, *M. genitalium*, *M. pneumoniae*, *S. aureus*, *S. pneumoniae*, *S. pyogenes*, *S. australis*, *U. parvum*, *L. innocua*, *L. monocytogenes*, *G. adiacens*, *L. gastricus*, *Paenibacillus* sp., *Lentibacillus* sp., *B. megaterium*, *E. casseliflavus*.

Sequences were aligned using CLUSTALW [2] and colored with ESPript [3]. The highly conserved and moderately conserved residues are represented by red boxes and red letters respectively. The secondary structure elements and the residue numbers of *B. subtilis* TrmK are indicated above the sequence alignment. α indicates α -helice, η : 3_{10} helices and β : β -strand.

A



B

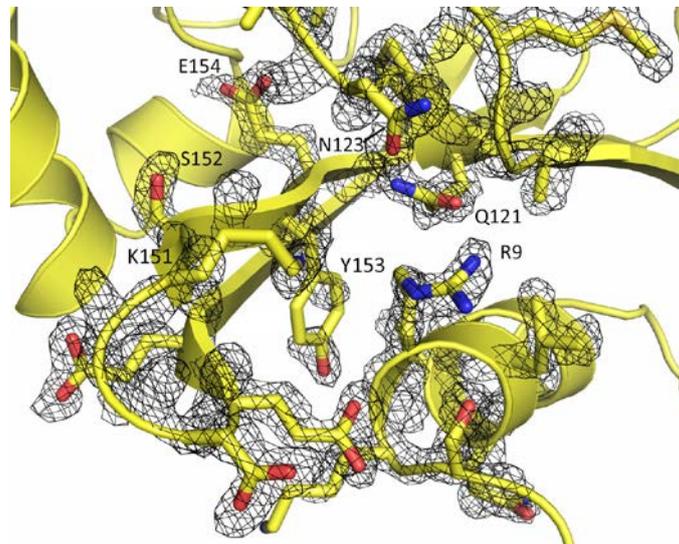


Figure S5: Crystal structure of *B. subtilis* TrmK. A) Four molecules of TrmK are present in the asymmetric unit. B) Density map ($2F_o - F_c$ at 2σ) of the adenine binding pocket. The density map of residues are very well defined, except for K151 and Y153 that are probably experiencing several conformations in the absence of SAM.

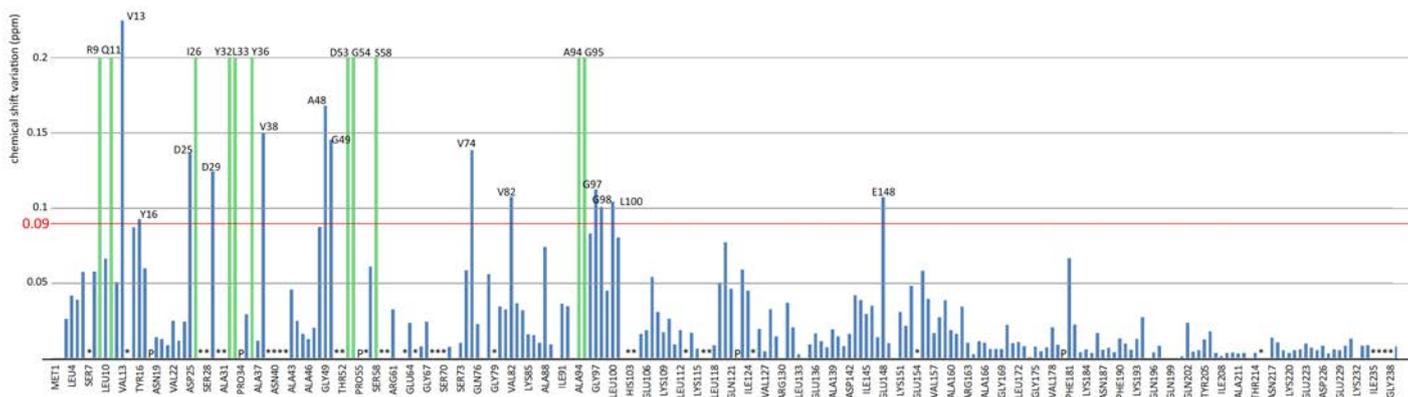


Figure S6: Chemical shift variations of TrmK amide groups upon SAH binding. Chemical shift variation are in ppm and the number of TrmK residues are in abscissa. Residues that disappeared upon addition of SAM are represented in green. Their chemical shift variations are arbitrary put to 0.2 ppm to easily identify them in the graph.

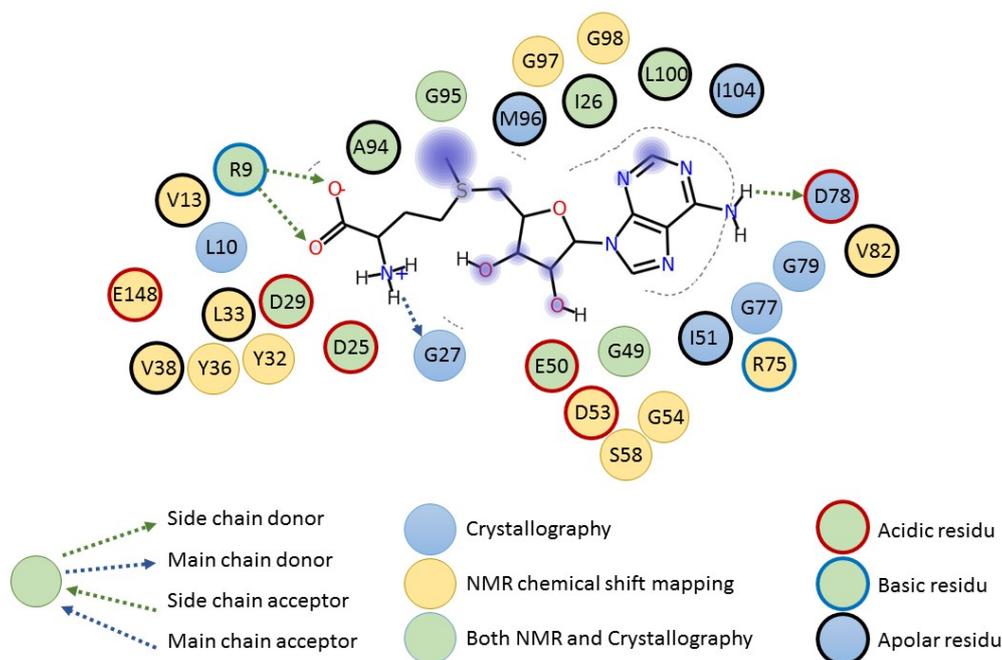


Figure S7: Analysis of the SAM binding site in TrmK structure and by NMR chemical shift mapping with the SAH. SAM was placed in the catalytic center of TrmK by superposing the TrmK structure with that of its ortholog from *S. pneumoniae* (PDB code 3KU1). Analysis of the contacts made between TrmK and SAH are indicated by blue circles or green ones if the contact were also observed by NMR. Residues experiencing chemical shift perturbations superior to 0.09 ppm are indicated in yellow circles. Hydrogen bonds are indicated by arrows whereas hydrophobic or van der Waals interactions are represented by circles.

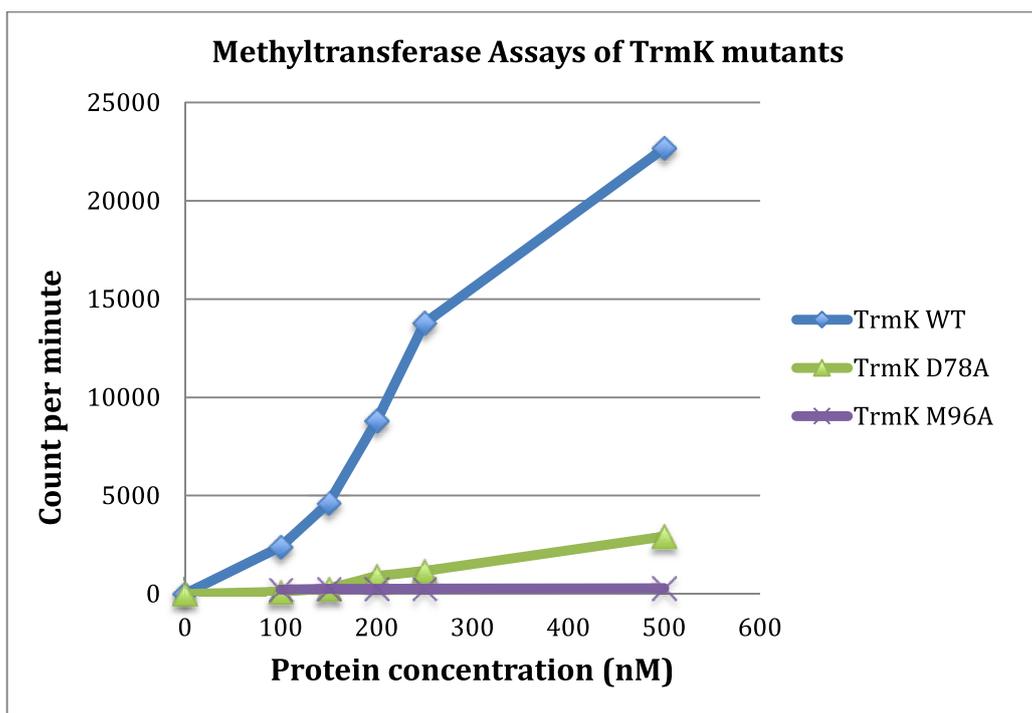


Figure S8: Methyltransferase assays of TrmK mutants.

The MTase activity at each concentration of TrmI mutant was determined by measuring the amount of ^{14}C transferred to total *E. coli* tRNA using methyl- ^{14}C SAM as methyl donor. The reaction mixture was 200 μL : 70 μg of total *E. coli* tRNA, 4 μL of methyl- ^{14}C SAM (130 dpm/pmol), the fraction containing TrmK oligomer and H_2O . The mixtures were incubated 1 hour at 37°C . The reaction was stopped by phenol extraction, and tRNA was TCA precipitated and captured on a GC/C Glass microfiber filter for scintillation counting.

References

1. Machnicka, M. A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., ... Grosjean, H. (2013). MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic acids research*, 41(Database issue), D262–7. doi:10.1093/nar/gks1007
2. Larkin, M. A., Blackshields, Brown, N. P., Chenna, McGettigan, P. A., McWilliam, ... Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947–2948. doi:10.1093/bioinformatics/btm404
3. Robert, & Gouet. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), W320–W324. doi:10.1093/nar/gku316

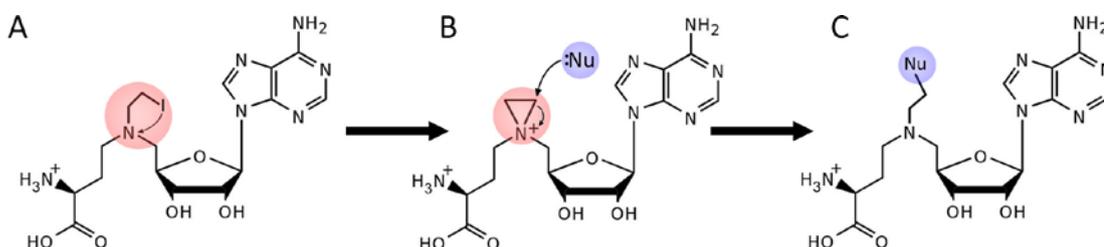
Chapitre III

Etudes de l'interaction de la méthyltransférase TrmK avec son substrat ARNt –
Article 4 Dégut et al (2015)

VII.3 Etude en cours et perspectives sur TrmK

Les études structurales sur TrmK vont se poursuivre. Afin d'affiner nos résultats, nous allons enregistrer l'empreinte de la protéine TrmK sur l'ARN. Pour cela, nous devons produire l'ARNt^{Ser} marqué à l'azote ¹⁵N par transcription *in vitro* en utilisant des UTP et GTP ¹⁵N afin d'effectuer l'attribution des groupements iminos de cet ARN. L'expérience d'empreinte RMN sur l'ARNt nous permettra de déterminer quels nucléotides de l'ARN sont en contact avec la protéine. Nous aurons alors la possibilité de proposer un modèle de l'interaction entre les deux partenaires.

