



Incorporation de la beta alanine dans des polypeptides

Giuliano Nigro

► To cite this version:

Giuliano Nigro. Incorporation de la beta alanine dans des polypeptides. Biochimie [q-bio.BM]. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLX064 . tel-02484790

HAL Id: tel-02484790

<https://theses.hal.science/tel-02484790>

Submitted on 19 Feb 2020

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.

Incorporation de β -méthionine dans des polypeptides

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Ecole polytechnique

École doctorale n°573 : Interface : approches interdisciplinaires, fon-
dements, applications et innovation
(interface)

Spécialité de doctorat: Biologie

Thèse présentée et soutenue à Palaiseau, le 18 octobre 2019, par

Giuliano Nigro

Composition du Jury :

Magali Remaud-Simeon	Rapporteur /Presidente du jury
Professeur, INSA (UMR 5504)	
Claude Sauter	Rapporteur
Directeur de recherche, Institut de Biologie Moléculaire et Cellulaire (UPR 9002)	
Dominique Fourmy	Examineur
Directeur de recherche, CNRS (URM 9198)	
Philippe Marlière	Examineur
Visiting scientist, ETH Zürich	
Emmanuelle Schmitt	Co-Directrice de thèse
Directeur de recherche, CNRS – Ecole Polytechnique (URM 7654)	
Yves Mechulam	Directeur de thèse
Directeur de recherche, CNRS – Ecole Polytechnique (URM 7654)	

Remerciements

Je souhaite tout d'abord remercier mon directeur et ma co-directrice de thèse Yves Mechulam et Emmanuelle Schmitt pour m'avoir donné l'opportunité de faire cette thèse avec eux et grâce à qui j'ai énormément appris que je ne remercierai probablement jamais assez.

J'exprime également tous mes remerciements à l'ensemble de mon jury de thèse : Magali Remaud-Simeon, Claude Sauter, Dominique Fourmy et Philippe Marlière.

Mes travaux ont été réalisés en collaboration avec Philippe Marlière et Valérie Pezo de l'ISSB, que je remercie profondément ainsi que leur équipe pour leur accueil et leur aide. Je remercie particulièrement Philippe Marlière de m'avoir orienté dans ce choix de thèse. Je tiens aussi à remercier mes autres collaborateurs Vaitea Opuu et Thomas Simonson du BIOIC, pour les échanges que nous avons eu.

Au cours de cette thèse j'ai pu travailler avec Sophie Bourcier du LCM ainsi qu'avec Alexis Archambeau et Vincent Jactel du LSO et je les remercie de l'aide qu'ils m'ont fournie.

Merci à tous les membres du BIOIC présent lors de ma thèse, en particulier les membres de B2 Titine, Michou, Gaby et Jérôme, grâce à qui j'ai pu faire cette thèse dans la bonne humeur.

Pierre Damien, et Marc G. pour leur soutien et leurs conseils notamment pour résoudre les structures.

Je souhaite aussi remercier tous les doctorants et Post-Doc du BIOIC que j'ai pu côtoyer, Clément, Karen, Nicolas, Nicolas M., Francesco, Amlan, Ditipriya, Artem, Vaitea, Irène, Can, Paula, Yanan, Gleb, Ramy et Christina pour la bonne ambiance au sein du laboratoire.

Merci aussi à Mimi pour les nombreuses longues conversations, à Clément, Marc D., aux Nathalie(s), à Stéphane, Clara ainsi qu'à David et Thomas.

Je tiens aussi à remercier les autres membres du laboratoire Guillaume, Mélanie, Chantal, Lydia et Layla sans qui tout serait tellement plus compliqué.

Sans oublier ceux qui sont partie à la retraite pendant ma thèse, Michou, Pierre Plateau et les Pascaux.

Je me dois de aussi de remercier Doc'Union et ses membres notamment Olga et Gauthier deux présidents assez patients pour me supporter et aussi Pauline qui m'a subit.

Sans oublier Ethan et Holl avec qui j'ai pu passer des soirées mémorables à oublier mes tracas.

Je remercie aussi les doctorants des autres laboratoires LOB, LSO, LADHYX avec qui j'ai pu sympathiser et avec qui nous avons pu échanger sur nos doctorats respectifs, de manière constructive, évidemment !

Un grand merci à mes anciens camarades de l'ESCOM, Max, Tom, Emilie, Flo, Jérémie, Laurine, Barth, Coco, Valou et leurs valeurs ajoutées qui m'ont permis de décompresser après certaines journées difficiles.

Merci à Camille, pour avoir relu et corrigé une partie de cette thèse sans comprendre ce qu'elle lisait.

Merci à mes cousins où qu'ils soient avec qui j'ai pu me changer les idées.

Je tiens à remercier mon frère et ma sœur qui ont partagé d'une certaine façon les plaisirs de la thèse avec moi.

Et finalement, je tiens à remercier mes parents qui ont toujours été là ces 28 dernières années et m'ont toujours poussé à faire ce que je voulais faire, même si je ne le savais pas toujours moi-même.

Préambule

La biologie de synthèse est une approche multidisciplinaire qui vise à la fois à mieux comprendre le fonctionnement des systèmes biologiques et à construire des organismes capables de répondre à des applications désirées, par exemple dans les domaines de la santé, de l'énergie, de la chimie, de l'agro alimentaire... Dans ce contexte, la conception de nouveaux catalyseurs protéiques capables de contribuer à la synthèse d'une molécule d'intérêt constitue un enjeu central. Or, les protéines naturellement fabriquées par les ribosomes sont constituées par seulement 20 acides aminés, tous de conformation L-. Dans de rares cas de catalyses particulières, les cellules vivantes peuvent également utiliser un 21^{ème} acide aminé, la sélénocystéine, utile par son potentiel rédox élevé. Encore plus rarement, certaines archées sont capables d'incorporer la pyrrolysine dans quelques chaînes polypeptidiques. Finalement, il s'agit d'une limitation sérieuse au répertoire des protéines possibles. Ainsi, dès que les études des aminoacyl-ARNt synthétases (enzymes catalysant l'amination spécifique des ARNt) sont arrivées à une maturité suffisante (fin des années 1990), l'incorporation par le ribosome d'acides aminés non naturels, ou « extension du code génétique », est devenue une préoccupation centrale de l'ingénierie biologique. Certains succès ont été remarquables, avec la création de « paires orthogonales » aminoacyl ARNt synthetase/ARN de transfert, permettant d'insérer plusieurs acides aminés au choix en réponse à un codon non-sens (Köhler *et al.* , 2004; Wang and Schultz, 2005). Les acides aminés utilisés restent toutefois relativement proches des acides aminés canoniques, en ce sens que ce sont des acides α aminés. Les protéines ainsi fabriquées adoptent donc des repliements classiques, à base d'hélices α et de feuillets β , avec une flexibilité du squelette peptidique limitée. Toutefois, l'introduction d'acides β -aminés dans une chaîne polypeptidique permettrait de changer ses propriétés de repliement et sa dynamique de manière importante, aboutissant à une plus large gamme de possibilités catalytiques. Plus généralement, l'utilisation d'acides β aminés devrait permettre de concevoir et biofabriquer des polymères complexes, bases de matériaux innovants.

Il a été montré très récemment (Katoh and Suga, 2018), grâce à un système de traduction ribosomique reconstitué *in vitro*, que le ribosome bactérien était capable d'incorporer plusieurs types d'acides β -aminés, dans des chaînes polypeptidiques. À ce stade, il n'a été possible de diriger l'incorporation d'acides β -aminés successifs que très récemment (Katoh and Suga, 2018), grâce à une modification du ribosome. Dans ces études, les ARNt aminoacylés sont préparés au préalable grâce à des systèmes d'aminacylation *in vitro* appelés flexizymes, basés sur un ribozyme, puis ajoutés au mélange réactionnel. Ainsi, l'obstacle majeur à une utilisation *in vivo* d'acides aminés β dans la synthèse ribosomique est la disponibilité d'une aminoacyl-ARNt synthétase capable de charger des acides aminés β sur un ARNt.

Table des matières

Remerciements.....	1
Préambule.....	3
Table des matières	5
Table des abréviations.....	9
Liste des figures	11
Table des tableaux	15
Publications et communications.....	16
Introduction.....	17
I Biologie de synthèse.....	17
I. 1 Généralités.....	17
II Système de traduction	19
II. 1. Généralités	19
II. 2 Rôle des partenaires.....	19
II. 2. 1 Ribosome.....	19
II. 2. 2 ARNt	20
II. 3 Aminoacyl ARNt synthétases (AARS).....	21
II. 3. 1 Généralités sur les AARS.....	21
II. 3. 1. 1 Réaction d'aminacylation.....	22
II. 3. 1. 2 Organisation des AARS	23
II. 3. 1. 2. 1 Deux classes d'AARS	23
II. 3. 1. 2. 1. 1 Classe I.....	25
II. 3. 1. 2. 1. 2 classe II	26
II. 3. 1. 2. 2 Sous classes d'AARS.....	28
II. 3. 1. 2. 2. 1 Classe I, a, b, c	28
II. 3. 1. 2. 2. 2 Classe II, a, b, c.....	30
II. 3. 1. 3 Reconnaissance de l'ARNt	31
II. 3. 1. 3. 1 identité des ARNts.....	32
II. 3. 1. 3. 2 Reconnaissance des déterminants par les AARS.....	33
II. 3. 1. 4 L'édition chez les AARS.....	34
III Méthionyl-ARNt synthétase	37
III. 1 Structure de la MetRS	37
III. 1. 1 Les différents modules de la MetRS	38
III. 1. 2 Rôle de l'ion Zinc	39

III. 1. 3	Domaine de dimérisation.....	40
III. 1. 4	Réaction d'activation de la méthionine	41
III. 2	Reconnaissance des substrats.....	42
III. 2. 1	La liaison de la méthionine.....	42
III. 2. 2	La liaison du méthionyl-adénylate.....	43
III. 2. 3	La fixation de l'ATP	44
III. 2. 4	Interaction avec l'ARNt	45
III. 3	Misacylation et édition	47
IV	Incorporation d'acides aminés non canoniques.....	49
IV. 1.	Les acides aminés non canoniques	49
IV. 2.	Incorporation <i>in vivo</i>	49
IV. 2. 1	Méthode utilisant le système de traduction sans modification.....	50
IV. 2. 2	Méthode utilisant un système de traduction modifié	50
IV. 2. 2. 1	Paires orthogonales AARS-ARNt.....	50
IV. 2. 2. 2	ARNt et ARNr orthogonaux.....	51
IV. 2. 2. 3	Les AARS orthogonales.....	53
IV. 2. 3	Modification des facteurs d'élongation et de terminaison	56
IV. 3.	Incorporation d'acides aminés non canoniques <i>in vitro</i>	59
IV. 3. 1	Système de traduction <i>in vitro</i>	59
IV. 3. 2	Méthodes de préparation des ARNt aminoacylés.....	61
V	Les acides aminés β	62
V. 1.	Intérêt des AA- β	63
V. 1. 1.	Spécificité des AA- β	63
V. 1. 2.	Applications	65
V. 2	Incorporation d'acides aminés β dans des peptides.....	67
	Objectifs.....	70
Partie I.....		71
I	Introduction	71
Partie II		89
I	Interaction de la MetRS avec des analogues de la méthionine.....	89
I. 1	Mesure des paramètres cinétiques de la MetRS vis-à-vis d'analogues de la méthionine ...	89
I. 2	Structures de la MetRS cocrystallisée avec des analogues de la méthionine	94
I. 2. 1	Structure du complexe MetRS:SeMet	95
I. 2. 2	Structure du complexe MetRS:Méthoxinine	96
I. 2. 3	Structure du complexe MetRS:Acide glutamique 5-méthyle ester	97
I. 2. 4	Structure du complexe MetRS:Desmeninol.....	99

I. 3 Conclusion.....	101
II Mutagenèse ponctuelle en vue d'améliorer l'activité de la MetRS pour la β -Met	102
II. 1 Sélection des positions.....	102
II. 2 Etude cinétique des variants de la MetRS	103
II. 3 Etude structurale des variants A12S, A12T, A12N et D52E.....	104
II. 3 Structures des variants A12T, A12N et D52E	106
II. 3. 2 Structure du variant A12S en présence de méthionine et de β -méthionine	108
II. 4 Conclusion	111
III Méthode de sélection <i>in Silico</i> en collaboration avec l'équipe de Thomas Simonson	112
III. 1 Sélection des variants.....	112
III. 1. 1 Principe.....	112
III. 1. 2 Sélection des positions à muter	112
III. 1. 3 Echantillonnage.....	113
III. 2 Liste des variants proposés par l'étude <i>in silico</i>	114
III. 2. 1 Variants spécifiques du méthionyl adénylate	114
III. 2. 2 Variants spécifiques du β -méthionyl adénylate.....	115
III. 3 Etude cinétique	116
III. 3. 1 Caractérisation en présence de méthionine.....	116
III. 3. 2 Evaluation de l'activité catalytique des variants en présence de β -Met.....	117
III. 4 Conclusion.....	120
Partie III.....	121
I Méthode d'évolution dirigée	121
I. 1 Principes	121
I. 2 Choix d'une enzyme de sélection.....	122
II Construction d'une NDT dépendante de l'incorporation d'une β -méthionine.....	123
II. 1 Choix des positions à muter	124
II. 1. 1 La N-Desoxyribosyltransferase	124
II. 1. 2 Choix des positions à muter	125
II. 2 Test des variants en présence de méthionine.....	129
III Perméabilité d' <i>E. coli</i> à la β -Met	131
III. 1 Test de toxicité	131
III. 2 Perméabilité à la β -Met via un dipeptide.....	132
III. 2. 1 Synthèse du dipeptide β -Met-His.....	132
III. 2. 2 Test de l'incorporation du dipeptide	133
IV Amélioration de l'incorporation d'acides aminés analogues de la méthionine <i>in vivo</i>	134
IV. 1 Recherche d'une mauvaise source de méthionine	134

IV. 2 Test du composés Boc-D-Met comme mauvaise source de méthionine	137
V Conclusion et perspectives.....	138
Discussion et perspectives.....	139
Matériels et Méthodes.....	142
I Matériels.....	142
I. 1 Liste des souches	142
I. 2 Liste des plasmides	142
I. 3 Liste des milieux de culture et tampons.....	143
II Méthodes.....	143
II. 1 Production et purification de protéines	143
II. 1. 1 Clonage.....	143
II. 1. 2 Mutagénèse	144
II. 2. 1 Production	144
II. 2. 2 Purification	145
II. 2 Enzymologie	145
II. 2. 1 Digestion de la β -méthionine à la γ -lyase (MGL)	145
II. 2 Caractérisation de la MetRS.....	145
II. 2. 1 Echange isotopique ATP-PPI.....	146
II. 2. 1. 1 Test d'aminoacylation	146
II. 2. 1. 2 Mesure de la vitesse de la réaction avec de la L-[^{14}C] méthionine.....	146
III. 2. 2 Aminoacylation de l'ARNt marqué au ^{32}P	146
II. 3. Etude par fluorescence	147
II. 3. 1 Fluorescence à l'équilibre.....	147
II. 3. 2. Fluorescence à l'état pré-stationnaire.....	148
II. 4 Mise en évidence de la catalyse par LC-MS et LC-MS/MS	149
II. 4. 1 Activation de l'acide aminé.....	149
II. 4. 2 Aminoacylation de l'ARNt ^{Met}	149
II. 5 Cristallisation de la MetRS et acquisition des données cristallographiques	150
II. 5. 1 Préparation des cristaux de MetRS.....	150
II. 5. 2 Principe de la cristallographie par diffraction des rayons X.....	150
II. 5. 3 Acquisition et traitement des données de diffraction.....	152
II. 5. 4 Résolution de la structure	153
III Synthèse peptidique	153
III. 1 Synthèse peptidique	153
III. 2 Test Kaiser	154
Bibliographie	155

Table des abréviations

AANC	Acide Aminé non canonique
AARS	Aminoacyl ARNt synthétase
AA-β	Acide aminé β
ADN	Acide désoxyribonucléique
Ado	Adénosine
AlaRS	Alanyl-ARNt synthetase
ARNm	Acide ribonucléique messenger
ARNr	Acide ribonucléique ribosomal
ARNt	Acide ribonucléique de transfert
AspRS	Aspartyl-ARNt synthetase
ATP	Adenosine triphosphate
Boc	Tert-Butoxycarbonyl
CCM	Chromatographie sur couche mince
CP1	Connective peptide
CTP	Cytidine triphosphate
DCM	Dichlorométhane
deoA	Thymidylate phosphorylase
Des	Desmeninol
DMF	N,N-Diméthylformamide
DMSO	Diméthylesulfoxyde
dT	Désoxythymidine
dTMP	Thymidylate
dU	Désoxyuridine
EDTA	acide éthylènediaminetétraacétique
EF-G	Facteur d'élongation G
EF-P	Facteur d'élongation P
EF-Ts	Facteur d'élongation Ts
EF-Tu	Facteur d'élongation Tu
EtOH	Ethanol
GlnRS	Glutaminyl-ARNt synthétase
GluRS	Glutamyl-ARNt synthétase
HisRS	Histidinyl-ARNt synthétase
HOAt	1-Hydroxy-7-azabenzotriazole
LC-MS	Liquide chromatographie-masse spectrometrie
LysRS	Lysyl ARNt synthetase
MAGE	Multiplex Automated Genome Engineering
Met	Méthionine
Met-ester	Acide glutamique 5-méthyle ester
Met-ether	Méthoxinine
MetRS	Méthionyl-ARNt synthétase
MGL	Méthionine Gama-lyase
NDT	Nucléoside deoxytransferase
PACE	Phage Assisted Continuous Evolution
PDT	Desoxyribosyltransferase

PEI	Polyétherimide
PheRS	Phenylalanyl-ARNt synthétase
PI	Point isoélectrique
PLP	Pyridoxal phosphate
PPi	Pyrophosphate
ProRS	Propyl-ARNt synthétase
PURE	Protein Synthesis Using Recombinant Element
PylRS	Pyrrolysyl-ARNt synthetase
RF	facteur de recyclage
RRF	facteur de recyclage du ribosome
SeMet	Sélenométhionine
SerRS	Seryl-ARNt synthetase
T	thymine
tdk	thymidylate kinase
TFA	Tétrafluoroacetate
thyA	thymidylate syntéhasse
TIS	Triisopropylsilane
Tris	2-amino-2-hydroxyméthylpropane-1,3-diol
TrpRS	Tryptophanyl-ARNt synthetase
β-Met	β-Méthionine