Des études cristallographiques sont également en cours, en particulier avec un substrat "suicide" : la SAM-N-mustard (**Figure VII-1**) qui a été synthétisée dans le groupe du Pr. Mélanie Ethève-Quellejeu avec qui nous collaborons. Cette molécule est un analogue de la SAM qui porte une fonction réactive (dite moutarde azotée*) à la place du sulfonium (**Figure VII-1 A**). Cette fonction forme spontanément un aziridium (**Figure VII-1 B**) qui est un cycle tendu dont les carbones sont rendus électrophiles par la charge positive sur l'azote. L'un des carbones est attaqué par un nucléophile et forme une liaison covalente avec celui-ci (**Figure VII-1 C**). Cette réaction développée pour modifier de l'ADN [300], et a été utilisée pour la génération *in situ* de bisubstrat pour une méthyltransférase arginine des protéines (PRMT) [301]. Dans le cas de l'utilisation avec une m¹A méthyltransférase et de son substrat ARNt, les contraintes de proximité font que la réaction devrait se faire soit sur le N₁ soit sur le NH₂ exocyclique de l'adénine 22 dans le cas de TrmK et 58 pour TrmI.



* Cette appellation vient du gaz moutarde de formule Cl-(CH₂)₂-S-(CH₂)₂-Cl, dans notre cas l'atome de soufre est remplacé par un azote, d'où le terme de moutarde azotée.

Figure VII-1 A) La SAM-N-mustard est un analogue de la SAM portant une fonction moutarde azotée à la place du sulfonium. B) Cette fonction est très réactive et forme spontanément un aziridium qui est attaquée par un nucléophile (en bleu). Dans notre cas, le nucléophile est l'un des azotes de l'adénine de l'ARNt (le NH₂ exocyclique, ou le N₁).

L'idée de l'utilisation de ce substrat suicide est de former un complexe covalent ARNt-SAM qui reste bloqué au site actif. Plusieurs tests ont été menés et des cristaux ont été obtenus, mais de tailles trop petites pour que leurs diffractions aient pu être testées.

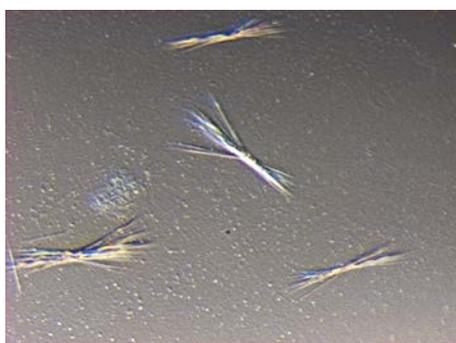
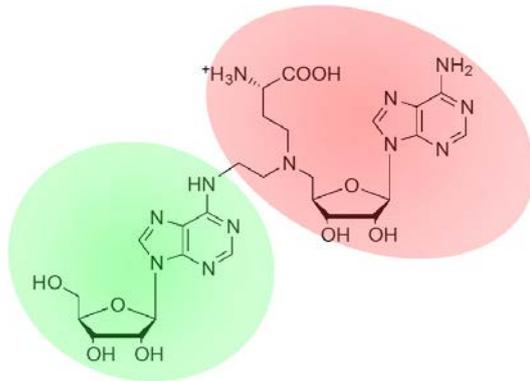


Figure VII-2 Cristaux du Complexe ARNt-SAM/TrmK obtenus dans 0.1M HEPES pH 7.5, 10% PEG 8000, 8% Ethylène Glycol.

Enfin, et toujours dans l'équipe du Pr Mélanie Ethève-Quellejeu des travaux sont en cours pour la production d'une molécule bisubstrat qui soit pour une part un analogue de la SAM, et pour une autre part une adénine (**Figure VII-3**). Cette molécule doit être utilisée aussi bien pour des essais de cristallogénèse que pour des empreintes RMN. Elle nous permettra une analyse fine du site actif de l'enzyme.

A



B

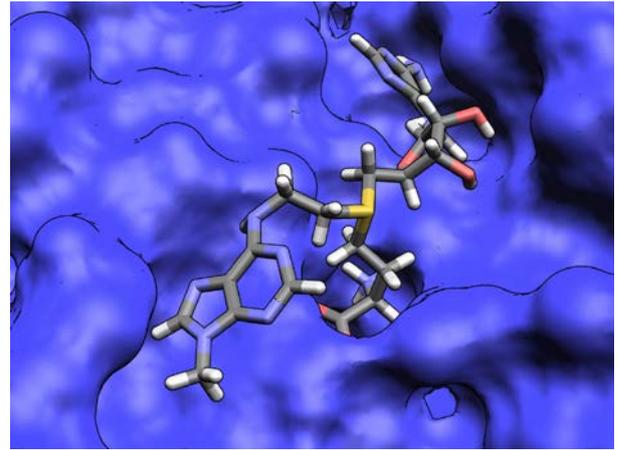


Figure VII-3 A) Analogue bisubstrat développé pour l'étude des méthyltransférases m^1A . La "partie SAM" de la molécule est en rouge, la "moitié adénosine" en vert. B) modèle d'un analogue bisubstrat dans le site actif de TrmK (le sucre de l'adénosine n'est pas modélisé et a été remplacé par un groupement méthyle).

Conclusion

La résolution de la structure des complexes entre enzyme de modification et ARNt est particulièrement difficile. La cristallographie aux rayons X est la méthode de choix pour obtenir la structure de ces complexes, mais leur cristallisation est souvent impossible du fait de la flexibilité des ARN et des faibles affinités mises en jeu entre ces deux partenaires. L'obtention de complexe covalent entre l'enzyme de modification et l'ARNt permet de maintenir ces partenaires assemblés afin d'aider à la cristallisation. La RMN peut aussi être un outil très puissant pour apporter des informations sur les interactions enzymes/ARNt en solution. Durant mon travail de thèse, j'ai exploré ces différentes voies pour étudier deux méthyltransférases m¹A des ARNt et leur interaction avec leurs substrats.

La première étude de mon travail de thèse a concerné l'enzyme TrmI de *T. thermophilus*, méthyltransférase m¹A₅₈ des ARNt, dont la structure cristallographique avait été résolue avant mon arrivée dans l'équipe de C. Tisné. J'ai résolu et affiné la structure de deux mutants du site actif de cette enzyme, qui avaient été précédemment identifiés comme essentiels pour la catalyse de la méthylation. Cette étude nous a permis, en particulier, de discuter de l'implication du domaine accessoire de TrmI dans l'accommodation de l'ARNt et de proposer deux mécanismes catalytiques compatibles avec nos données. Nous travaillons, également, en collaboration avec l'équipe de Ronald Micura (Université d'Innsbruck, Autriche), à la production et à la cristallisation de complexes covalents entre TrmI et des ARN substrats modifiés. Ces substrats sont des tiges-boucles analogues du bras T qui portent un nucléotide modifié capable de faire un pont disulfure avec une cystéine libre de la protéine. J'ai optimisé la formation de complexes covalents TrmI/bras T et mes travaux ont permis d'obtenir des premiers cristaux. Ils doivent être maintenant optimisés afin de pouvoir enregistrer des jeux de données de

diffraction aux rayons X. Ces résultats sont très encourageants et les tests doivent être maintenant généralisés à tous les mutants cystéine de TrmI que nous avons produits afin de mener de manière systématique des essais de cristallogénèse des complexes covalents.

La deuxième partie de mon travail de thèse a concerné l'étude structurale de l'enzyme TrmK de *B. subtilis*, méthyltransférase m¹A₂₂ des ARNt. Peu de données sont actuellement disponibles sur cette protéine. J'ai résolu la structure cristallographique de TrmK et ai étudié son interaction avec son substrat ARNt par la méthode d'empreinte RMN. Ces résultats alliés aux nombreux tests *in vitro* effectués dans l'équipe de Louis Droogmans (Université libre de Bruxelles, Belgique) nous ont permis de mieux comprendre les déterminants de cette interaction. J'ai ainsi montré que le domaine C-terminal non catalytique de TrmK, structuré en "coiled-coil", n'interagit pas avec le substrat ARNt. Des comparaisons structurales avec d'autres protéines interagissant avec les ARN nous laissaient pourtant supposer le contraire (**Figure C-1**). Par exemple, rpl29 est une petite protéine ribosomale structurée en trois hélices α , une très courte, et deux longues en "coiled-coil". Ces deux hélices interagissent avec l'ARN ribosomique et le reconnaissent par de multiples interactions. L'aaRS-seryl possède un domaine coiled-coil qui reconnaît l'ARNt^{Ser}, il est composé comme le domaine accessoire de TrmK de quatre hélices, dont deux très courtes et deux longues. Une autre fonction de ce domaine est donc à rechercher. L'enzyme pourrait méthyler un autre substrat ARN ou protéine reconnu par ce domaine C_{Ter}. Cette seconde fonction est peut-être en lien avec la phosphorylation postulée de TrmK dans certaines conditions cellulaires.

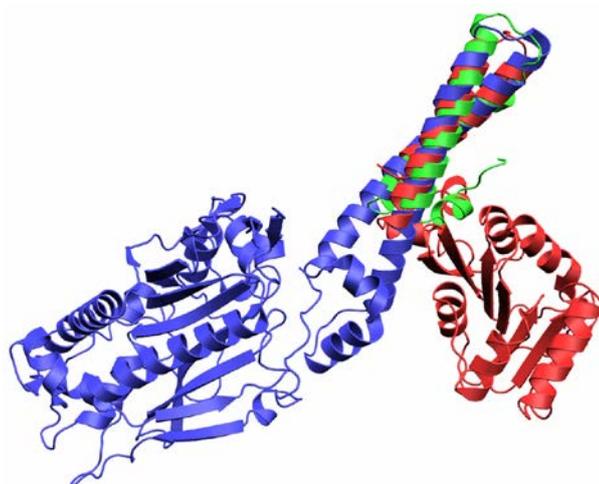


Figure C-1 : Superposition du domaine C_{Ter} de TrmK (en rouge), avec la protéine ribosomale rpl29 en vert (code d'accèsion 1S72), et le domaine accommodateur de l'ARN de l'aaRS serine en bleu (code d'accèsion pdb 1sry). La bonne superposition avec ces domaines laissait imaginer que le domaine C_{ter} de TrmK participait l'interaction avec l'ARNt.

Ces études structurales vont se poursuivre avec l'utilisation d'un substrat suicide, la SAM-N-mustard, ou d'un analogue bisubstrat SAM/Adénosine, pour des essais de co-cristallisation avec TrmI ou TrmK et des expériences d'empreinte RMN avec TrmK (collaboration avec le Pr Melanie Ethève-Quelquejeu, Université Paris Descartes).

TrmI et TrmK catalysant la même réaction, une comparaison structurale semblait évidente. On remarque tout d'abord que les domaines catalytiques se superposent très bien (RMSD de 2.6 Å sur les carbones α), à l'exception des brins β_6 et β_7 qui sont les plus variables dans les méthyltransférases de classe I. Pour TrmI, ces brins forment en particulier le feuillet central qui maintient le tétramère. Pour TrmK, ces brins sont en interaction étroite avec le domaine C_{Ter} que ces interactions rigidifient (**Figure C-2**). Il s'avère en fait que ce bon alignement structural est surtout en rapport avec leur appartenance à la classe I des méthyltransférases plutôt qu'à l'identité de la réaction que ces deux protéines catalysent. Si l'on considère la structure de COMT qui est une méthyltransférase de l'anabolisme des catécholamines, ne comportant qu'un domaine catalytique, la superposition

structurale est très bonne pour les deux protéines (2.4 Å pour TrmK et 3.8 Å pour TrmI). Par ailleurs, le site actif des deux protéines ne comporte aucun résidu en commun. On repère assez simplement dans la structure de TrmI une poche de fixation de l'adénine qui la positionne en face du méthyle de la SAM, la structure de TrmK ne semble pas comporter de telle poche. C'est certainement la raison pour laquelle l'ARNt entier est nécessaire à la méthylation chez TrmK. La superposition de la SAM pour les deux structures montre une conformation légèrement différente de celle-ci. Si la conformation générale est celle étendue adoptée dans toutes les MTase de classe I, on notera que le cycle de l'adénine est tourné de 180°, et l'acide aminé terminal de 45° (**Figure C-1 B**). Enfin les domaines supplémentaires au domaine catalytique sont très différents d'un point de vue structural. Celui de TrmK est constitué de quatre hélices α en C_{ter} de TrmK et n'a presque aucune mobilité par rapport au reste de la protéine. Pour TrmI, le domaine additionnel est une structure α/β , situé en N_{ter} qui est relativement flexible par rapport au reste de la protéine. De plus, TrmI est l'une des rares MTase de classe I tétramérique, alors que TrmK est monomérique. Il semble donc que ces protéines n'ont pas d'autre lien entre elles que celui qui relie toutes les méthyltransférases de classe I.