Liste des figures

Figure 1 : Code génétique	18
Figure 2 : Processus de traduction (Chin, 2017)	19
Figure 3 : ARNt ^{Met} en forme « feuille de trèfle » (gauche) et en forme repliée (droite).....	20
Figure 4 : Mécanisme d'aminacylation en deux étapes. La première est l'activation de l'acide aminé qui permet de former le complexe E:AA~AMP (1). La seconde est l'estérification de l'acide aminé sur l'ARNt qui permet de charger l'acide aminé sur ce dernier (2).	23
Figure 5 : Repliement de Rossmann de la structure 3D de la MetRS de <i>E. coli</i> (A) et feuillet β antiparallèle de la structure 3D de l'AspRS de <i>E. coli</i> (B)	24
Figure 6 : Positionnement de l'ATP dans la TrpRS de <i>Bacillus stearothermophilus</i> (Classe I) (5DK4) à gauche et dans la SerRS de <i>Bradyrhizobium diazoefficiens</i> (Classe II) (3MEY) à droite. Sur la structure de la TrpRS deux lysines du motif KMSKS en interaction avec l'ATP sont visibles. Dans la structure de la SerRS deux arginines (motifs 2 et 3) interagissent avec l'ATP.....	27
Figure 7 : Sites de reconnaissance des acides aminés de trois AARS de classe II. A : HisRS de <i>Burkholderia thailandensis</i> (4E51), B : PheRS de <i>T. thermophilus</i> (1B70) et C : AspRS de <i>T. thermophilus</i> (3NEL)	28
Figure 8 : Structures 3D d'enzymes appartenant à différentes sous classes de la classe I. La structure de la MetRS d' <i>E. coli</i> est une version monomérique de celle-ci. Le repliement de Rossmann est en cyan, un domaine CP1 (connective peptide) en jaune, un domaine de reconnaissance de l'anticodon en bleu. Les domaines spécifiques des AARS Ib, domaine 3, 4 et 5 sont respectivement en orange rouge et bleu.....	29
Figure 9 : Structures 3D d'enzymes appartenant à différentes sous classes de la classe II. .	31
Figure 10 : Distribution des éléments d'identité de l'ARNt. (Giege <i>et al.</i> , 1998).	33
Figure 11 : Symétrie de l'interaction des AARS avec l'ARNt. A gauche une AARS de classe I et à droite une AARS de classe II (Arnez and Moras, 1997).....	34
Figure 12 : Schéma représentant les différentes méthodes d'édition chez les AARS (Reynolds <i>et al.</i> , 2010).	36
Figure 13 : Structure 3D d'un variant monomérique de la MetRS d' <i>E. coli</i> avec le repliement de Rossmann en jaune, le CP (connective peptide) en vert, le domaine KMSKS en rose ainsi que le domaine de reconnaissance de l'ARNt en bleu (3H9C).	38
Figure 14 : Organisation modulaire de la MetRS. Le corps de l'enzyme est représenté en vert, le domaine C-terminal responsable de la dimérisation est représenté en bleu, les domaines possibles en N-terminal sont représentés en violet ou en rouge et le domaine situé à l'extrémité C-terminal de l'enzyme de type hélice-tour-hélice est représenté en orange.....	40
Figure 15 : Etapes élémentaires de l'activation de la méthionine par l'ATP	41
Figure 16 : Site de reconnaissance de la méthionine de la MetRS d' <i>E. coli</i>	43
Figure 17 : Structure de la MetRS d' <i>E. coli</i> liée au méthioninyl-adiénylate (analogue du méthioninyl-adiénylate, PDB : 1PG0).....	44

Figure 18 : Structure du complexe MetRS:méthionyl adénylate:PPi:Mg ²⁺ de <i>Leishmania major</i> (PDB : 3KFL). Cette figure illustre les interactions entre les phosphates, le magnésium et les lysines (332 et 335). Ces dernières permettent la délocalisation des charges négatives des phosphates.....	45
Figure 19 : Vue du domaine C-terminal en hélice α de la MetRS d' <i>E. coli</i> . Les principaux résidus impliqués dans la liaison de l'anticodon CAU des ARNt ^{Met} sont représentés en bâtons.....	46
Figure 20 : Structure de la MetRS monomérique d' <i>Aquifex aeolicus</i> complexée avec l'ARNt ^{Met} (PDB : 2CT8). Dans cette structure l'extrémité CCA de l'ARNt n'est pas visible.	46
Figure 21 : Mécanisme de cyclisation de l'homocystéine (Jakubowski, 2011)	48
Figure 22 : Système de traduction avec le codon UAG réassigné. Figure tirée de (Chin, 2017).....	51
Figure 23 : Système de traduction faisant intervenir un codon quadruplet. Figure tirée de (Chin, 2017)	52
Figure 24 : Représentation schématique du mécanisme d'évolution dirigée spécifique aux AANC portant un groupement azido, ici l'azidonorleucine. Figure tirée de (Tanrikulu <i>et al.</i> , 2009).....	54
Figure 25 : Représentation schématique des étapes du cycle de « Multiplexe Automated Genome Engineering » (MAGE). Figure tirée de (Wang <i>et al.</i> , 2009).	55
Figure 26 : Représentation schématique du processus de « Phage Assisted Continuous Evolution » PACE. Figure tirée de (Packer and Liu, 2015). Dans un réacteur (bleu), les cellules hôtes d' <i>E. coli</i> (vert) permettent les développements et l'évolution d'une population de bactériophages (bleu).	56
Figure 27 : Rôle d'EF-Tu dans les différentes étapes du mécanisme bactérien de traduction. Figure tirée de (Uhlenbeck and Schrader, 2018).....	57
Figure 28 : ARNt ^{Pro} modifié pour faciliter l'incorporation des acides aminés D ou β . Figure tirée de (Kato and Suga, 2018).....	59
Figure 29 : Représentation schématique du système de traduction <i>in vitro</i> (PURE system. Figure tirée de (Shimizu <i>et al.</i> , 2001))	60
Figure 30 : Représentation schématique du système de traduction <i>in vitro</i> réduit. Figure tirée de (Forster <i>et al.</i> , 2003)	61
Figure 31 : Schéma montrant la synthèse d'ARNt-aminoacylé catalysée par un flexizyme fixé sur résine. Figure tirée de (Murakami <i>et al.</i> , 2003b). Le flexizyme reconnaît d'une part l'acide aminé et d'autre part la tige acceptrice de l'ARNt.....	62
Figure 32 : Acides aminés α (A), β^2 (B), β^3 (C), Acide trans-2-amino cyclopropanoïque (D), Acide trans-2-amino cyclopentanoïque (E), Acide trans-2-amino cyclohexanoïque (F) ...	63
Figure 33 : Structure chimique des acides aminés α et β^3	63
Figure 34 : Structure chimique de peptides contenant un et deux acides aminés β	64
Figure 35 : Représentation des différents types d'hélices obtenus pour des β -peptides (Raynal, 2002).....	65
Figure 36 A : Séquence peptidique du neuropeptide Y et d'un de ses analogues contenant des AA- β écrits X. B : Formule développée de l'analogue du neuropeptide Y contenant deux acides aminés β (Koglin <i>et al.</i> , 2003).....	66

Figure 37 : Fagot de peptides β , octamérique et tetramérique (Daniels <i>et al.</i> , 2007; Horne <i>et al.</i> , 2007).....	66
Figure 38 : Structure de la MetRS monomérique cocrystallisée avec la β^3 -L- méthionine. Les environnements de la β^3 -L- méthionine et ceux des molécules de citrate et glycérol observées dans le cristal sont montrés.	72
Figure 39 : Analogues de la méthionine, A : Méthionine 1 : Norvaline 2 : Homocystéine 3 : Norleucine 4 : Ethionine.....	90
Figure 40 : Analogues de la méthionine, A : Méthionine ; 1 : Norvaline ; 4 : Norleucine ; 5 : Acide 2-amino-5-hexenoïque 6 : Acide 2-amino-5-hexynoïque ; 7 : Acide <i>cis</i> -2-amino-4-hexenoïque ; 8 : Acide <i>trans</i> -2-amino-4-hexenoïque ; 9 : Acide 6,6,6-trifluoro-2-aminohexanoïque ; 10 : Acide 2-aminoheptanoïque.....	90
Figure 41 : Méthionine (Met, A), Selenométhionine (SeMet, B), méthoxinine (MOX, C), β -méthionine (β -Met, D), Acide glutamique 5-méthyle ester (Glu-ester, E), Desmeninol (DES, F)	91
Figure 42 : Spectre LC-HRMS des composés C, E et F.....	91
Figure 43 : Mécanisme réactionnel de la réaction d'échange isotopique [32 P] PPi-ATP.....	92
Figure 44 structure d'analogue de la méthionine A: Ethionine AC: superposition de l'Ethionine et de l'acide glutamique 5-méthyle ester C: Acide glutamique 5-méthyle ester	93
Figure 45 : Liaison de la sélénométhionine à la MetRS d' <i>E. coli</i>	96
Figure 46 : Liaison de la méthoxinine à la MetRS d' <i>E. coli</i>	97
Figure 47 : Liaison de l'acide glutamique 5-méthyle ester à la MetRS d' <i>E. coli</i>	98
Figure 48 : Structure du complexe MetRS: β -Met (rose) superposée à la structure du complexe MetRS :Glu-ester (bleu)	99
Figure 49 : Liaison du Desmeninol à la MetRS d' <i>E. coli</i>	100
Figure 50 : Structure du complexe MetRS: β -Met (rose) superposée à la structure de complexe MetRS : méthioninyl-adénylate (bleu).....	102
Figure 51 : Structure de la poche de reconnaissance de la méthionine de variants de la MetRS d' <i>E. coli</i>	106
Figure 52 : Structure du variant A12N de la MetRS d' <i>E. coli</i> (<i>violet</i>) superposée à celle du complexe WT:Met (bleu).....	107
Figure 53 : Influence de la mutation D52E sur la structure de la MetRS d' <i>E. coli</i>	108
Figure 54 : Influence de la Mutation A12S sur la structure de la MetRS apo et celle du complexe MetRS :Met.....	109
Figure 55 : Influence de la Mutation A12S sur la structure du complexe MetRS : β -Met	110
Figure 56 : Structure complexe MetRS: β -Met superposée à celle du complexe méthioninyl-adénylate.....	111
Figure 57 : Site de reconnaissance de la méthionine complexé à la β -Met. Les chaînes latérales des résidus sélectionnés pour être modifiés sont colorées en orange.	113
Figure 58 : Graphique comparant les rapports des (kfKI) β – Met(kcatKm)Met des différent variants produits et testé à partir de l'étude <i>in silico</i>	120
Figure 59 : Schéma de la voie métabolique du dTMP en noir chez <i>E. coli</i> en bleu la voie alternative présente chez <i>Lactobacillus leichmannii</i> et <i>Lactobacillus helveticus</i>	122
Figure 60 : Représentation schématique de la méthode d'évolution dirigée	123

Figure 61 : Correspondance du remplacement de deux acides aminés par un AA- α (gauche) ou un AA- β (droite)	124
Figure 62 : Structure de la NDT hexamérique de <i>Lactobacillus leichmannii</i> cocrystallisée avec de la 5-méthyl-2'-deoxypseudouridine (analogue de la thymidine) à gauche. Structure d'un dimère de NDT à droite (1F8Y).....	124
Figure 63 : Structure la NDT de <i>L. leichmannii</i> à proximité des résidus 24 à 25.	125
Figure 64 : Structure la NDT de <i>L. leichmannii</i> à proximité des résidus 96-99.....	126
Figure 65 : Structure la NDT de <i>L. leichmannii</i> à proximité des résidus 124-126 (A) et des résidus 136 et 137 (B).	126
Figure 66 : Liaison du 5-méthyl-2'-deoxypseudouridine à la NDT.....	128
Figure 67 : Structure d'un dimère de NDT.	129
Figure 68 : Culture sur boîte MH dT (300 μ M) à gauche et DU+T (300 μ M) à droite après 48h.....	130
Figure 69 : Schéma représentant le concept de la synthèse peptidique en phase solide...	133
Figure 70 : Test d'incorporation du dipeptide L- β^3 -Met-His dans une souche d' <i>E. coli</i> Δ His	133
Figure 71 : Boîte de culture avec 100 μ l de précurseur de la méthionine	136
Figure 72 : Test de croissance d'une souche Δ met A en présence de Boc-D-Met (3D).....	137
Figure 73 : Schéma des états pré-stationnaires de la réaction d'activation de la méthionine	148

Table des tableaux

Tableau 1 : Classification des aminoacyl-ARNt synthétases	24
Tableau 2 : Correction d'erreurs chez les AARS (Jakubowski and Goldman, 1992; Jakubowski, 1999; Mascarenhas <i>et al.</i> , 2009; Perona and Gruic-Sovulj, 2013)	35
Tableau 3 : Paramètres cinétiques de la réaction d'échange ATP-[³² P] PPi et constantes de dissociation	93
Tableau 4 : Données collectées et statistiques d'affinement	95
Tableau 5 : Vitesses initiales d'activation de la Méthionine et de la β-méthionine par différents variants (2mM Met ou 2mM β-Met, 2mM ATP et 2mM PPi)	103
Tableau 6 : Constantes d'aminacylation de la MetRS WT et de son variant A12S pour la Met et la β-Met (contaminée).....	104
Tableau 7 : Données collectées et statistiques d'affinement	105
Tableau 8 : Douze meilleurs variants proposés par l'étude <i>in silico</i> classés en fonction de leur affinité pour le méthionyl adénylate.....	114
Tableau 9 : Paramètres thermodynamique calculés pour les vingt meilleurs variants vis-à-vis du β-Met-adénylate.....	115
Tableau 10 : Paramètres cinétiques de la réaction d'échange [32P] PPi-ATP avec la méthionine classé. L'ordre des variants est l'ordre décroissant des affinités calculées.	116
Tableau 11 : Paramètres de la réaction d'échange isotopique ATP :PPi mesurés en présence de méthionine ou de β-méthionine. Dans ces mesures le stock de β-Met utilisé est contaminé à 0.09% par de la méthionine.	117
Tableau 12 : Efficacités catalytiques d'activation de la méthionine ou de la β-méthionine (ND : non déterminé)	119
Tableau 13 : Liste des mutations réalisées dans le gène codant pour la NDT (Figure 67)	128
Tableau 14 : Résultats des tests de criblage négatif.....	130
Tableau 15 : Liste des précurseurs potentiels de la méthionine testés.....	135
Tableau 16 : Liste des précurseurs qui ne sont pas de bonnes sources de méthionine	136

Publications et communications

Nigro, G., Bourcier, S., Lazennec-Schurdevin, C., Schmitt, E., Marlière, P., and Mechulam, Y. (2019) Use of β 3-methionine as an amino acid substrate of *Escherichia coli* methionyl-tRNA synthetase. *J Struct Biol* 107435.

Opuu, V., Nigro, G., Gaillard, T., Schmitt, E., Mechulam, Y., and Simonson, T. (2020) Adaptive landscape flattening allows the design of both enzyme: Substrate binding and catalytic power. *PLOS Comput Biol* 16: e1007600.

Présentation orale : Engineering Conference International in Enzyme Engineering XXIV 24-28 September 2017, Toulouse. G. Nigro, E. Schmitt, P. Marlière, Y. Mechulam : « Recognition of l- β -homomethionine by methionyl-tRNA synthetase ».

Présentation par affiche à la conférence organisée par la FRENCH INFRASTRUCTURE FOR INTEGRATED STRUCTURAL BIOLOGY le 13 décembre 2018 à Gif-sur-Yvette. Giuliano Nigro, Emmanuelle Schmitt and Yves Mechulam : « Recognition of l- β -homomethionine by methionyl-tRNA synthetase ».

Introduction

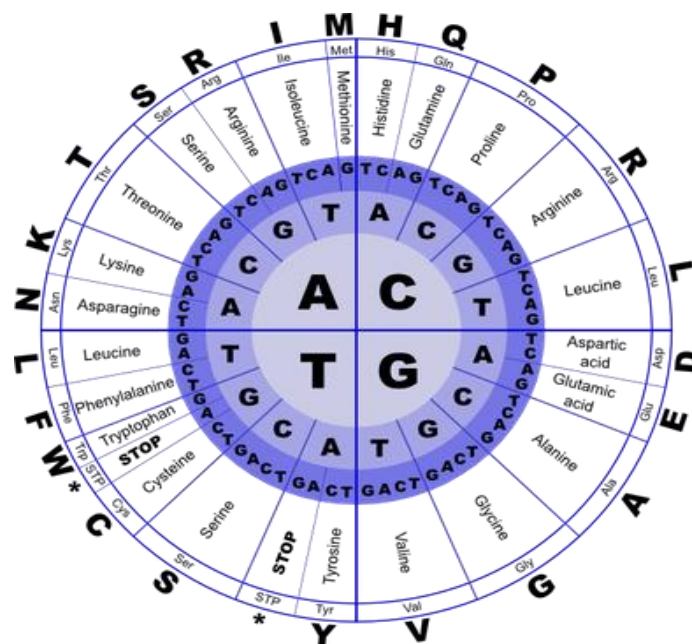
I Biologie de synthèse

I. 1 Généralités

Par sa richesse et sa diversité, la nature a produit des organismes vivants dont les fonctions et la capacité à évoluer et à s'adapter, dépendent des protéines qu'ils produisent. On constate en effet une grande diversité dans les protéines existantes et leurs activités catalytiques. En 1912, Stéphane Leduc (Leduc, 1912) pose les bases de la biologie synthétique en argumentant qu'il n'y a pas de différences entre le vivant et les phénomènes physiques. Ainsi, si le fonctionnement d'un phénomène est compris, il est alors possible de le reproduire et de l'utiliser. Dans cet esprit, de nombreux chercheurs ont contribué par leurs recherches sur le code génétique et le vivant à confirmer cette hypothèse, qui a mené à l'actuelle biologie de synthèse. Dans un premier temps, les enzymes utiles ont pu être extraites et leurs fonctions ont été caractérisées au niveau biochimique et structural (Sedlacek and Smith, 1988). Dans un deuxième temps (Bornscheuer *et al.*, 2012) des méthodes d'ingénierie des protéines ont été développées pour extraire mieux et produire plus. Toutefois, la base proposée par la nature est limitée à certaines fonctions catalytiques. En effet, certaines réactions sont très difficiles, voire impossible dans la nature, telles que des réactions redox, des réactions électrophiles ou encore des réactions avec des molécules non naturelles. Pour y remédier, la nature utilise une variété de cofacteurs ainsi que des modifications post traductionnelles (Morinaka *et al.*, 2018). De par la nécessité d'obtenir des protéines avec des activités nouvelles, des méthodes comme l'évolution dirigée (Agostini *et al.*, 2017) ont été développées. Malgré ces nouvelles techniques, certaines fonctions chimiques et activités catalytiques ne peuvent pas être faites par des enzymes, à cause de la nature même des acides aminés qui composent ces protéines. Il y a donc un besoin d'utiliser des acides aminés non canoniques. Ces derniers sont déjà utilisés en synthèse peptidique pour produire des peptides qui peuvent marquer des anticorps (Li *et al.*, 2016).

Notre objectif est d'ajouter des acides aminés non canoniques *in vivo* en utilisant le système de traduction. La modification de celui-ci est primordiale si l'on veut introduire de tels acides aminés à des positions spécifiques.

Le code génétique définit comment les 64 codons formés par les 4 nucléotides sont décodés de façon unique et redondante par le ribosome pour produire des protéines avec une structure tridimensionnelle complexe, constituées de 20 acides aminés (figure 1). La machinerie qui permet de traduire un ARN messager est universellement conservée. Le code génétique n'est toutefois pas tout à fait universel. Par exemple, certains organismes utilisent des acides aminés atypiques tels que la sélénocystéine ou la pyrrolysine (Ambrogelly *et al.*, 2007). Malgré cela, la plus grande partie des protéines est constituée des 20 mêmes acides aminés canoniques qui en sont les briques élémentaires (Figure 1).



II Système de traduction

II. 1. Généralités

La lecture de l'information génétique pour produire les protéines est un processus primordial pour chaque cellule vivante. La complexité de ce processus reflète à quel point il est évolué. La traduction nécessite en effet de nombreux partenaires.

Le mécanisme d'interprétation fonctionne en deux étapes distinctes. Tout d'abord, la transcription qui assure la copie de l'information génétique portée par l'acide désoxyribonucléique (ADN) en une copie, l'acide ribonucléique messenger, qui permet de transporter l'information jusqu'à la machinerie qui exécutera la seconde étape, la traduction. Cette dernière permet la transformation du code génétique codé en 4 lettres porté par l'ARNm en code protéique à 20 lettres. Ce processus complexe a besoin de plusieurs partenaires pour être exécuté (figure 2) :

- Les ARNms supports de l'information génétique,
- les ribosomes constitués d'ARNs ribosomiques (ARNr) associés à plusieurs protéines, formant deux sous unités distinctes,
- les ARNs de transfert (ARNt).

Tous ces éléments sont importants pour l'extension du code génétique (figure 2).

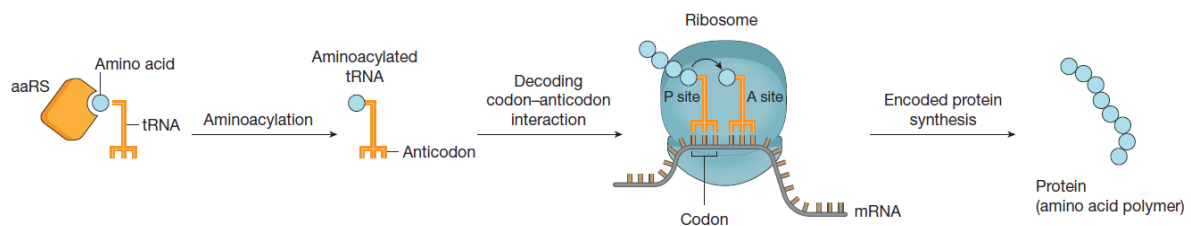


Figure 2 : Processus de traduction (Chin, 2017)

II. 2 Rôle des partenaires

II. 2. 1 Ribosome

Chez les procaryotes, le ribosome est constitué de deux sous unités caractérisées par leurs coefficients de sédimentation en « Svedberg » 30S et 50S qui ensemble forment des particules 70S. Cet ensemble se dissocie à chaque fois que la traduction d'un ARNm est terminée. Cette synthèse catalysée par le ribosome peut être aussi effectuée dans des systèmes de traduction reconstitués *in vitro* (Shimizu *et al.* , 2001; Shimizu *et al.* , 2006).