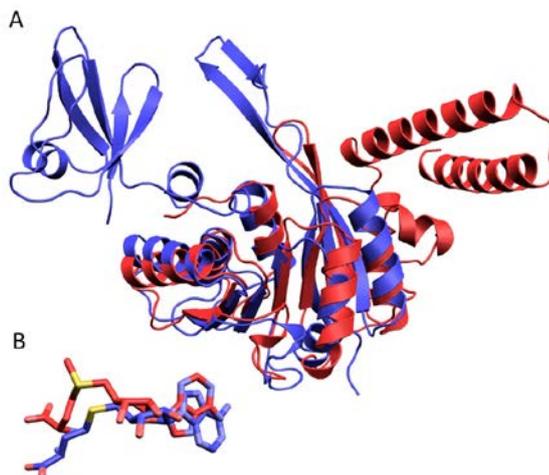


Figure C-2 : Comparaison structurale entre TrmI (bleu) et TrmK (rouge). A) Superposition des deux protéines, le domaine méthyltransférase est très facilement superposable avec un RMSD de 2.6Å, les domaines additionnelles sont très différents. B) Superposition de la SAM dans le site actif des deux protéines, la conformation

générale est la même, mais le cycle adénine est tourné de 180°, et le groupement acide aminé de 45°.

On peut rechercher l'origine de cette disparité dans le passé évolutif des méthyltransférases et des ARNt. Les ARNt sont des molécules centrales chez tous les êtres vivants connus, ils sont uniques par leur capacité à lier l'information d'un polymère d'acide nucléique et à transmettre cette information à un nouveau polymère protéique. La conservation extrême de cette structure indique qu'elle était chez LUCA* (last unique common ancestor) identique à ce qu'elle est aujourd'hui [302]. Les cinq familles d'ARN traçable jusqu'à cet ancêtre sont toutes liées à la traduction (ARNt, ARNm, ARNr, et RNaseP) et au trafic des protéines (SRP). Toutes les autres familles étant postérieures à LUCA [303]. Il semblerait que parmi ces cinq familles, la famille des ARNt soit la plus ancienne [302, 303]. La découverte récente de microribozyme, capable d'aminocycler des ARN [304], et la présence parmi les cinq familles d'ARN conservées de deux ribozymes, dont l'une dédiée à la maturation des ARNt (RNase P voir p.31) laisse entrevoir un monde ARN dont le métabolisme était déjà en partie centré sur les ARNt. Ces faits permettent de se rendre compte que les ARNt ne se sont pas adaptés à la machinerie actuelle de traduction, c'est au contraire tout le système de traduction, ribosome y compris, qui semble s'être structuré autour des ARNt.

Certaines études proposent de remonter plus loin encore dans le passé évolutif des ARNt. Certains gènes d'archée sont séparés en deux parties [305, 306]. Un gène code pour les 37 premiers nucléotides, un autre pour les nucléotides restants, les transcrits étant alors liés dans la cellule. Si cette architecture était tout d'abord vue comme un phénomène de dégradation du génome, il semblerait qu'il s'agisse en réalité d'un caractère ancestral [307]. Les proto-ARNt primordiaux seraient dans ce modèle des mini hélices correspondant à la moitié des ARNt actuels. Cette hypothèse est supportée par le fait que de nombreux ARNt actuels portent à la

* LUCA, le dernier ancêtre commun unique, est la cellule (ou le type cellulaire) qui est à l'origine de tous les règnes du vivant actuels. A différencier de la première cellule qui est très probablement très antérieure à LUCA.

position 37 un intron qui serait alors le résidu de la mise côte à côte de deux gènes de proto-ARNt [307]. De façon particulièrement intéressante, l'aminoacylation spontanée des mini hélices d'ARN montre une préférence très forte pour les acides aminés L [308]. La chiralité du ribose des ARN serait alors la source de la chiralité des protéines. Par ailleurs l'implication de demi-ARNt dans de très nombreux processus métaboliques (coupé précisément par l'angiogénine en position 37) évoque fortement ces proto-ARNt.

Les ARNt ont donc un passé évolutivement très riche, et il est probable que les modifications des bases aient toujours accompagné ces ARN. Huit modifications des ARNt, dont la m¹A₅₈, sont présentes dans tous les règnes (mitochondries comprises) signe qu'elles étaient très probablement présentes chez LUCA [125, 276]. La capacité de TrmI à méthyler un mini substrat analogue du bras T uniquement est peut-être à rapprocher de son statut de méthyltransférase primordiale [276], qui aurait alors été une méthyltransférase des mini hélices proto-ARNt ?

De nombreuses modifications sont apparues plus tard, après les différenciations des trois règnes. C'est probablement le cas de la m¹A₂₂ catalysée par TrmK. Les fonctions de ces modifications non primordiales sont à rechercher dans les spécificités des organismes qui les portent. Le problème est alors que nos connaissances des microorganismes sont limitées par les moyens d'études de ceux-ci. Il est probable que des systèmes essentiels en milieu naturel soient complètement masqués par les conditions de culture en laboratoire. Les études structurales des enzymes de modifications peuvent permettre d'identifier des fonctions et/ou de caractériser des mécanismes difficilement observables lorsque le micro-organisme est sorti de son milieu naturel.

Bibliographie

1. Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I., & Zamecnik, P. C. (1958). A soluble ribonucleic acid intermediate in protein synthesis. *The Journal of biological chemistry*, 231(1), 241–57.
2. Holley, R. W., Everett, G. A., Madison, J. T., & Zamir, A. (1965). Nucleotide sequences in the yeast alanine transfer ribonucleic acid. *The Journal of biological chemistry*, 240, 2122–8.
3. Sprinzl, M., & Vassilenko, K. S. (2005). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic acids research*, 33(Database issue), D139–40. doi:10.1093/nar/gki012
4. Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A., & Steinberg, S. (1998). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic acids research*, 26(1), 148–53.
5. Sprinzl, M., & Gauss, D. H. (1982). Compilation of sequences of tRNA genes. *Nucleic acids research*, 10(2), r57–81.
6. Rich, A., & RajBhandary, U. L. (1976). Transfer RNA: molecular structure, sequence, and properties. *Annual review of biochemistry*, 45, 805–60. doi:10.1146/annurev.bi.45.070176.004105
7. Suddath, F. L., Quigley, G. J., McPherson, A., Sneden, D., Kim, J. J., Kim, S. H., & Rich, A. (1974). Three-dimensional structure of yeast phenylalanine transfer RNA at 3.0 angstroms resolution. *Nature*, 248(5443), 20–4.
8. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., ... Rich, A. (1974). Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science (New York, N.Y.)*, 185(4149), 435–40.
9. Kim, S. H., Quigley, G., Suddath, F. L., McPherson, A., Sneden, D., Kim, J. J., ... Rich, A. (1972). The three-dimensional structure of yeast phenylalanine transfer RNA: shape of the molecule at 5.5-Å resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 69(12), 3746–50.

10. Kim, S. H., Quigley, G. J., Suddath, F. L., McPherson, A., Sneden, D., Kim, J. J., ... Rich, A. (1973). Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain. *Science (New York, N.Y.)*, *179*(4070), 285–8.
11. Oliva, R., Cavallo, L., & Tramontano, A. (2006). Accurate energies of hydrogen bonded nucleic acid base pairs and triplets in tRNA tertiary interactions. *Nucleic acids research*, *34*(3), 865–79. doi:10.1093/nar/gkj491
12. Palioura, S., Sherrer, R. L., Steitz, T. A., Söll, D., & Simonovic, M. (2009). The human SepSecS-tRNA^{Sec} complex reveals the mechanism of selenocysteine formation. *Science (New York, N.Y.)*, *325*(5938), 321–5. doi:10.1126/science.1173755
13. Shi, H., & Moore, P. B. (2000). The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA (New York, N.Y.)*, *6*(8), 1091–105.
14. Giegé, R. (2008). Toward a more complete view of tRNA biology. *Nature structural & molecular biology*. doi:10.1038/nsmb.1498
15. Barraud, P., Schmitt, E., Mechulam, Y., Dardel, F., & Tisné, C. (2008). A unique conformation of the anticodon stem-loop is associated with the capacity of tRNA^{fMet} to initiate protein synthesis. *Nucleic acids research*, *36*(15), 4894–901. doi:10.1093/nar/gkn462
16. Suzuki, T., Nagao, A., & Suzuki, T. (2011). Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. *Annual review of genetics*, *45*, 299–329. doi:10.1146/annurev-genet-110410-132531
17. Villanueva, M. T. (2015). Genetics: The importance of mitochondrial tRNA. *Nature reviews. Cancer*, *15*(8), 456. doi:10.1038/nrc3997
18. Hanada, T., Suzuki, T., Yokogawa, T., Takemoto-Hori, C., Sprinzl, M., & Watanabe, K. (2001). Translation ability of mitochondrial tRNA^{Ser} with unusual secondary structures in an in vitro translation system of bovine mitochondria. *Genes to cells : devoted to molecular & cellular mechanisms*, *6*(12), 1019–30.
19. Ishitani, R., Nureki, O., Nameki, N., Okada, N., & Nishimura, S. (2003). Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme, *113*(May 2), 383–394.
20. Selmer, M., Dunham, C. M., Murphy, F. V., Weixlbaumer, A., Petry, S., Kelley, A. C., ... Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. *Science (New York, N.Y.)*, *313*(5795), 1935–42. doi:10.1126/science.1131127
21. Marck, C., & Grosjean, H. (2002). tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies

and domain-specific features. *RNA (New York, N.Y.)*. doi:10+1017/S1355838202022021

22. Grosjean, H., Crécy-Lagard, V., & Marck, C. (2010). Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS letters*, *584*(2), 252–64. doi:10.1016/j.febslet.2009.11.052

23. Crick, F. H. (1966). Codon--anticodon pairing: the wobble hypothesis. *Journal of molecular biology*, *19*(2), 548–55.

24. Agris, P. F., Vendeix, F. A., & Graham, W. D. (2007). tRNA's wobble decoding of the genome: 40 years of modification. *Journal of molecular biology*, *366*(1), 1–13. doi:10.1016/j.jmb.2006.11.046

25. Murphy, F. V., & Ramakrishnan, V. (2004). Structure of a purine-purine wobble base pair in the decoding center of the ribosome. *Nature structural & molecular biology*, *11*(12), 1251–2. doi:10.1038/nsmb866

26. Plotkin, J.B., & Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*. doi:10.1038/nrg2899

27. Andersson, S. G., & Kurland, C. G. (1990). Codon preferences in free-living microorganisms. *Microbiological reviews*, *54*(2), 198–210.

28. Sharp, P. M., Tuohy, T. M., & Mosurski, K. R. (1986). Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic acids research*, *14*(13), 5125–43.

29. Bulmer, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics*, *129*(3), 897–907.

30. Angov, E. (2011). Codon usage: nature's roadmap to expression and folding of proteins. *Biotechnology journal*, *6*(6), 650–659. doi:10.1002/biot.201000332

31. Ban, N., Nissen, P., Hansen, J., Moore, P. B., & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science (New York, N.Y.)*, *289*(5481), 905–20.

32. Muth, G. W., Chen, L., Kosek, A. B., & Strobel, S. A. (2001). pH-dependent conformational flexibility within the ribosomal peptidyl transferase center. *RNA (New York, N.Y.)*, *7*(10), 1403–15.