Les ribosomes de bactéries sont constitués à 60% d'ARNr, le reste étant des protéines ribosomiques. Ils sont composés de 3 types d'ARNr et de 55 protéines. La sous unité 30S bactérienne renferme une molécule d'ARNr 16S et 21 protéines. Elle assure l'appariement correct entre l'anticodon de l'ARNt et le codon présent sur l'ARNm. Elle constitue le centre de décodage du ribosome. La sous unité 50S, elle, est constituée de deux autres ARNr 5S et 23S ainsi que de 34 protéines. Elle constitue le site de formation de la liaison peptidique.

II. 2. 2 ARNt

Les ARNt sont des molécules constituées de nucléotides dont certains sont modifiés après la formation de l'ARNt (eg : pseudouridine, dihydrouracile ...). Beaucoup de ces modifications sont spécifiques des ARNt. Ces molécules présentent une structure bien définie qui comprend 4 parties de bases complémentaires appariées, séparées par des boucles (la boucle TΨC, la boucle anticodon et la boucle DHU), le tout formant une structure dite en « feuille de trèfle » (figure 3). À cette structure caractéristique des ARNt correspond une organisation dans l'espace particulière. Les ARNt ont une forme de L avec les boucles DHU (D) et TΨC (T) au niveau du coude, à l'extrémité du bras long du L, l'anticodon, et à l'extrémité courte, la tige acceptrice 3' (Figure 3).

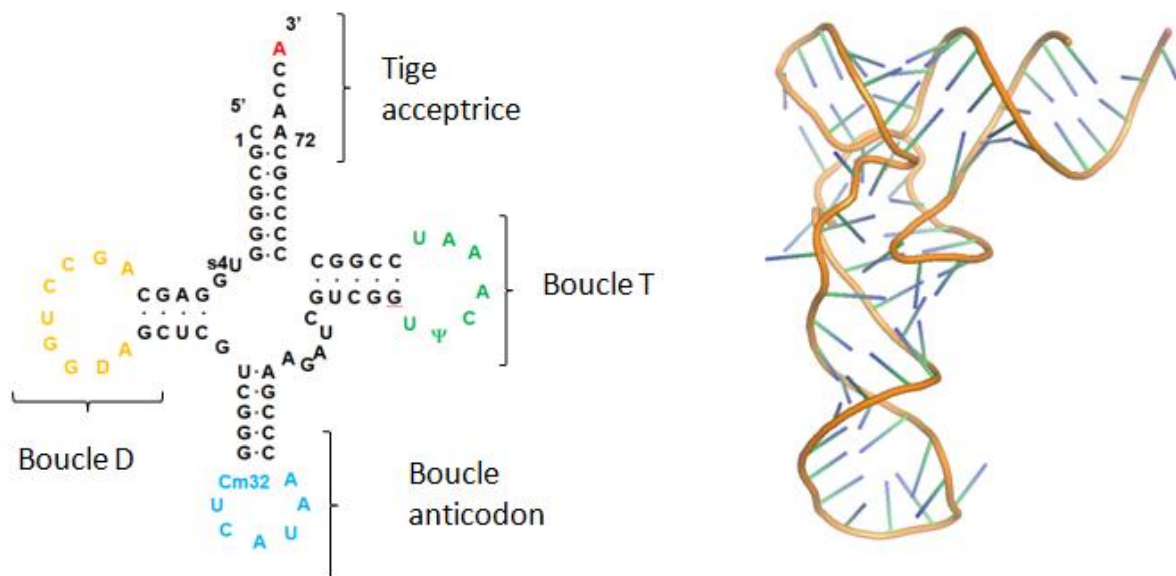


Figure 3 : ARNt^{Met} en forme « feuille de trèfle » (gauche) et en forme repliée (droite)

Les ARNt sont caractérisés par une boucle anticodon qui leur est spécifique. Ces 3 nucléotides, venant s'apparier spécifiquement à leurs codons complémentaires portés par l'ARNm, assurent la correspondance avec l'acide aminé fixé à l'extrémité 3' de la tige ac-

ceptrice. C'est pourquoi la disponibilité d'ARNt aminoacylé est primordiale pour la traduction. Cette réaction d'aminacylation qui permet de fixer les acides aminés à leurs ARNt correspondants est catalysée par une famille d'enzymes : les aminoacyl-ARNt synthétases.

II. 3 Aminoacyl ARNt synthétases (AARS)

Les aminoacyl-ARNt synthétases constituent une famille d'enzymes qui ont toujours la même fonction et qui sont conservées chez tous les organismes. Elles sont indispensables à la vie cellulaire, en assurant la fixation des acides aminés à leurs ARNt correspondants. Pour des raisons de survie, la spécificité de chaque AARS doit être très élevée. En effet, la précision de la traduction de l'information génétique est intrinsèquement liée à la capacité des AARS à aminoacyler chaque ARNt avec l'acide aminé qui lui correspond.

II. 3. 1 Généralités sur les AARS

Les AARS sont des enzymes qui montrent une certaine variété, mais quelques grands traits sont conservés. Il y a chez les AARS deux grandes catégories distinctes, les enzymes de classes I et les enzymes de classes II. Il y a cependant une variabilité dans leurs séquences, leurs longueurs et leurs oligoméries. Autour des deux domaines catalytiques des AARS de classe I et II qui seront décrits plus loin, on peut trouver des modules additionnels pour la reconnaissance de l'ARNt, la correction d'erreurs ou d'autres fonctions spécifiques à certains organismes. Chez les eucaryotes, la complexité des AARS augmente avec l'ajout d'autres domaines plus ou moins grands comme dans l'AspRS (Wolf *et al.* , 1999). De plus les AARS bactériennes diffèrent des autres AARS par leurs gammes de point isoélectrique entre 5 et 6 en contraste avec d'autres espèces comme les AARS mitochondriales ou les AARS cytosoliques humaines (Sissler *et al.* , 2008). La plupart des AARS sont monomériques ou dimériques, mais il y a des exceptions : par exemple, chez *E. coli* l'AlaRS est homotétramérique α_4 (Putney and Schimmel, 1981) et la PheRS est un tétramère $\alpha_2\beta_2$ chez *E. coli* (Fayat *et al.* , 1974a) ou *T. thermophilus* (Safro *et al.* , 1993) tandis que la PheRS mitochondriale est monomérique (Sanni *et al.* , 1991).

La plupart des organismes possèdent 20 AARS, une pour chaque acide aminé canonique. On observe toutefois des organismes où ce n'est pas le cas. Chez *E. coli*, il y a deux gènes

codant pour la LysRS (Leveque *et al.* , 1990), et chez certaines bactéries et une grande partie des archées la GlnRS et/ou l'AsnRS sont absentes. Ces organismes possèdent des GluRS et AspRS non discriminantes capables de reconnaître l'ARNt^{Gln} et l'ARNt^{Asn} ainsi qu'une voie métabolique indirecte qui passe par une transamidation des ARNt^{Gln} et des ARNt^{Asn} chargés par un Glu ou un Asp (Schön *et al.* , 1988b; Schön *et al.* , 1988a; Ibba and Soll, 2001; Schmitt *et al.* , 2005). Malgré ces voies de synthèses alternatives les AARS restent la clé de voute du système de traduction qui relie les acides aminés à leurs anticodons respectifs.

Pour cette raison, cette famille d'enzymes a beaucoup été étudiée que ce soit au niveau structural ou fonctionnel. Le mécanisme qui régit le fonctionnement des enzymes est ainsi très bien connu.

II. 3. 1. 1 Réaction d'aminoacylation

La réaction catalysée par les AARS est une réaction d'aminoacylation qui se fait en deux étapes distinctes (Figure 4). La première étape permet la formation d'un complexe stable constitué d'un aminoacyl adenylate et de l'enzyme. Il est le produit de la reconnaissance spécifique de l'ATP et de l'acide aminé correspondant à l'AARS. La réaction chimique consiste en une attaque nucléophile du carboxyle sur le phosphate α de l'ATP. La présence du cation Mg^{2+} est indispensable à la réaction. Cette étape d'activation de l'acide aminé forme un anhydride mixte très réactif et peu stable lorsqu'il n'est pas protégé par l'enzyme. Cette étape conduit également au relargage du PPi, dégradé dans la cellule par la pyrophosphatase.

La seconde réaction est l'aminoacylation de l'ARNt. Elle permet le transfert de l'acide aminé activé sur l'ARNt ainsi que la libération des différents éléments dans le milieu. Cette réaction consiste en la récupération d'un proton du groupement hydroxyle 2' ou 3' du ribose de l'adénine situé en 3' de la tige acceptrice de l'ARNt au phosphate de l'adenylate, formant un état de transition. Ensuite, un des oxygènes libres de la partie phosphate de l'adenylate attire le proton du groupement -OH qui réagit formant un intermédiaire cyclique qui est converti en ester liant l'acide aminé au ribose de l'ARNt (Giegé and Springer, 2016).

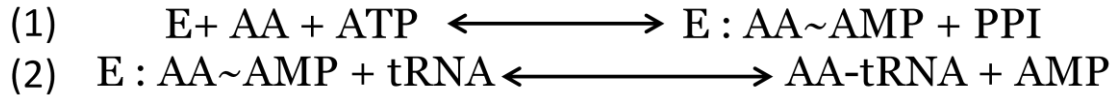


Figure 4 : Mécanisme d'aminocyclation en deux étapes. La première est l'activation de l'acide aminé qui permet de former le complexe E:AA~AMP (1). La seconde est l'estérification de l'acide aminé sur l'ARNt qui permet de charger l'acide aminé sur ce dernier (2).

Dans quelques cas (GlnRS, ArgRS, GluRS), la présence de l'ARNt est nécessaire pour la première étape car d'une part il facilite la formation de l'adenylate et d'autre part évite la réaction de celui-ci avec les chaînes latérales des acides aminés proches (Giegé and Springer, 2016).

II. 3. 1. 2 Organisation des AARS

Malgré leur unité fonctionnelle, les aminoacyl ARNt synthétases sont très diversifiées sur différents plans, la taille, la séquence, et le niveau de multimérisation. C'est pourquoi de nombreux efforts ont été fournis pour les catégoriser.

II. 3. 1. 2. 1 Deux classes d'AARS

La partition des AARS en deux grandes familles est basée sur leurs différences structurales, principalement autour du site actif. Il a été remarqué que chez certaines AARS, le site de liaison de l'ATP se trouve dans un domaine de type repliement de Rossmann (Blow *et al.* , 1983) alors que chez d'autres AARS, l'ATP se lie au niveau d'un feuillet de brins β antiparallèles (Cusack *et al.* , 1990). En parallèle, il a été observé que les AARS à repliement de Rossmann (Classe I, figure 5a) possèdent deux motifs peptidiques caractéristiques HIGH et KMSKS alors que celles à feuillet β antiparallèle (Classe II, figure 5b) possèdent trois autres motifs signatures (Eriani *et al.* , 1990). Ces trois motifs distincts sont décrits de la manière suivante : le motif 1 (g $\Phi\Phi$ XX Φ XPP $\Phi\Phi$), le motif 2 (fRx ϵ) et le motif 3 (gxgxgfd/eR), Φ représentant un résidu hydrophobe et X un résidu quelconque. Les motifs 2 et 3 ont des fonctions similaires aux motifs KSMKS et HIGH alors que le motif 1 est responsable de la dimérisation.

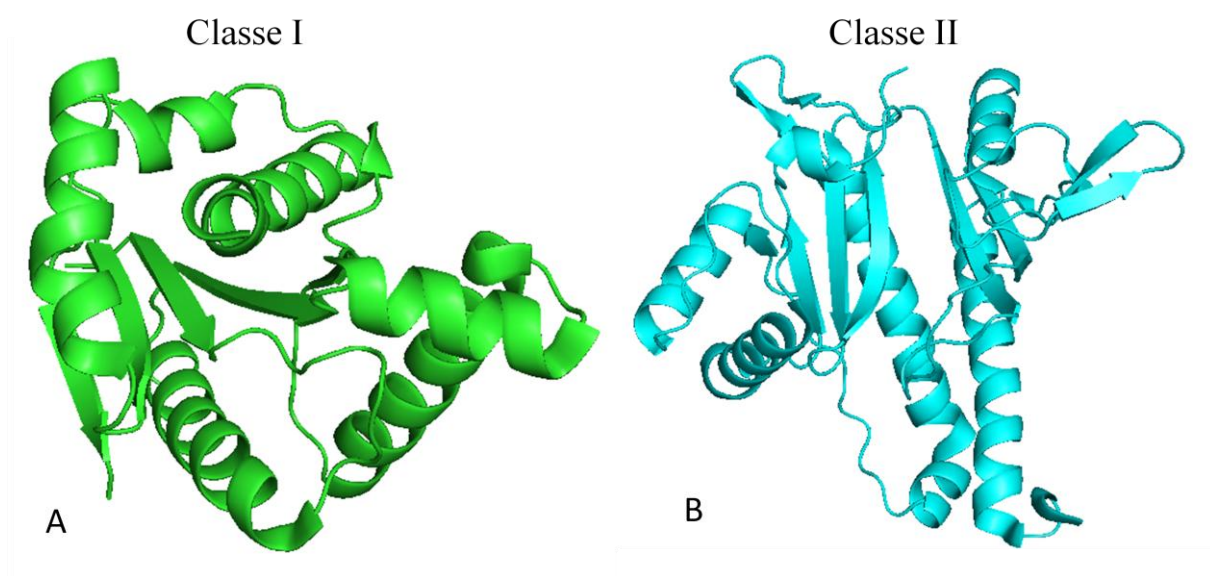


Figure 5 : Repliement de Rossmann de la structure 3D de la MetRS de *E. coli* (A) et feuillet β antiparallèle de la structure 3D de l'AspRS de *E. coli* (B)

Il y a une grande dégénérescence dans les motifs des AARS de classe II, mais une proline et deux arginines sont rigoureusement conservées. Les structures tridimensionnelles ont joué un rôle important dans la découverte des motifs des AARS (Blow *et al.*, 1983; Eriani *et al.*, 1990). Ceci signifie que la conservation de la structure ne nécessite pas forcément de conservation de séquence importante (Figure 5). Il a été montré que 15% d'identités de séquence sont suffisants pour conserver la structure dans une série d'AARS spécifiques d'un même acide aminé (O'Donoghue and Luthey-Schulten, 2003). Des rapprochements entre les différentes structures d'AARS de même classe ont pu être faits, ce qui a permis une sous classification, a, b et c représentée dans le tableau 1 (Cusack *et al.*, 1990).

Les deux classes d'AARS montrent le même mécanisme de réaction. Toutefois les AARS de classe I lient l'acide aminé sur le groupement 2'OH de l'adénine 76 de l'ARNt alors que les AARS de classe II le lient en 3'OH de l'ARNt.

Tableau 1 : Classification des aminoacyl-ARNt synthétases

Classe I		Classe II	
AARS	Structure oligomérique	AARS	Structure oligomérique
Classe Ia		Classe IIa	
LeuRS	α	SerRS	α_2
IleRS	α	ThrRS	α_2
ValRS	α	ProRS	α_2
ArgRS	α	HisRS	α_2
CysRS	α	GlyRS	α_2

MetRS	α ou α_2	Classe IIb	
Classe Ib		AspRS	α_2
GluRS	α	AsnRS	α_2
GlnRS	α	LysRS	α_2
LysRS	α	Classe IIc	
Classe Ic		AlaRS	α ou α_4
TyrRS	α_2	PheRS	α_2 , $\alpha_2\beta_4$ ou α
TrpRS	α_2	PylRS	α_2
		SepRS	α_4

II. 3. 1. 2. 1. 1 Classe I

Les AARS de classe I se caractérisent par leurs structures 3D construites autour d'un repliement de Rossmann, au sein duquel se trouvent deux motifs particuliers, le motif KMSKS et le motif HIGH. Ceux-ci sont toujours positionnés au même endroit dans le repliement de Rossmann. Dans cette classe on retrouve les AARS suivantes : MetRS, ValRS, IleRS, ArgRS, CysRS, LeuRS, GlnRS, GluRS, TyrRS, TrpRS ainsi qu'une LysRS très particulière trouvée chez certaines archées (Ibba *et al.* , 1997; Terada *et al.* , 2002).

Le repliement de Rossmann a été initialement décrit chez certaines déshydrogénases (Rossmann *et al.* , 1974). Cette structure qui caractérise les AARS de Classe I est formée d'une succession de brins β parallèles qui composent un feuillet. Ces brins sont reliés par des hélices α (Figure 5). La poche formée à la jonction entre les deux moitiés du repliement constitue le site de liaison des nucléotides et acides aminés substrats. Les deux motifs caractéristiques HIGH et KMSKS sont spécifiques des AARS et liés à la réaction d'aminacylation.

Le motif HIGH est positionné dans la première boucle du repliement de Rossmann. Des expériences de mutagenèse dirigées sur les différents acides aminés du motif ont permis de comprendre le rôle de chacun d'eux (Leatherbarrow *et al.* , 1985; Schmitt *et al.* , 1995). Les deux histidines se distinguent par le rôle de leur chaîne latérale qui maintient l'ATP et stabilise l'état de transition lors de l'activation de l'acide aminé. Récemment une étude structurale a montré l'importance d'une des hélices du repliement de Rossmann (Kaiser *et al.* , 2018). En effet, cette étude montre que deux liaisons hydrogène sont très conservées entre l'ATP et la chaîne principale de l'hélice α_3 chez les AARS de classe I. Cela ne forme

toutefois pas un troisième motif caractéristique car ce sont les atomes de la chaîne principale qui interagissent avec l'adénosine de l'ATP.

Le second motif caractéristique KMSKS a été étudié tout autant que le motif HIGH. Ce motif avait au départ été identifié par expériences de marquage d'affinité de la MetRS et de la TyrRS de *E. coli* avec des ARNt oxydés par du periodate (Hountondji *et al.* , 1985; Hountondji *et al.* , 1986). Les études structurales des différentes AARS de classe I ont montré que ce motif possédait une localisation dans l'espace constante, à la sortie du repliement de Rossmann et proche du motif HIGH. Des expériences de mutagenèse ont montré que ce motif était très impliqué dans l'état de transition lors de la formation de l'aminoacyl adénylate (Fersht, 1988; Mechulam *et al.* , 1991; First and Fersht, 1993a; First and Fersht, 1993b; First and Fersht, 1993c; Schmitt *et al.* , 1994). Les lysines interviennent pour neutraliser les charges négatives des phosphates, ce qui facilite l'attaque nucléophile du carboxyle de l'acide aminé au niveau du phosphate α de l'ATP. De plus, ce motif est entouré d'acides aminés tels que des glycines, ce qui lui confère une certaine mobilité, et permet son basculement vers l'ATP lors de la catalyse (Schmitt *et al.* , 1994). Ce basculement de la boucle KMSKS a été confirmé au niveau structural chez la MetRS de *Leishmania major* (Kaiser *et al.* , 2018). Sur cette même structure, il est possible de voir l'importance des ions Mg^{2+} qui contribuent à dissiper les charges négatives du pyrophosphate.

II. 3. 1. 2. 1. 2 classe II

Les AARS de classe II se caractérisent par leurs structures 3D qui possèdent un feuillet β antiparallèle (figure 5), initialement mis en évidence par la résolution des structures de l'AspRS de levure (Ruff *et al.* , 1991) et de la SerRS d'*E. coli* (Cusack *et al.* , 1990). Le feuillet est composé de sept brins β agencés de façon antiparallèle et flanqués par des hélices α . De plus, les motifs conservés 1, 2 et 3 possèdent des positions constantes dans la structure tridimensionnelle. Dans cette classe on retrouve les AARS suivantes : GlyRS, HisRS, ProRS, SerRS, ThrRS, AsnRS, AspRS, LysRS et la PheRS.