33. Thompson, J., Kim, D. F., O'Connor, M., Lieberman, K. R., Bayfield, M. A., Gregory, S. T., ... Dahlberg, A. E. (2001). Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 9002–7. doi:10.1073/pnas.151257098

34. Polacek, N., Gaynor, M., Yassin, A., & Mankin, A. S. (2001). Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature*, *411*(6836), 498–501. doi:10.1038/35078113
35. Weinger, JS, Parnell, KM, Dorner, S, & Green, R. (2004). Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nature structural & ...* doi:10.1038/nsmb841
36. Beringer, M, & Rodnina, MV. (2007). The ribosomal peptidyl transferase. *Molecular Cell*. doi:10.1016/j.molcel.2007.03.015
37. Francklyn, CS, & Minajigi, A. (2010). tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Letters*. doi:10.1016/j.febslet.2009.11.045
38. Von der Haar, F., & Cramer, F. (1976). Hydrolytic action of aminoacyl-tRNA synthetases from baker's yeast: "chemical proofreading" preventing acylation of tRNA(I1e) with misactivated valine. *Biochemistry*, *15*(18), 4131–8.
39. Nordin, B. E., & Schimmel, P. (2002). Plasticity of recognition of the 3'-end of mischarged tRNA by class I aminoacyl-tRNA synthetases. *The Journal of biological chemistry*, *277*(23), 20510–7. doi:10.1074/jbc.M202023200
40. Minajigi, A, & Francklyn, CS. (2008). RNA-assisted catalysis in a protein enzyme: The 2'-hydroxyl of tRNAThr A76 promotes aminoacylation by threonyl-tRNA synthetase. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.0804247105
41. Sankaranarayanan, R., Dock-Bregeon, A. C., Romby, P., Caillet, J., Springer, M., Rees, B., ... Moras, D. (1999). The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell*, *97*(3), 371–81.
42. Barreteau, H., Kovac, A., Boniface, A., Sova, M., Gobec, S., & Blanot, D. (2008). Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS microbiology reviews*, *32*(2), 168–207. doi:10.1111/j.1574-6976.2008.00104.x
43. Niyomporn, B., Dahl, J. L., & Strominger, J. L. (1968). Biosynthesis of the peptidoglycan of bacterial cell walls. IX. Purification and properties of glycyl transfer ribonucleic acid synthetase from *Staphylococcus aureus*. *The Journal of biological chemistry*, *243*(4), 773–8.
44. Matsushashi, M., Dietrich, C. P., & Strominger, J. L. (1965). Incorporation of glycine into the cell wall glycopeptide in *Staphylococcus aureus*: role of sRNA and lipid intermediates. *Proceedings of the National Academy of Sciences of the United States of America*, *54*(2), 587–94.
45. Bumsted, R. M., Dahl, J. L., Söll, D., & Strominger, J. L. (1968). Biosynthesis of the peptidoglycan of bacterial cell walls. X. Further study of the glycyl transfer

ribonucleic acids active in peptidoglycan synthesis in *Staphylococcus aureus*. *The Journal of biological chemistry*, 243(4), 779–82.

46. Vetting, M. W., S de Carvalho, L. P., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., & Blanchard, J. S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. *Archives of biochemistry and biophysics*, 433(1), 212–26. doi:10.1016/j.abb.2004.09.003

47. Fonvielle, M., Li de La Sierra-Gallay, I., El-Sagheer, A. H., Lecerf, M., Patin, D., Mellal, D., ... Arthur, M. (2013). The structure of FemX(Wv) in complex with a peptidyl-RNA conjugate: mechanism of aminoacyl transfer from Ala-tRNA(Ala) to peptidoglycan precursors. *Angewandte Chemie (International ed. in English)*, 52(28), 7278–81. doi:10.1002/anie.201301411

48. Varshavsky, A. (1997). The N-end rule pathway of protein degradation, (2), 13–28.

49. Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., & Tomita, K. (2006). Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *The EMBO journal*, 25(24), 5942–50. doi:10.1038/sj.emboj.7601433

50. Mogk, A., Schmidt, R., & Bukau, B. (2007). The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends in Cell Biology*, 17(4). doi:10.1016/j.tcb.2007.02.001

51. Kraus, D., & Peschel, A. (2006). Molecular mechanisms of bacterial resistance to antimicrobial peptides. *Antimicrobial Peptides and Human Disease*. doi:10.1007/3-540-29916-5_9

52. Garg, R. P., Qian, X. L., Alemany, L. B., Moran, S., & Parry, R. J. (2008). Investigations of valanimycin biosynthesis: elucidation of the role of seryl-tRNA. *Proceedings of the National Academy of Sciences of the United States of America*, 105(18), 6543–7. doi:10.1073/pnas.0708957105

53. Garg, R. P., Ma, Y., Hoyt, J. C., & Parry, R. J. (2002). Molecular characterization and analysis of the biosynthetic gene cluster for the azoxy antibiotic valanimycin. *Molecular microbiology*, 46(2), 505–17.

54. Garg, R. P., Gonzalez, J. M., & Parry, R. J. (2006). Biochemical characterization of VlmL, a Seryl-tRNA synthetase encoded by the valanimycin biosynthetic gene cluster. *The Journal of biological chemistry*, 281(37), 26785–91. doi:10.1074/jbc.M603675200

55. Gondry, M., Sauguet, L., Belin, P., & Thai, R. (2009). Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. *Nature Chemical Biology*. doi:10.1038/nchembio.175

56. Schön, A., Krupp, G., Gough, S., Berry-Lowe, S., Kannangara, C. G., & Söll, D. (1986). The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. *Nature*, *322*(6076), 281–4. doi:10.1038/322281a0
57. Huang, D. D., & Wang, W. Y. (1986). Chlorophyll biosynthesis in *Chlamydomonas* starts with the formation of glutamyl-tRNA. *The Journal of biological chemistry*, *261*(29), 13451–5.
58. Sleiman, D., Goldschmidt, V., Barraud, P., Marquet, R., Paillart, J.-C. C., & Tisné, C. (2012). Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus research*, *169*(2), 324–39. doi:10.1016/j.virusres.2012.06.006
59. Burnett, B. P., & McHenry, C. S. (1997). Posttranscriptional modification of retroviral primers is required for late stages of DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(14), 7210–5.
60. Auxilien, S., Keith, G., Le Grice, S. F., & Darlix, J. L. (1999). Role of post-transcriptional modifications of primer tRNA^{Lys,3} in the fidelity and efficacy of plus strand DNA transfer during HIV-1 reverse transcription. *The Journal of biological chemistry*, *274*(7), 4412–20.
61. Renda, M. J., Rosenblatt, J. D., Klimatcheva, E., Demeter, L. M., Bambara, R. A., & Planelles, V. (2001). Mutation of the methylated tRNA(Lys)(3) residue A58 disrupts reverse transcription and inhibits replication of human immunodeficiency virus type 1. *Journal of virology*, *75*(20), 9671–8. doi:10.1128/JVI.75.20.9671-9678.2001
62. Ohno, S. (1972). So much “junk” DNA in our genome. *Brookhaven symposia in biology*, *23*, 366–70.
63. Amaral, P. P., Dinger, M. E., Mercer, T. R., & Mattick, J. S. (2008). The eukaryotic genome as an RNA machine. *Science (New York, N.Y.)*, *319*(5871), 1787–9. doi:10.1126/science.1155472
64. Mattick, J. S. (2004). RNA regulation: a new genetics? *Nature reviews. Genetics*, *5*(4), 316–23. doi:10.1038/nrg1321
65. Hüttenhofer, A., Schattner, P., & Polacek, N. (2005). Non-coding RNAs: hope or hype? *Trends in genetics : TIG*, *21*(5), 289–97. doi:10.1016/j.tig.2005.03.007
66. Levitz, R., Chapman, D., Amitsur, M., Green, R., Snyder, L., & Kaufmann, G. (1990). The optional *E. coli* prr locus encodes a latent form of phage T4-induced anticodon nuclease. *The EMBO journal*, *9*(5), 1383–9.
67. Gebetsberger, J., & Polacek, N. (2013). Slicing tRNAs to boost functional ncRNA diversity. *RNA biology*, *10*(12), 1798–806. doi:10.4161/rna.27177

68. Emara, M. M., Ivanov, P., Hickman, T., Dawra, N., Tisdale, S., Kedersha, N., ... Anderson, P. (2010). Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *The Journal of biological chemistry*, *285*(14), 10959–68. doi:10.1074/jbc.M109.077560
69. Yamasaki, S., Ivanov, P., Hu, G.-F. F., & Anderson, P. (2009). Angiogenin cleaves tRNA and promotes stress-induced translational repression. *The Journal of cell biology*, *185*(1), 35–42. doi:10.1083/jcb.200811106
70. Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985). Isolation and characterization of angiogenin, an angiogenic protein from human carcinoma cells. *Biochemistry*, *24*(20), 5480–6.
71. Zhao, H., Bojanowski, K., Ingber, D. E., Panigrahy, D., Pepper, M. S., Montesano, R., & Shing, Y. (1999). New role for tRNA and its fragment purified from human urinary bladder carcinoma conditioned medium: inhibition of endothelial cell growth. *Journal of cellular biochemistry*, *76*(1), 109–17.
72. Durdevic, Z., Mobin, M. B., Hanna, K., Lyko, F., & Schaefer, M. (2013). The RNA methyltransferase Dnmt2 is required for efficient Dicer-2-dependent siRNA pathway activity in *Drosophila*. *Cell reports*, *4*(5), 931–7. doi:10.1016/j.celrep.2013.07.046
73. Honda, S., Loher, P., Shigematsu, M., Palazzo, J. P., Suzuki, R., Imoto, I., ... Kirino, Y. (2015). Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), E3816–25. doi:10.1073/pnas.1510077112
74. Raina, M., & Ibba, M. (2014). tRNAs as regulators of biological processes. *Frontiers in genetics*, *5*. doi:10.3389/fgene.2014.00171
75. Durdevic, Z., & Schaefer, M. (2013). tRNA modifications: necessary for correct tRNA-derived fragments during the recovery from stress? *BioEssays: news and reviews in molecular, cellular and developmental biology*, *35*(4), 323–7. doi:10.1002/bies.201200158
76. Gebetsberger, J., Zywicki, M., Künzi, A., & Polacek, N. (2012). tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Haloferax volcanii*. *Archaea (Vancouver, B.C.)*, *2012*, 260909. doi:10.1155/2012/260909
77. Sobala, A., & Hutvagner, G. (2013). Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. *RNA biology*, *10*(4), 553–63. doi:10.4161/rna.24285
78. Yeung, M. L., Bennasser, Y., Watashi, K., Le, S.-Y. Y., Houzet, L., & Jeang, K.-T. T. (2009). Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence

for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic acids research*, 37(19), 6575–86. doi:10.1093/nar/gkp707

79. Maute, R. L., Schneider, C., Sumazin, P., Holmes, A., Califano, A., Basso, K., & Dalla-Favera, R. (2013). tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*, 110(4), 1404–9. doi:10.1073/pnas.1206761110

80. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096), 816–21. doi:10.1126/science.1225829

81. Lee, Y. S., Shibata, Y., Malhotra, A., & Dutta, A. (2009). A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes & development*, 23(22), 2639–49. doi:10.1101/gad.1837609

82. Anderson, P., & Ivanov, P. (2014). tRNA fragments in human health and disease. *FEBS letters*, 588(23), 4297–304. doi:10.1016/j.febslet.2014.09.001

83. Greenway, M. J., Andersen, P. M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., ... Hardiman, O. (2006). ANG mutations segregate with familial and “sporadic” amyotrophic lateral sclerosis. *Nature genetics*, 38(4), 411–3. doi:10.1038/ng1742

84. Es, M. A., Schelhaas, H. J., Vught, P. W., Ticozzi, N., Andersen, P. M., Groen, E. J., ... Diekstra, F. P. (2011). Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. *Annals of neurology*, 70(6), 964–973. doi:10.1002/ana.22611

85. Wu, D., Yu, W., Kishikawa, H., & Folkerth, R. D. (2007). Angiogenin loss-of-function mutations in amyotrophic lateral sclerosis. *Annals of neurology*. doi:10.1002/ana.21221

86. Kieran, D., Sebastia, J., Greenway, M. J., King, M. A., Connaughton, D., Concannon, C. G., ... Prehn, J. H. (2008). Control of motoneuron survival by angiogenin. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 28(52), 14056–61. doi:10.1523/JNEUROSCI.3399-08.2008

87. Sebastià, J., Kieran, D., Breen, B., King, M. A., Netteland, D. F., Joyce, D., ... Prehn, J. H. (2009). Angiogenin protects motoneurons against hypoxic injury. *Cell death and differentiation*, 16(9), 1238–47. doi:10.1038/cdd.2009.52

88. Saikia, M., Jobava, R., Parisien, M., Putnam, A., Krokowski, D., Gao, X.-H. H., ... Hatzoglou, M. (2014). Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. *Molecular and cellular biology*, 34(13), 2450–63. doi:10.1128/MCB.00136-14

89. Brzezicha, B., Schmidt, M., Makalowska, I., Jarmolowski, A., Pienkowska, J., & Szweykowska-Kulinska, Z. (2006). Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic acids research*, *34*(20), 6034–43. doi:10.1093/nar/gkl765
90. Hussain, S., Sajini, A. A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., ... Frye, M. (2013). NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell reports*, *4*(2), 255–61. doi:10.1016/j.celrep.2013.06.029
91. Khoddami, V., & Cairns, B. R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nature biotechnology*, *31*(5), 458–64. doi:10.1038/nbt.2566
92. Khan, M. A., Rafiq, M. A., Noor, A., Hussain, S., Flores, J. V., Rupp, V., ... Vincent, J. B. (2012). Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. *American journal of human genetics*, *90*(5), 856–63. doi:10.1016/j.ajhg.2012.03.023
93. Abbasi-Moheb, L., Mertel, S., Gonsior, M., Nouri-Vahid, L., Kahrizi, K., Cirak, S., ... Kuss, A. W. (2012). Mutations in NSUN2 cause autosomal-recessive intellectual disability. *American journal of human genetics*, *90*(5), 847–55. doi:10.1016/j.ajhg.2012.03.021
94. Martinez, F. J., Lee, J. H., Lee, J. E., Blanco, S., Nickerson, E., Gabriel, S., ... Gleeson, J. G. (2012). Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *Journal of medical genetics*, *49*(6), 380–5. doi:10.1136/jmedgenet-2011-100686
95. Blanco, S., Dietmann, S., Flores, J. V., Hussain, S., Kutter, C., Humphreys, P., ... Frye, M. (2014). Aberrant methylation of tRNAs links cellular stress to neurodevelopmental disorders. *The EMBO journal*, *33*(18), 2020–39. doi:10.15252/embj.201489282
96. Nakanishi, K., & Nureki, O. (2005). Recent progress of structural biology of tRNA processing and modification. *Molecules and cells*, *19*(2), 157–166.
97. Motorin, Y., & Helm, M. (2010). tRNA stabilization by modified nucleotides. *Biochemistry*, *49*(24), 4934–44. doi:10.1021/bi100408z
98. Li, H. (2007). Complexes of tRNA and maturation enzymes: shaping up for translation. *Current Opinion in Structural Biology*, *17*, 293–301. doi:10.1016/j.sbi.2007.05.002
99. Hopper, AK, & Phizicky, EM. (2003). tRNA transfers to the limelight. *Genes & development*, *17*, 162–180. doi: 10.1101/gad.1049103

100. Torres-Larios, A, Swinger, KK, & Pan, T. (2006). Structure of ribonuclease P—a universal ribozyme. *Current Opinion in Structural Biology*, (16), 327–335. doi:10.1016/j.sbi.2006.04.002
101. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35(3), 849–857.
102. Chamberlain, JR, Lee, Y, & Lane, WS. (1998). Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes & Development*, (12), 1678–1690. doi:10.1101/gad.12.11.1678
103. Kikovska, E., Svärd, S. G., & Kirsebom, L. A. (2007). Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proceedings of the National Academy of Sciences of the United States of America*, 104(7), 2062–7. doi:10.1073/pnas.0607326104
104. True, HL, & Celander, DW. (1998). Protein components contribute to active site architecture for eukaryotic ribonuclease P. *Journal of Biological Chemistry*. doi:10.1074/jbc.273.13.7193
105. Holzmann, J, Frank, P, Löffler, E, Bennett, KL, & Gerner, C. (2008). RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme, (135), 462–474. doi:10.1016/j.cell.2008.09.013
106. Li, Z., & Deutscher, M. P. (2002). RNase E plays an essential role in the maturation of Escherichia coli tRNA precursors. *RNA (New York, N.Y.)*, 8(1), 97–109.
107. Li de la Sierra-Gallay, I., Pellegrini, O., & Condon, C. (2005). Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z. *Nature*, 433(7026), 657–61. doi:10.1038/nature03284
108. Tomita, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, N., Vassylyev, D. G., & Nureki, O. (2004). Structural basis for template-independent RNA polymerization. *Nature*, 430(7000), 700–4. doi:10.1038/nature02712
109. Xiong, Y., & Steitz, T. A. (2004). Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature*, 430(7000), 640–5. doi:10.1038/nature02711
110. Li, F., Xiong, Y., Wang, J., Cho, H. D., Tomita, K., Weiner, A. M., & Steitz, T. A. (2002). Crystal structures of the Bacillus stearothermophilus CCA-adding enzyme and its complexes with ATP or CTP. *Cell*, 111(6), 815–24.
111. Reinhold-Hurek, B, & Shub, DA. (1992). Self-splicing introns in tRNA genes of widely divergent bacteria. doi:10.1038/357173a0

112. Peebles, CL, Gegenheimer, P, & Abelson, J. (1983). Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease, *32*, 525–536.
113. Culver, G. M., McCraith, S. M., Consaul, S. A., Stanford, D. R., & Phizicky, E. M. (1997). A 2'-phosphotransferase implicated in tRNA splicing is essential in *Saccharomyces cerevisiae*. *The Journal of biological chemistry*, *272*(20), 13203–10.
114. Phizicky, E. M., Schwartz, R. C., & Abelson, J. (1986). *Saccharomyces cerevisiae* tRNA ligase. Purification of the protein and isolation of the structural gene. *The Journal of biological chemistry*, *261*(6), 2978–86.
115. Ferrer, E. (2006). Aminoacyl-tRNA Synthetases. *Drug news & perspectives*, *19*(6).
116. Cusack, S. (1997). Aminoacyl-tRNA synthetases, *7*, 881–889.
117. Lund, E, & Dahlberg, JE. (1998). Proofreading and aminoacylation of tRNAs before export from the nucleus, *282*, 2082–2085. doi:10.1126/science.282.5396.2082
118. Wilcox, M, & Nirenberg, M. (1968). Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proceedings of the National Academy of Sciences of the United States of America*, *61*(1), 229.
119. Sheppard, K, Yuan, J, Hohn, MJ, & Jester, B. (2008). From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Research*, *36*(6), 1813–1825.
120. Schmitt, E, Panvert, M, Blanquet, S, & Mechulam, Y. (2005). Structural basis for tRNA-dependent amidotransferase function, *13*, 1421–1433. doi: 10.1016/j.str.2005.06.016
121. Dumelin, C. E., Chen, Y., Leconte, A. M., Chen, G. Y., & Liu, D. R. (2012). Discovery and biological characterization of geranylated RNA in bacteria. *Nature chemical biology*, *8*(11), 913–919. doi:10.1038/nchembio.1070
122. Cantara, W. A., Crain, P. F., Rozenski, J., McCloskey, J. A., Harris, K. A., Zhang, X., ... Agris, P. F. (2011). The RNA Modification Database, RNAMDB: 2011 update. *Nucleic acids research*, *39*(Database issue), D195–201. doi:10.1093/nar/gkq1028
123. Grosjean, H., & Benne, R. (1998). *Modification and Editing of Rna*. ASM Press.
124. Jackman, J., & Alfonzo, J. (2013). Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley Interdisciplinary Reviews: RNA*, *4*(1), 35–48. doi:10.1002/wrna.1144
125. Björk, GR, Jacobsson, K, & Nilsson, K. (2001). A primordial tRNA modification required for the evolution of life? *The EMBO Journal*. doi:10.1093/emboj/20.1.231

126. Ge, J., & Yu, Y.-T. T. (2013). RNA pseudouridylation: new insights into an old modification. *Trends in biochemical sciences*, 38(4), 210–8. doi:10.1016/j.tibs.2013.01.002
127. King, T. H., Liu, B., McCully, R. R., & Fournier, M. J. (2003). Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Molecular cell*, 11(2), 425–35.
128. Yang, C., McPheeters, D. S., & Yu, Y.-T. T. (2005). Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*. *The Journal of biological chemistry*, 280(8), 6655–62. doi:10.1074/jbc.M413288200
129. Lin, Y., & Kielkopf, C. L. (2008). X-ray structures of U2 snRNA-branchpoint duplexes containing conserved pseudouridines. *Biochemistry*, 47(20), 5503–14. doi:10.1021/bi7022392
130. Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., & Yi, C. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nature Chemical Biology*, 11(8), 592–597. doi:10.1038/nchembio.1836
131. Schwartz, S., Bernstein, D., Mumbach, M., Jovanovic, M., Herbst, R., León-Ricardo, B., ... Regev, A. (2014). Transcriptome-wide Mapping Reveals Widespread Dynamic-Regulated Pseudouridylation of ncRNA and mRNA. *Cell*, 159(1). doi:10.1016/j.cell.2014.08.028
132. Carlile, T., Rojas-Duran, M., Zinshteyn, B., Shin, H., Bartoli, K., & Gilbert, W. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*, 515(7525), 143–146. doi:10.1038/nature13802
133. Karijolich, J., & Yu, Y.-T. T. (2011). Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature*, 474(7351), 395–8. doi:10.1038/nature10165
134. Grosjean, H, Sprinzl, M, & Steinberg, S. (1995). Posttranscriptionally modified nucleosides in transfer RNA: their locations and frequencies, 77, 139–141.
135. Dunin-Horkawicz, S, & Czerwoniec, A. (2006). MODOMICS: a database of RNA modification pathways. *Nucleic Acids Research*, 34, D145–D149. doi:10.1093/nar/gkj084
136. Czerwoniec, A, Dunin-Horkawicz, S, & Purta, E. (2009). MODOMICS: a database of RNA modification pathways. 2008 update. *Nucleic Acids Research*, 37, D118–D121. doi:10.1093/nar/gkn710
137. Machnicka, M. A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., ... Grosjean, H. (2013). MODOMICS: a database of RNA modification

pathways--2013 update. *Nucleic acids research*, 41(Database issue), D262–7. doi:10.1093/nar/gks1007

138. Sanbonmatsu, KY. (2006). Alignment/misalignment hypothesis for tRNA selection by the ribosome, *88*, 1075–1089. doi:10.1016/j.biochi.2006.07.002

139. Ogle, JM, Brodersen, DE, Clemons, WM, & Tarry, MJ. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit, *292*, 897–902. doi:10.1126/science.1060612

140. Peng, C, Qian, Q, Zhang, S, Isaksson, LA, & Björk, GR. (2002). A cytosolic tRNA with an unmodified adenosine in the wobble position reads a codon ending with the non-complementary nucleoside cytidine, *317*, 481–492. doi:10.1006/jmbi.2001.5435

141. Lim, VI. (1995). Analysis of action of the wobble adenine on codon reading within the ribosome, *252*, 277–282.

142. Lietzke, SE, Barnes, CL, Berglund, JA, & Kundrot, CE. (1996). The structure of an RNA dodecamer shows how tandem U–U base pairs increase the range of stable RNA structures and the diversity of recognition sites, *4*, 917–930.

143. Grosjean, H, Söll, DG, & Crothers, DM. (1976). Studies of the complex between transfer RNAs with complementary anticodons: I. Origins of enhanced affinity between complementary triplets, *103*, 499–519.

144. Jenner, LB, Demeshkina, N, & Yusupova, G. (2010). Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nature Structural & Molecular Biology*. doi:10.1038/nsmb.1790

145. Kowalak, JA, Dalluge, JJ, McCloskey, JA, & Stetter, KO. (1994). The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry*. doi:10.1021/bi00191a014

146. Sampson, JR, & Uhlenbeck, OC. (1988). Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro, *85*, 1033–1037. doi:10.1073/pnas.85.4.1033

147. Derrick, WB, & Horowitz, J. (1993). Probing structural differences between native and in vitro transcribed Escherichia coli valine transfer RNA: evidence for stable base modification-dependent conformers, *21(21)*, 4948–4953. doi:10.1093/nar/21.21.4948

148. Maglott, EJ, Deo, SS, Przykorska, A, & Glick, GD. (1998). Conformational transitions of an unmodified tRNA: implications for RNA folding. *Biochemistry*. doi:10.1021/bi981722u

149. Perret, V, Garcia, A, Puglisi, J, Grosjean, H, & Ebel, JP. (1990). Conformation in solution of yeast tRNA Asp transcripts deprived of modified nucleotides, *72*, 735–744.
150. Perret, V., Garcia, A., Grosjean, H., Ebel, J. P., Florentz, C., & Giegé, R. (1990). Relaxation of a transfer RNA specificity by removal of modified nucleotides. *Nature*, *344*(6268), 787–9. doi:10.1038/344787a0
151. Nobles, K. N., Yarian, C. S., Liu, G., Guenther, R. H., & Agris, P. F. (2002). Highly conserved modified nucleosides influence Mg²⁺-dependent tRNA folding. *Nucleic acids research*, *30*(21), 4751–60.
152. Arnez, J. G., & Steitz, T. A. (1994). Crystal structure of unmodified tRNA(Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry*, *33*(24), 7560–7.
153. Hall, K. B., Sampson, J. R., Uhlenbeck, O. C., & Redfield, A. G. (1989). Structure of an unmodified tRNA molecule. *Biochemistry*, *28*(14), 5794–801.
154. Yue, D., Kintanar, A., & Horowitz, J. (1994). Nucleoside modifications stabilize Mg²⁺ binding in Escherichia coli tRNA(Val): an imino proton NMR investigation. *Biochemistry*, *33*(30), 8905–11.
155. Durant, P. C., Bajji, A. C., Sundaram, M., Kumar, R. K., & Davis, D. R. (2005). Structural effects of hypermodified nucleosides in the Escherichia coli and human tRNA^{Lys} anticodon loop: the effect of nucleosides s²U, mcm⁵U, mcm⁵s²U, mnm⁵s²U, t⁶A, and ms²t⁶A. *Biochemistry*, *44*(22), 8078–89. doi:10.1021/bi050343f
156. Helm, M., Brulé, H., Degoul, F., Capanec, C., Leroux, J. P., Giegé, R., & Florentz, C. (1998). The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic acids research*, *26*(7), 1636–43.
157. Yue, D, Kintanar, A, & Horowitz, J. (1994). Nucleoside modifications stabilize Mg²⁺ binding in Escherichia coli tRNA^{Val}: an imino proton NMR investigation. *Biochemistry*. doi:10.1021/bi00196a007
158. Helm, M. (2006). Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic acids research*. doi:10.1093/nar/gkj471
159. Garcia-Barrio, M. T., Naranda, T., Vazquez de Aldana, C. R., Cuesta, R., Hinnebusch, A. G., Hershey, J. W., & Tamame, M. (1995). GCD10, a translational repressor of GCN4, is the RNA-binding subunit of eukaryotic translation initiation factor-3. *Genes & development*, *9*(14), 1781–96.
160. Calvo, O, Cuesta, R, & Anderson, J. (1999). GCD14p, a repressor of GCN4 translation, cooperates with Gcd10p and Lhp1p in the maturation of initiator methionyl-tRNA in Saccharomyces cerevisiae, *19*(6), 4167–4181.