La compréhension des AARS de classe II et de leurs processus catalytiques est principalement due à la résolution de leurs structures Apo et Holo. Ces structures ont permis

d'établir le positionnement des substrats dans la poche. L'ATP prend ainsi une conformation coudée, la partie pyrophosphate se repliant au dessus de l'adénine. Cette dernière forme une interaction d'empilement avec un résidu aromatique conservé dans les AARS de classe II.

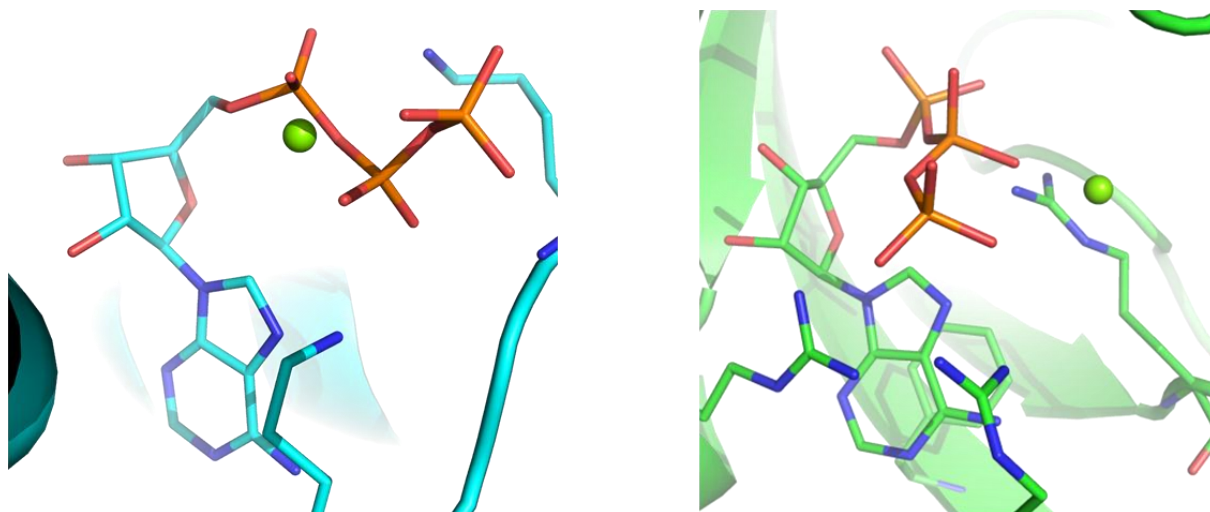


Figure 6 : Positionnement de l'ATP dans la TrpRS de *Bacillus stearothermophilus* (Classe I) (5DK4) à gauche et dans la SerRS de *Bradyrhizobium diazoefficiens* (Classe II) (3MEY) à droite. Sur la structure de la TrpRS deux lysines du motif KMSKS en interaction avec l'ATP sont visibles. Dans la structure de la SerRS deux arginines (motifs 2 et 3) interagissent avec l'ATP.

Les motifs 2 et 3 participent activement à la réaction d'activation, les arginines présentes dans ces motifs ont pour rôle de stabiliser l'ATP en formant une sorte de pince (Kaiser *et al.*, 2018) (Figure 6). L'arginine du motif 3 interagit d'une part avec l'adénosine pour stabiliser sa position et d'autre part avec un des phosphates de l'ATP, tandis que l'arginine du motif 2 interagit avec le phosphate α et stabilise l'état de transition (Schmitt *et al.*, 1998). Pour différentes AARS de classe II comme la SerRS de *Thermus thermophilus* ou l'AspRS de *Pyrococcus kodakaraensis*, des ions magnésium ont été observés autour de l'ATP formant des interactions ioniques avec les phosphates. Les ions Mg^{2+} , permettent de stabiliser la structure coudée de l'ATP. Un des ions magnésium assure le pontage de la charge entre les phosphates α et β . Cet atome de magnésium qui est présent aussi chez les AARS de classe I a pour rôle d'attirer les électrons loin du groupement phosphate α facilitant ainsi la réaction nucléophile (Belrhali *et al.*, 1995; Schmitt *et al.*, 1998).

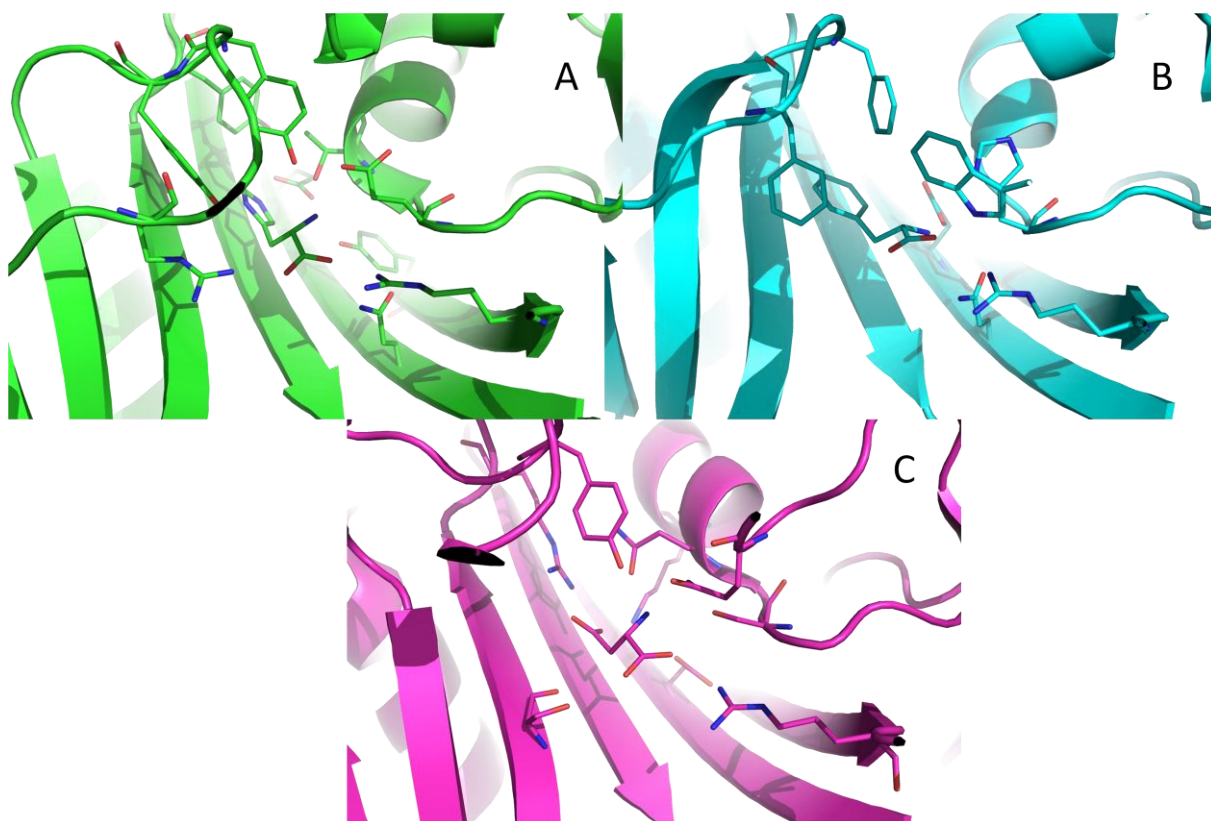


Figure 7 : Sites de reconnaissance des acides aminés de trois AARS de classe II. A : HisRS de *Burkholderia thailandensis* (4E51), B : PheRS de *T. thermophilus* (1B70) et C : AspRS de *T. thermophilus* (3NEL)

La reconnaissance de l'acide aminé chez les AARS de classe II se fait par l'intermédiaire d'interactions spécifiques avec les résidus du site actif. L'amine de l'acide aminé est maintenue par un résidu acide et des résidus chargés (figure 7). Cette interaction contribue au bon positionnement du carboxylate de l'acide aminé pour la réaction d'activation.

II. 3. 1. 2. 2 Sous classes d'AARS

Autour des structures principales des AARS de classe I et II on retrouve différents modules qui permettent de les classer dans trois sous catégories pour les deux classes. Ces sous classes reflètent des homologies de séquences et des ressemblances structurales (Cusack *et al.*, 1990).

II. 3. 1. 2. 2. 1 Classe I, a, b, c

Les sous classes d'AARS de classe I dépendent de la structure autour du domaine catalytique et du type d'acide aminé que l'AARS lie (Burbaum and Schimmel, 1991). En particulier, quand la séquence du repliement de Rossmann du site actif est interrompue par un autre élément structural (CP1), celui-ci est caractéristique des sous classes de la classe I. La sous classe Ia regroupe les synthétases d'acide aminé hydrophobe ainsi que l'ArgRS, les Ib

regroupent les AARS des gros acides aminés et les AARS Ic regroupent celles qui utilisent les acides aminés hydrophiles (Figure 8).