161. Basavappa, R., & Sigler, P. B. (1991). The 3 A crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination. *The EMBO journal*, *10*(10), 3105–11.
162. Wang, X., Jia, H., Jankowsky, E., & Anderson, J. T. (2008). Degradation of hypomodified tRNA(iMet) in vivo involves RNA-dependent ATPase activity of the DExH helicase Mtr4p. *RNA (New York, N.Y.)*, *14*(1), 107–16. doi:10.1261/rna.808608
163. Pintard, L., Lecointe, F., Bujnicki, J. M., Bonnerot, C., Grosjean, H., & Lapeyre, B. (2002). Trm7p catalyses the formation of two 2'-O-methylriboses in yeast tRNA anticodon loop. *The EMBO journal*, *21*(7), 1811–20. doi:10.1093/emboj/21.7.1811
164. Urbonavicius, J., Qian, Q., Durand, J. M., Hagervall, T. G., & Björk, G. R. (2001). Improvement of reading frame maintenance is a common function for several tRNA modifications. *The EMBO journal*, *20*(17), 4863–73. doi:10.1093/emboj/20.17.4863
165. Gutgsell, N., Englund, N., Niu, L., Kaya, Y., Lane, B. G., & Ofengand, J. (2000). Deletion of the Escherichia coli pseudouridine synthase gene truB blocks formation of pseudouridine 55 in tRNA in vivo, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. *RNA (New York, N.Y.)*, *6*(12), 1870–81.
166. Hopper, A. K., Furukawa, A. H., Pham, H. D., & Martin, N. C. (1982). Defects in modification of cytoplasmic and mitochondrial transfer RNAs are caused by single nuclear mutations. *Cell*, *28*(3), 543–50.
167. Nordlund, M. E., Johansson, J. O., von Pawel-Rammingen, U., & Byström, A. S. (2000). Identification of the TRM2 gene encoding the tRNA(m5U54)methyltransferase of Saccharomyces cerevisiae. *RNA (New York, N.Y.)*, *6*(6), 844–60.
168. Johansson, M. J., & Byström, A. S. (2002). Dual function of the tRNA(m(5)U54)methyltransferase in tRNA maturation. *RNA (New York, N.Y.)*, *8*(3), 324–35.
169. Watanabe, K., Shinma, M., Oshima, T., & Nishimura, S. (1976). Heat-induced stability of tRNA from an extreme thermophile, Thermus thermophilus. *Biochemical and biophysical research communications*, *72*(3), 1137–44.
170. Horie, N., Hara-Yokoyama, M., Yokoyama, S., Watanabe, K., Kuchino, Y., Nishimura, S., & Miyazawa, T. (1985). Two tRNA^{Ala1} species from an extreme thermophile, Thermus thermophilus HB8: effect of 2-thiolation of ribothymidine on the thermostability of tRNA. *Biochemistry*, *24*(21), 5711–5.
171. Yokoyama, S., Watanabe, K., & Miyazawa, T. (1987). Dynamic structures and functions of transfer ribonucleic acids from extreme thermophiles. *Advances in biophysics*, *23*, 115–47.

172. Hori, H. (2014). Methylated nucleosides in tRNA and tRNA methyltransferases. *Frontiers in genetics*, 5, 144. doi:10.3389/fgene.2014.00144
173. Shigi, N., Suzuki, T., Tamakoshi, M., Oshima, T., & Watanabe, K. (2002). Conserved bases in the TΨi C loop of tRNA are determinants for thermophile-specific 2-thiouridylation at position 54. *The Journal of biological chemistry*, 277(42), 39128–35. doi:10.1074/jbc.M207323200
174. Tomikawa, C., Yokogawa, T., Kanai, T., & Hori, H. (2010). N7-Methylguanine at position 46 (m7G46) in tRNA from *Thermus thermophilus* is required for cell viability at high temperatures through a tRNA modification network. *Nucleic acids research*, 38(3), 942–57. doi:10.1093/nar/gkp1059
175. Ishida, K., Kunibayashi, T., Tomikawa, C., Ochi, A., Kanai, T., Hirata, A., ... Hori, H. (2011). Pseudouridine at position 55 in tRNA controls the contents of other modified nucleotides for low-temperature adaptation in the extreme-thermophilic eubacterium *Thermus thermophilus*. *Nucleic acids research*, 39(6), 2304–18. doi:10.1093/nar/gkq1180
176. Yamagami, R., Yamashita, K., Nishimasu, H., Tomikawa, C., Ochi, A., Iwashita, C., ... Hori, H. (2012). The tRNA recognition mechanism of folate/FAD-dependent tRNA methyltransferase (TrmFO). *The Journal of biological chemistry*, 287(51), 42480–94. doi:10.1074/jbc.M112.390112
177. Kalhor, H. R., & Clarke, S. (2003). Novel methyltransferase for modified uridine residues at the wobble position of tRNA. *Molecular and cellular biology*, 23(24), 9283–92.
178. Begley, T. J., Rosenbach, A. S., Ideker, T., & Samson, L. D. (2002). Damage recovery pathways in *Saccharomyces cerevisiae* revealed by genomic phenotyping and interactome mapping. *Molecular cancer research : MCR*, 1(2), 103–12.
179. Bennett, C. B., Lewis, L. K., Karthikeyan, G., Lobachev, K. S., Jin, Y. H., Sterling, J. F., ... Resnick, M. A. (2001). Genes required for ionizing radiation resistance in yeast. *Nature genetics*, 29(4), 426–34. doi:10.1038/ng778
180. Ilyas, M., Straub, J., Tomlinson, I. P., & Bodmer, W. F. (1999). Genetic pathways in colorectal and other cancers. *European journal of cancer (Oxford, England : 1990)*, 35(3), 335–51.
181. Flanagan, J. M., Healey, S., Young, J., Whitehall, V., Trott, D. A., Newbold, R. F., & Chenevix-Trench, G. (2004). Mapping of a candidate colorectal cancer tumor-suppressor gene to a 900-kilobase region on the short arm of chromosome 8. *Genes, chromosomes & cancer*, 40(3), 247–60. doi:10.1002/gcc.20039
182. Begley, U., Dyavaiah, M., Patil, A., Rooney, J. P., DiRenzo, D., Young, C. M., ... Begley, T. J. (2007). Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Molecular cell*, 28(5), 860–70. doi:10.1016/j.molcel.2007.09.021

183. Weissenbach, J., & Dirheimer, G. (1978). Pairing properties of the methylester of 5-carboxymethyl uridine in the wobble position of yeast tRNA^{3Arg}. *Biochimica et biophysica acta*, 518(3), 530–4.
184. Chan, C., Dyavaiah, M., DeMott, M., Taghizadeh, K., Dedon, P., & Begley, T. (2010). A Quantitative Systems Approach Reveals Dynamic Control of tRNA Modifications during Cellular Stress. *PLoS Genetics*, 6(12). doi:10.1371/journal.pgen.1001247
185. Chan, C. T., Pang, Y., Deng, W., Babu, R. I., Dyavaiah, M., Begley, T. J., & Dedon, P. C. (2012). Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nature communications*, 3, 937. doi:10.1038/ncomms1938
186. Torres, AG, Batlle, E, & de Poupiana, LR. (2014). Role of tRNA modifications in human diseases, 20(6), 306–314.
187. Novoa, EM, & de Poupiana, LR. (2012). Speeding with control: codon usage, tRNAs, and ribosomes, 28(11), 574–581. doi:10.1016/j.tig.2012.07.006
188. Guy, M. P., & Phizicky, E. M. (2015). Conservation of an intricate circuit for crucial modifications of the tRNA^{Phe} anticodon loop in eukaryotes. *RNA (New York, N.Y.)*, 21(1), 61–74. doi:10.1261/rna.047639.114
189. Freude, K., Hoffmann, K., Jensen, L.-R. R., Delatycki, M. B., des Portes, V., Moser, B., ... Ropers, H.-H. H. (2004). Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. *American journal of human genetics*, 75(2), 305–9. doi:10.1086/422507
190. Takano, K., Nakagawa, E., Inoue, K., Kamada, F., Kure, S., & Goto, Y. (2008). A loss-of-function mutation in the FTSJ1 gene causes nonsyndromic X-linked mental retardation in a Japanese family. *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 147B(4), 479–84. doi:10.1002/ajmg.b.30638
191. Ramser, J., Winnepenninckx, B., Lenski, C., Errijgers, V., Platzer, M., Schwartz, C. E., ... Kooy, R. F. (2004). A splice site mutation in the methyltransferase gene FTSJ1 in Xp11.23 is associated with non-syndromic mental retardation in a large Belgian family (MRX9). *Journal of medical genetics*, 41(9), 679–83. doi:10.1136/jmg.2004.019000
192. Slangen Haupt, S. A., Blumenfeld, A., Gill, S. P., Leyne, M., Mull, J., Cuajungco, M. P., ... Gusella, J. F. (2001). Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. *American journal of human genetics*, 68(3), 598–605.

193. Anderson, S. L., Coli, R., Daly, I. W., Kichula, E. A., Rork, M. J., Volpi, S. A., ... Rubin, B. Y. (2001). Familial dysautonomia is caused by mutations of the IKAP gene. *American journal of human genetics*, *68*(3), 753–8. doi:10.1086/318808
194. Karlsborn, T., Tükenmez, H., Mahmud, A. K., Xu, F., Xu, H., & Byström, A. S. (2014). Elongator, a conserved complex required for wobble uridine modifications in eukaryotes. *RNA biology*, *11*(12), 1519–28. doi:10.4161/15476286.2014.992276
195. Nguyen, L., Humbert, S., Saudou, F., & Chariot, A. (2010). Elongator - an emerging role in neurological disorders. *Trends in molecular medicine*, *16*(1), 1–6. doi:10.1016/j.molmed.2009.11.002
196. Strug, L. J., Clarke, T., Chiang, T., Chien, M., Baskurt, Z., Li, W., ... Pal, D. K. (2009). Centrotemporal sharp wave EEG trait in rolandic epilepsy maps to Elongator Protein Complex 4 (ELP4). *European journal of human genetics : EJHG*, *17*(9), 1171–81. doi:10.1038/ejhg.2008.267
197. Takeoka, S., Unoki, M., Onouchi, Y., Doi, S., Fujiwara, H., Miyatake, A., ... Tamari, M. (2001). Amino-acid substitutions in the IKAP gene product significantly increase risk for bronchial asthma in children. *Journal of human genetics*, *46*(2), 57–63. doi:10.1007/s100380170109
198. Frye, M., & Watt, F. M. (2006). The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Current biology : CB*, *16*(10), 971–81. doi:10.1016/j.cub.2006.04.027
199. Rodriguez, V., Chen, Y., Elkahoun, A., Dutra, A., Pak, E., & Chandrasekharappa, S. (2007). Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of TRMT12 in breast cancer. *Genes, chromosomes & cancer*, *46*(7), 694–707. doi:10.1002/gcc.20454
200. Shimada, K., Nakamura, M., Anai, S., De Velasco, M., Tanaka, M., Tsujikawa, K., ... Konishi, N. (2009). A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. *Cancer research*, *69*(7), 3157–64. doi:10.1158/0008-5472.CAN-08-3530
201. Schaefer, M., Hagemann, S., Hanna, K., & Lyko, F. (2009). Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. *Cancer research*, *69*(20), 8127–32. doi:10.1158/0008-5472.CAN-09-0458
202. Frye, M., Dragoni, I., Chin, S.-F. F., Spiteri, I., Kurowski, A., Provenzano, E., ... Watt, F. M. (2010). Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer letters*, *289*(1), 71–80. doi:10.1016/j.canlet.2009.08.004
203. Pathak, C., Jaiswal, Y. K., & Vinayak, M. (2005). Hypomodification of transfer RNA in cancer with respect to queuosine. *RNA biology*, *2*(4), 143–8.

204. Towns, W. L., & Begley, T. J. (2012). Transfer RNA methyltransferases and their corresponding modifications in budding yeast and humans: activities, predications, and potential roles in human health. *DNA and cell biology*, 31(4), 434–54. doi:10.1089/dna.2011.1437
205. Bartlett, J. M., Thomas, J., Ross, D. T., Seitz, R. S., Ring, B. Z., Beck, R. A., ... Chetty, U. (2010). Mammostrat as a tool to stratify breast cancer patients at risk of recurrence during endocrine therapy. *Breast cancer research : BCR*, 12(4), R47. doi:10.1186/bcr2604
206. Saxena, R., Voight, B. F., Lyssenko, V., Burt, N. P. P., de Bakker, P. I., Chen, H., ... Purcell, S. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science (New York, N.Y.)*, 316(5829), 1331–6. doi:10.1126/science.1142358
207. Steinhorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G. B., ... Stefansson, K. (2007). A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nature genetics*, 39(6), 770–5. doi:10.1038/ng2043
208. Stancáková, A., Pihlajamäki, J., Kuusisto, J., Stefan, N., Fritsche, A., Häring, H., ... Laakso, M. (2008). Single-nucleotide polymorphism rs7754840 of CDKAL1 is associated with impaired insulin secretion in nondiabetic offspring of type 2 diabetic subjects and in a large sample of men with normal glucose tolerance. *The Journal of clinical endocrinology and metabolism*, 93(5), 1924–30. doi:10.1210/jc.2007-2218
209. Zhou, B., Wei, F.-Y. Y., Kanai, N., Fujimura, A., Kaitsuka, T., & Tomizawa, K. (2014). Identification of a splicing variant that regulates type 2 diabetes risk factor CDKAL1 level by a coding-independent mechanism in human. *Human molecular genetics*, 23(17), 4639–50. doi:10.1093/hmg/ddu184
210. Wei, F.-Y. Y., & Tomizawa, K. (2011). Functional loss of Cdkal1, a novel tRNA modification enzyme, causes the development of type 2 diabetes. *Endocrine journal*, 58(10), 819–25.
211. Yasukawa, T., Suzuki, T., Ueda, T., Ohta, S., & Watanabe, K. (2000). Modification defect at anticodon wobble nucleotide of mitochondrial tRNAs(Leu)(UUR) with pathogenic mutations of mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes. *The Journal of biological chemistry*, 275(6), 4251–7.
212. Yasukawa, T., Suzuki, T., Ishii, N., Ueda, T., Ohta, S., & Watanabe, K. (2000). Defect in modification at the anticodon wobble nucleotide of mitochondrial tRNA(Lys) with the MERRF encephalomyopathy pathogenic mutation. *FEBS letters*, 467(2-3), 175–8.

213. Yasukawa, T., Hino, N., Suzuki, T., Watanabe, K., Ueda, T., & Ohta, S. (2000). A pathogenic point mutation reduces stability of mitochondrial mutant tRNA(Ile). *Nucleic acids research*, *28*(19), 3779–84.
214. Kirino, Y., Goto, Y.-I., Campos, Y., Arenas, J., & Suzuki, T. (2005). Specific correlation between the wobble modification deficiency in mutant tRNAs and the clinical features of a human mitochondrial disease. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(20), 7127–32. doi:10.1073/pnas.0500563102
215. Suzuki, T., Nagao, A., & Suzuki, T. (2011). Human mitochondrial diseases caused by lack of taurine modification in mitochondrial tRNAs. *Wiley interdisciplinary reviews. RNA*, *2*(3), 376–86. doi:10.1002/wrna.65
216. Liu, Y., & Whitman, W. B. (2008). Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Annals of the New York Academy of Sciences*, *1125*(1), 171–189. doi:10.1196/annals.1419.019
217. Cantoni, GL. (1975). Biological methylation: selected aspects, *44*, 435–451.
218. Schubert, HL, Blumenthal, RM, & Cheng, X. (2003). Many paths to methyltransfer: a chronicle of convergence. *Trends in biochemical sciences*. doi:10.1016/S0968-0004(03)00090-2
219. Beis, I., & Newsholme, E. A. (1975). The contents of adenine nucleotides, phosphagens and some glycolytic intermediates in resting muscles from vertebrates and invertebrates. *The Biochemical journal*, *152*(1), 23–32.
220. Lu, S. C., & Mato, J. M. (2012). S-adenosylmethionine in liver health, injury, and cancer. *Physiological reviews*, *92*(4), 1515–1542. doi:10.1152/physrev.00047.2011
221. Loenen, W. (2006). S-adenosylmethionine: jack of all trades and master of everything?, *34*(2), 331–333.
222. Chiang, P. K., Gordon, R. K., Tal, J., Zeng, G. C., Doctor, B. P., Pardhasaradhi, K., & McCann, P. P. (1996). S-Adenosylmethionine and methylation. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, *10*(4), 471–80.
223. Cheng, X, Kumar, S, Posfai, J, Pflugrath, JW, & Roberts, RJ. (1993). Crystal structure of the Hhal DNA methyltransferase complexed with S-adenosyl-L-methionine, *74*, 299–307.
224. Rossmann, MG, Moras, D, & Olsen, KW. (1974). Chemical and biological evolution of a nucleotide-binding protein, *250*(5463), 194–199.
225. Vidgren, J., Svensson, L. A., & Liljas, A. (1994). Crystal structure of catechol O-methyltransferase. *Nature*, *368*(6469), 354–8. doi:10.1038/368354a0

226. Skinner, M. M., Puvathingal, J. M., Walter, R. L., & Friedman, A. M. (2000). Crystal structure of protein isoaspartyl methyltransferase: a catalyst for protein repair. *Structure (London, England : 1993)*, *8*(11), 1189–201.
227. Zhang, X., Zhou, L., & Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *The EMBO journal*, *19*(14), 3509–19. doi:10.1093/emboj/19.14.3509
228. Martin, JL, & McMillan, FM. (2002). SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold, *12*, 783–793.
229. Fujii, Galivan, & Huennekens. (1977). Activation of methionine synthase: further characterization of the flavoprotein system.
230. Koutmos, Datta, & Pattridge. (2009). Insights into the reactivation of cobalamin-dependent methionine synthase.
231. Matthews, RG. (2001). Cobalamin-dependent methyltransferases. *Accounts of chemical research*. doi:10.1021/ar0000051
232. Lexa, Saveant, & Zickler. (1977). Electrochemistry of vitamin B12. 2. Redox and acid-base equilibria in the B12a/B12r system. doi:10.1021/ja00450a061
233. Toohey, JI. (2006). Vitamin B12 and methionine synthesis: a critical review. Is nature's most beautiful cofactor misunderstood? *Biofactors*. doi:10.1002/biof.5520260105
234. Schubert, H. L., Wilson, K. S., Raux, E., Woodcock, S. C., & Warren, M. J. (1998). The X-ray structure of a cobalamin biosynthetic enzyme, cobalt-precorrin-4 methyltransferase. *Nature structural biology*, *5*(7), 585–92. doi:10.1038/846
235. Michel, G, Sauvé, V, Larocque, R, Li, Y, Matte, A, & Cygler, M. (2002). The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot, *10*, 1303–1315.
236. Kozbial, PZ, & Mushegian, AR. (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC structural biology*. doi:10.1186/1472-6807-5-19
237. Yokoyama, A, Wang, Z, & Wysocka, J. (2004). Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. doi:10.1128/MCB.24.13.5639-5649.2004
238. Xiao, B, Jing, C, Wilson, JR, Walker, PA, & Vasisht, N. (2003). Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature*. doi:10.1038/nature01378
239. Manzur, KL, Farooq, A, Zeng, L, & Plotnikova, O. (2003). A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. doi:10.1038/nsb898

240. Chuikov, S, Kurash, JK, Wilson, JR, Xiao, B, & Justin, N. (2004). Regulation of p53 activity through lysine methylation. *Nature*. doi:10.1038/nature03117
241. Polastro, E, Schneck, AG, Leonis, J, & Kim, S. (1978). Cytochrome c methylation.
242. Akiva, E., Brown, S., Almonacid, D. E., Barber, A. E., Custer, A. F., Hicks, M. A., ... Babbitt, P. C. (2014). The Structure-Function Linkage Database. *Nucleic acids research*, 42(Database issue), D521–30. doi:10.1093/nar/gkt1130
243. Bauerle, M. R., Schwalm, E. L., & Booker, S. J. (2015). Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *The Journal of biological chemistry*, 290(7), 3995–4002. doi:10.1074/jbc.R114.607044
244. Zhang, Q., van der Donk, W. A., & Liu, W. (2012). Radical-mediated enzymatic methylation: a tale of two SAMs. *Accounts of chemical research*, 45(4), 555–64. doi:10.1021/ar200202c
245. Toh, S.-M. M., Xiong, L., Bae, T., & Mankin, A. S. (2008). The methyltransferase YfgB/RlmN is responsible for modification of adenosine 2503 in 23S rRNA. *RNA (New York, N.Y.)*, 14(1), 98–106. doi:10.1261/rna.814408
246. Benítez-Páez, A., Villarroya, M., & Armengod, M.-E. E. (2012). The Escherichia coli RlmN methyltransferase is a dual-specificity enzyme that modifies both rRNA and tRNA and controls translational accuracy. *RNA (New York, N.Y.)*, 18(10), 1783–95. doi:10.1261/rna.033266.112
247. Stojković, V., & Fujimori, D. G. (2015). Radical SAM-Mediated Methylation of Ribosomal RNA. *Methods in enzymology*, 560, 355–76. doi:10.1016/bs.mie.2015.03.002
248. Fujimori, D. G. (2013). Radical SAM-mediated methylation reactions. *Current opinion in chemical biology*, 17(4), 597–604. doi:10.1016/j.cbpa.2013.05.032
249. Broom, AD, Townsend, LB, Jones, JW, & Robins, RK. (1964). Purine Nucleosides. VI. Further Methylation Studies of Naturally Occurring Purine Nucleosides*. *Biochemistry*. doi:10.1021/bi00892a005
250. Dunn, DB. (1961). The occurrence of 1-methyladenine in ribonucleic acid, 46, 197–198.
251. Kempenaers, M, Roovers, M, & Oudjama, Y. (2010). New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Research*. doi:10.1093/nar/gkq451
252. Tolkunova, E., Park, H., Xia, J., King, M. P., & Davidson, E. (2000). The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial

enzymes by means of an unusual alternative splicing of the primary transcript. *The Journal of biological chemistry*, 275(45), 35063–9. doi:10.1074/jbc.M006265200

253. Sakurai, M, Ohtsuki, T, & Watanabe, Y. (2001). Requirement of modified residue m1A9 for EF-Tu binding to nematode mitochondrial tRNA lacking the T arm. doi:10.1093/nass/1.1.237

254. Jühling, F., Mörl, M., Hartmann, R. K., Sprinzl, M., Stadler, P. F., & Pütz, J. (2009). tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic acids research*, 37(Database issue), D159–62. doi:10.1093/nar/gkn772

255. Salas, CE, & Dirheimer, G. (1979). In vitro methylation of yeast tRNA^{Asp} by rat brain cortical tRNA-(adenine-1) methyltransferase. *Nucleic acids research*. doi:10.1093/nar/6.3.1123

256. Grosjean, H., Auxilien, S., Constantinesco, F., Simon, C., Corda, Y., Becker, H. F., ... Fourrey, J. L. (1996). Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: a review. *Biochimie*, 78(6), 488–501.

257. Raettig, R, Kersten, H, & Weissenbach, J. (1977). Methylation of an adenosine in the D-loop of specific transfer RNAs from yeast by a procaryotic tRNA (adenine-1) methyltransferase. doi:10.1093/nar/4.6.1769

258. Roovers, M, Kaminska, KH, Tkaczuk, KL, Gigot, D, Droogmans, L, & Bujnicki, M. (2008). The YqfN protein of *Bacillus subtilis* is the tRNA: m1A22 methyltransferase (TrmK). *Nucleic Acids Research*, 36(10), 3252–3262. doi:10.1093/nar/gkn169

259. Thanassi, J. A., Neumann, S. L., Dougherty, T. J., Dougherty, B. A., & Pucci, M. J. (2002). Identification of 113 conserved essential genes using a high throughput gene disruption system in *Streptococcus pneumoniae*. *Nucleic acids research*, 30(14), 3152–3162. doi:10.1093/nar/gkf418

260. Macek, B, Mijakovic, I, Olsen, JV, & Gnad, F. (2007). The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Molecular & Cellular Proteomics*. doi:10.1074/mcp.M600464-MCP200

261. Nguyen, HA. (2012). Découverte d'une nouvelle famille de protéine kinases bactériennes: mécanismes de fonctionnement et rôle cellulaire de YdiB, un archétype chez *Bacillus subtilis*. Retrieved from <http://www.theses.fr/2012GRENV017>

262. Karst, J, Stroebel, D, Mangat, CS, Brown, ED, & Jault, JM. (2009). The ATPase activity of an "essential" *Bacillus subtilis* enzyme, YdiB, is required for its cellular function and is modulated by oligomerization, 155, 944–956. doi:10.1099/mic.0.021543-0

263. Eymann, C, Becher, D, Bernhardt, J, & Gronau, K. (2007). Dynamics of protein phosphorylation on Ser/Thr/Tyr in *Bacillus subtilis*. *Proteomics*, 7(19).

264. Sun, X, Ge, F, Xiao, CL, Yin, XF, & Ge, R. (2009). Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae*. *Journal of Proteome Research*. doi:10.1021/pr900612v
265. Lévine, A, Vannier, F, Absalon, C, & Kuhn, L. (2006). Analysis of the dynamic *Bacillus subtilis* Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes. doi/10.1002/pmic.200500352
266. Ravikumar, V, Shi, L, Krug, K, Derouiche, A, & Jers, C. (2014). Quantitative phosphoproteome analysis of *Bacillus subtilis* reveals novel substrates of the kinase PrkC and phosphatase PrpC, *13*(8), 1965–1978. doi:DOI 10.1074/mcp.M113.035949
267. Kobir, A, Shi, L, Boskovic, A, & Grangeasse, C. (2011). Protein phosphorylation in bacterial signal transduction, *1810*, 988–994.
268. Ta, H. M., & Kim, K. K. (2010). Crystal structure of *Streptococcus pneumoniae* Sp1610, a putative tRNA methyltransferase, in complex with S-adenosyl-L-methionine. *Protein science: a publication of the Protein Society*, *19*(3), 617–24. doi:10.1002/pro.319
269. Kersten, H, Raettig, R, & Weissenbach, J. (1978). Recognition of individual procaryotic and eucaryotic transfer-ribonucleic acids by *B. subtilis* adenine-1-methyltransferase specific for the dihydrouridine loop. doi:10.1093/nar/5.8.3033
270. Yamazaki, N, Hori, H, Ozawa, K, Nakanishi, S, Ueda, T, Kumagai, I, ... Nishikawa, K. (1992). Purification and characterization of tRNA(adenosine-1)-methyltransferase from *Thermus thermophilus* HB27. *Nucleic acids symposium series*, (27), 141–142.
271. Anderson, J., Phan, L., Cuesta, R., Carlson, B. A., Pak, M., Asano, K., ... Hinnebusch, A. G. (1998). The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes & development*, *12*(23), 3650–62.
272. Anderson, J., Phan, L., & Hinnebusch, A. G. (2000). The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(10), 5173–8. doi:10.1073/pnas.090102597
273. Ozanick, S., Krecic, A., Andersland, J., & Anderson, J. T. (2005). The bipartite structure of the tRNA m1A58 methyltransferase from *S. cerevisiae* is conserved in humans. *RNA (New York, N.Y.)*, *11*(8), 1281–90. doi:10.1261/rna.5040605
274. Bujnicki, J. M. (2001). In silico analysis of the tRNA:m1A58 methyltransferase family: homology-based fold prediction and identification of new members from Eubacteria and Archaea. *FEBS letters*, *507*(2), 123–7.

275. Droogmans, L., Roovers, M., Bujnicki, J. M., Tricot, C., Hartsch, T., Stalon, V., & Grosjean, H. (2003). Cloning and characterization of tRNA (m1A58) methyltransferase (TrmI) from *Thermus thermophilus* HB27, a protein required for cell growth at extreme temperatures. *Nucleic acids research*, *31*(8), 2148–56.
276. Roovers, M, Stalon, V, Grosjean, H, & Droogmans, L. (2004). A primordial RNA modification enzyme: the case of tRNA (m1A) methyltransferase, *32*(2), 465–476.
277. Varshney, U., Ramesh, V., Madabushi, A., Gaur, R., Subramanya, H. S., & RajBhandary, U. L. (2004). Mycobacterium tuberculosis Rv2118c codes for a single-component homotetrameric m1A58 tRNA methyltransferase. *Nucleic acids research*, *32*(3), 1018–27. doi:10.1093/nar/gkh207
278. Guy, MP, & Phizicky, EM. (2014). Two-subunit enzymes involved in eukaryotic post-transcriptional tRNA modification. *RNA biology*. doi:10.1080/15476286.2015.1008360
279. Guelorget, A., Roovers, M., Guérineau, V., Barbey, C., Li, X., & Golinelli-Pimpaneau, B. (2010). Insights into the hyperthermostability and unusual region-specificity of archaeal *Pyrococcus abyssi* tRNA m1A57/58 methyltransferase. *Nucleic acids research*, *38*(18), 6206–18. doi:10.1093/nar/gkq381
280. Kuratani, M., Yanagisawa, T., Ishii, R., Matsuno, M., Si, S.-Y. Y., Katsura, K., ... Yokoyama, S. (2014). Crystal structure of tRNA m(1)A58 methyltransferase TrmI from *Aquifex aeolicus* in complex with S-adenosyl-L-methionine. *Journal of structural and functional genomics*, *15*(3), 173–80. doi:10.1007/s10969-014-9183-0
281. Barraud, P., Golinelli-Pimpaneau, B., Atmanene, C., Sanglier, S., Van Dorsselaer, A., Droogmans, L., ... Tisné, C. (2008). Crystal structure of *Thermus thermophilus* tRNA m1A58 methyltransferase and biophysical characterization of its interaction with tRNA. *Journal of molecular biology*, *377*(2), 535–50. doi:10.1016/j.jmb.2008.01.041
282. Kowalak, J. A., Dalluge, J. J., McCloskey, J. A., & Stetter, K. O. (1994). The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry*, *33*(25), 7869–76.
283. Barraud, P. (2008). Etudes structurales de différents processus biologiques impliquant les ARN de transfert.
284. Guelorget, A., Barraud, P., Tisné, C., & Golinelli-Pimpaneau, B. (2011). Structural comparison of tRNA m(1)A58 methyltransferases revealed different molecular strategies to maintain their oligomeric architecture under extreme conditions. *BMC structural biology*, *11*, 48. doi:10.1186/1472-6807-11-48
285. Sengupta, R, Vainauskas, S, & Yarian, C. (2000). Modified constructs of the tRNA TΨC domain to probe substrate conformational requirements of m1A58 and m5U54 tRNA methyltransferases, *28*(6), 1374–1380.

286. Constantinesco, F., & Motorin, Y. (1999). Transfer RNA modification enzymes from *Pyrococcus furiosus*: detection of the enzymatic activities in vitro. *Nucleic acids research*. doi:10.1093/nar/27.5.1308
287. Takuma, H., Ushio, N., Minoji, M., Kazayama, A., Shigi, N., Hirata, A., ... Hori, H. (2015). Substrate tRNA recognition mechanism of eubacterial tRNA (m1A58) methyltransferase (TrmI). *The Journal of biological chemistry*, 290(9), 5912–25. doi:10.1074/jbc.M114.606038
288. Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nature reviews Molecular cell biology*, 8(6), 479–490. doi:10.1038/nrm2178
289. Huang, H., Chopra, R., Verdine, G. L., & Harrison, S. C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science (New York, N.Y.)*, 282(5394), 1669–75.
290. Huang, H., Harrison, S. C., & Verdine, G. L. (2000). Trapping of a catalytic HIV reverse transcriptase*template:primer complex through a disulfide bond. *Chemistry & biology*, 7(5), 355–64.
291. Hanwell, M. D., Curtis, D., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics*, 4(1), 17. doi:10.1186/1758-2946-4-17
292. Dominguez, C., Boelens, R., & Bonvin, A. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, 125(7), 1731–1737. doi:10.1021/ja026939x
293. Hsu, V., Wassenaar, T., & Bonvin, A. (2007). HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins: structure*. doi:10.1002/prot.21723
294. Dégut, C., Monod, A., Brachet, F., Crépin, T., & Tisné, C. (2016). In Vitro/In Vivo Production of tRNA for X-Ray Studies. *Methods in molecular biology (Clifton, N.J.)*, 1320, 37–57. doi:10.1007/978-1-4939-2763-0_4
295. Salzmann, M., Wider, G., & Pervushin, K. (1999). TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *Journal of the American Chemical Society*. doi:10.1021/ja9834226
296. Talluri, S., & Wagner, G. (1996). An Optimized 3D NOESY–HSQC. *Journal of Magnetic Resonance*, 112(2), 200–205.
297. Bracken, C. III, P. A., & Cavanagh, J. (1997). (H) N (COCA) NH and HN (COCA) NH experiments for 1H-15N backbone assignments in 13C/15N-labeled proteins. *Journal of biomolecular NMR*.

298. Panchal, SC, Bhavesh, NS, & Hosur, RV. (2001). Improved 3D triple resonance experiments, HNN and HN (C) N, for HN and ¹⁵N sequential correlations in (¹³C, ¹⁵N) labeled proteins: application to unfolded Protein. *Journal of biomolecular NMR*. doi:10.1023/A:1011239023422
299. Sun, Z., Frueh, DP, Selenko, P, & Hoch, JC. (2005). Fast assignment of ¹⁵N-HSQC peaks using high-resolution 3D HNCcαNH experiments with non-uniform sampling. *Journal of biomolecular NMR*. doi:10.1007/s10858-005-1284-4
300. Weller, R., & Rajski, S. (2006). Design, synthesis, and preliminary biological evaluation of a DNA methyltransferase-directed alkylating agent. *Chembiochem: a European journal of chemical biology*, 7(2), 243–5. doi:10.1002/cbic.200500362
301. Osborne, T., Roska, R., Rajski, S., & Thompson, P. (2008). In situ generation of a bisubstrate analogue for protein arginine methyltransferase 1. *Journal of the American Chemical Society*, 130(14), 4574–5. doi:10.1021/ja077104v
302. Fujishima, K., & Kanai, A. (2014). tRNA gene diversity in the three domains of life. *Frontiers in Genetics*, 5. doi:10.3389/fgene.2014.00142
303. Hoepfner, M. P., Gardner, P. P., & Poole, A. M. (2012). Comparative analysis of RNA families reveals distinct repertoires for each domain of life. *PLoS computational biology*, 8(11), e1002752. doi:10.1371/journal.pcbi.1002752
304. Turk, R. M., Chumachenko, N. V., & Yarus, M. (2010). Multiple translational products from a five-nucleotide ribozyme. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10), 4585–9. doi:10.1073/pnas.0912895107
305. Randau, L. (2012). RNA processing in the minimal organism Nanoarchaeum equitans. *Genome biology*, 13(7), R63. doi:10.1186/gb-2012-13-7-r63
306. Randau, L., Münch, R., Hohn, M. J., Jahn, D., & Söll, D. (2005). Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature*, 433(7025), 537–41. doi:10.1038/nature03233
307. Giulio, D. M. (2009). Formal proof that the split genes of tRNAs of Nanoarchaeum equitans are an ancestral character. *Journal of molecular evolution*. doi:10.1007/s00239-009-9280-z
308. Tamura, K., & Schimmel, P. (2004). Chiral-selective aminoacylation of an RNA minihelix. *Science (New York, N.Y.)*, 305(5688), 1253. doi:10.1126/science.1099141