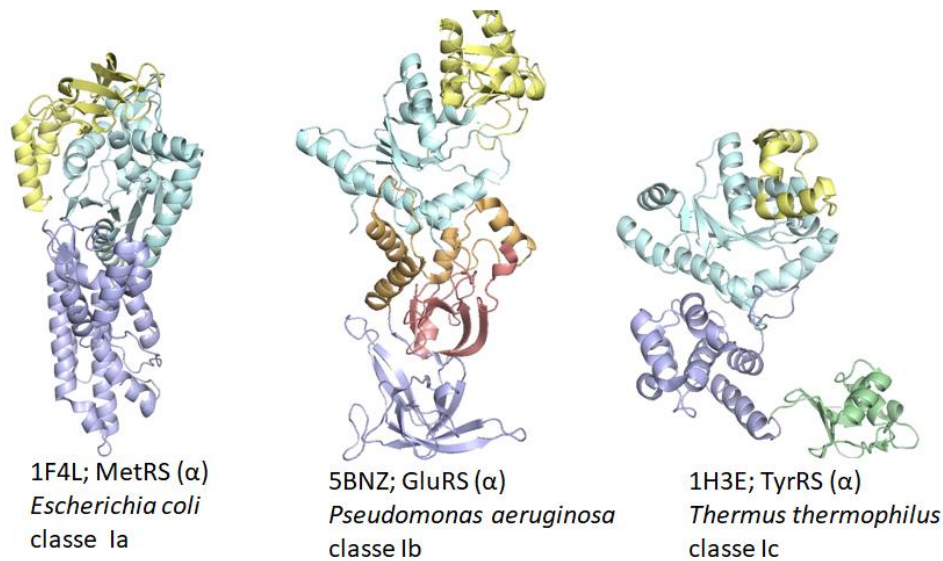


Figure 8 : Structures 3D d'enzymes appartenant à différentes sous classes de la classe I. La structure de la MetRS d'*E. coli* est une version monomérique de celle-ci. Le repliement de Rossmann est en cyan, un domaine CP1 (connective peptide) en jaune, un domaine de reconnaissance de l'anticodon en bleu. Les domaines spécifiques des AARS Ib, domaine 3, 4 et 5 sont respectivement en orange rouge et bleu.

Les AARS Ia ont trois domaines spécifiques. Le repliement de Rossmann, un domaine CP1 (connective peptide) et un domaine de reconnaissance de l'anticodon. Le domaine CP1 (figure 8) qui relie les deux moitiés du repliement de Rossmann est lié à une activité de correction d'erreurs pour l'IleRS, la ValRS et la LeuRS. De plus, toutes les AARS sont monomériques exceptées certaines MetRS qui ont un module supplémentaire en C terminal d'une centaine d'acides aminés dédié à la dimérisation. L'ArgRS a aussi une spécificité : pour l'activation de l'arginine, l'ARNt^{Arg} doit être présent.

Les AARS Ib se caractérisent par la présence de cinq domaines, le repliement de Rossmann (domaine 1), le domaine de reconnaissance du bras accepteur (domaine 2), un domaine qui interagit avec le domaine variable de l'ARNt (domaine 3) ainsi que deux autres domaines (domaines 4 et 5) l'un interagissant avec l'anticodon de l'ARNt et l'autre avec le bras variable. Ces domaines sont clairement visibles sur les structures de la GlnRS :tRNA^{Gln} (Rould *et al.*, 1989; Nureki *et al.*, 1995) et la GluRS de *T. thermophilus*. Il est à noter que les deux AARS de la sous classe Ib ont une spécificité réactionnelle : pour l'activation de l'acide aminé la présence de l'ARNt est nécessaire, comme pour l'ArgRS. Les trois premiers domaines des GluRS et des GlnRS sont très similaires mais les domaines 3 et 4 sont très

différents. Ils sont composés de faisceaux d'hélices α chez la GluRS et d'un feuillet de brins β antiparallèles formant un tonneau chez la GlnRS. La LysRS particulière rencontrée chez certaines archées fait partie de la catégorie Ib à cause de sa ressemblance structurale avec la GluRS. Comme les autres AARS de classe Ib, la présence de l'ARNt est requise pour la formation de l'adenylate (Ibba, 1999). De façon assez surprenante, la LysRS de classe I et celle de classe II ont des poches de reconnaissance de la lysine possédant des caractéristiques semblables malgré la différence de structure de leurs sites actifs.

La sous classe Ic réunit la TyrRS et la TrpRS. Ces enzymes sont constituées de 4 domaines, le repliement de Rossmann (cyan), un domaine de reconnaissance de l'anticodon (vert), un domaine de reconnaissance de la longue boucle variable de l'ARNt (bleu) et un domaine de dimérisation (en jaune). Ce sont des dimères qui, malgré leurs ressemblances de séquence assez faibles, ont des structures semblables. Chez les archées et les eucaryotes, il y a des modules supplémentaires par rapport à ces mêmes AARS chez les bactéries. Ces modules supplémentaires sont liés à des activités indépendantes de l'aminocyclation. Ces modules complémentaires peuvent être liés à des voies métaboliques du signal de transduction, notamment pour la stimulation et l'inhibition de l'angiogenèse comme c'est le cas pour la TyrRS et la TrpRS chez l'humain (Tzima and Schimmel, 2006).

II. 3. 1. 2. 2 Classe II, a, b, c

Autour du feuillet β antiparallèle (en cyan sur la figure 9) qui caractérise les AARS de classe 2 viennent se greffer des domaines supplémentaires qui permettent de définir des sous-classes (Figure 9).

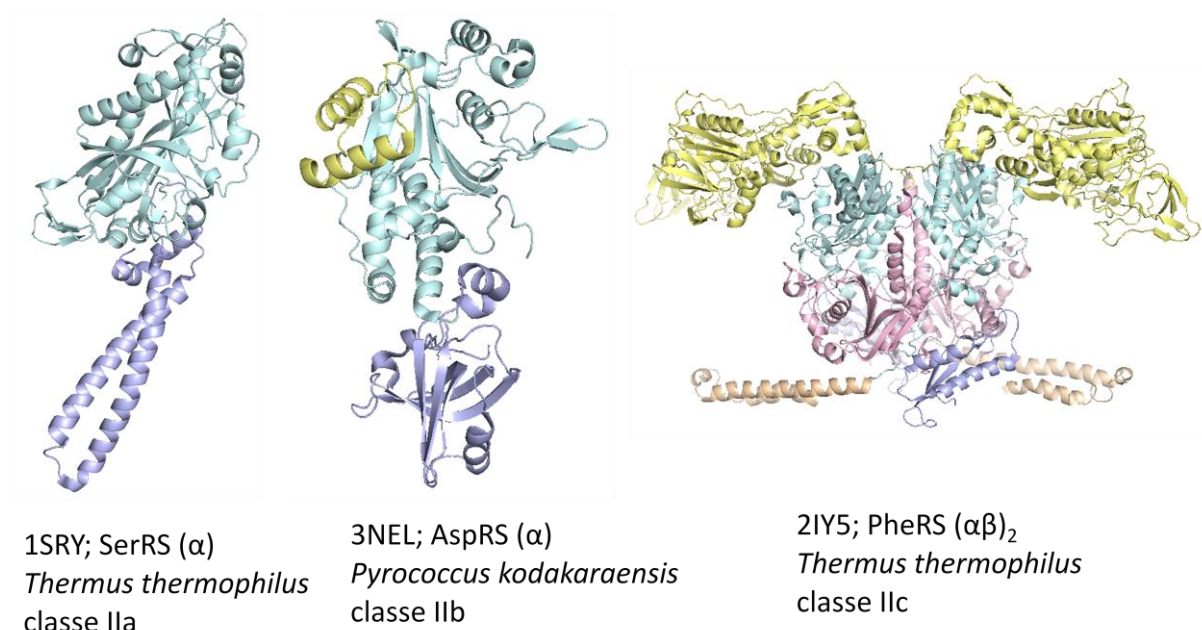


Figure 9 : Structures 3D d'enzymes appartenant à différentes sous classes de la classe II.

Le feuillet β antiparallèle est en cyan, le domaine de reconnaissance de l'ARNt en bleu

Les AARS IIa forment une sous classe hétérogène d'homodimères α_2 , exceptée la GlyRS qui peut être un tétramère $\alpha_2\beta_2$. Ces AARS chargent les acides aminés petits et polaires sur les ARNt (Gly, Ser, Thr, Pro, His). Cette sous classe se caractérise aussi par son domaine de reconnaissance de l'anticodon en C terminal. Une exception est la SerRS qui lie le long domaine variable de l'ARNt^{Ser} par une structure en hélices. La ThrRS possède un domaine supplémentaire qui interagit avec le bras accepteur.

Les AARS de classe IIb AspRS, AsnRS et LysRS se caractérisent par un domaine de reconnaissance de l'anticodon, qui forme un repliement de type « OB fold » (en bleu), De plus ce sont toutes des homodimères α_2 .

Les AARS de classe IIc se caractérisent par le fait que ce sont des tétramères, l'AlaRS étant α_4 et la PheRS $\alpha_2\beta_2$. Elles ont été classées dans cette sous classe par défaut, car elles ne ressemblent pas assez aux AARS des autres sous classes. Le dimère $\alpha\beta$ de la PheRS se caractérise par 5 domaines, 2 pour la partie α et 3 pour la partie β .

II. 3. 1. 3 Reconnaissance de l'ARNt

La reconnaissance des ARNt ne sera pas traitée en profondeur car ce n'est pas l'élément principal qui sera développé dans cette thèse, mais il convient d'en donner les grands principes.

II. 3. 1. 3. 1 identité des ARNt

Les ARNt ayant des structures 3D très similaires, le problème de leur reconnaissance par les AARS a longtemps suscité des interrogations. A priori, on pouvait imaginer que des nucléotides particuliers ou encore des motifs structuraux étaient reconnus par les AARS. En 1990, grâce à la possibilité de produire des ARNt de séquence voulue, une première étude approfondie a été menée sur l'ARNt^{Ala}. De façon étonnante, il a été montré que la présence de la paire G3-U70 était nécessaire et suffisante pour que l'ARNt soit reconnu par l'AlaRS (Hou and Schimmel, 1989). Cette avancée sur la reconnaissance des ARNt a été suivie de nombreuses études sur tous les différents ARNt. Les résultats de ces recherches ont permis d'identifier trois types de déterminants sur les différents ARNt ;

- les déterminants majeurs, qui sont les éléments principaux de reconnaissance de l'ARNt par les AARS ;

- les déterminants mineurs, qui sont des éléments secondaires permettant d'augmenter la spécificité d'un facteur 10 à 100 ;

- les anti-déterminants ou déterminants négatifs, qui permettent d'empêcher la reconnaissance d'un ARNt donné par une autre AARS.

La figure 10 montre la position des éléments d'identité des ARNt chargés par les AARS de classe I et II. La taille des sphères est proportionnelle à la fréquence de reconnaissance du nucléotide considéré par les AARS. On peut constater que deux régions des ARNt sont majoritairement reconnues : la boucle anticodon et la tige acceptrice. Ces deux régions sont finalement assez logiques puisque l'anticodon caractérise fonctionnellement un ARNt et que la tige acceptrice est nécessairement en contact avec l'AARS.

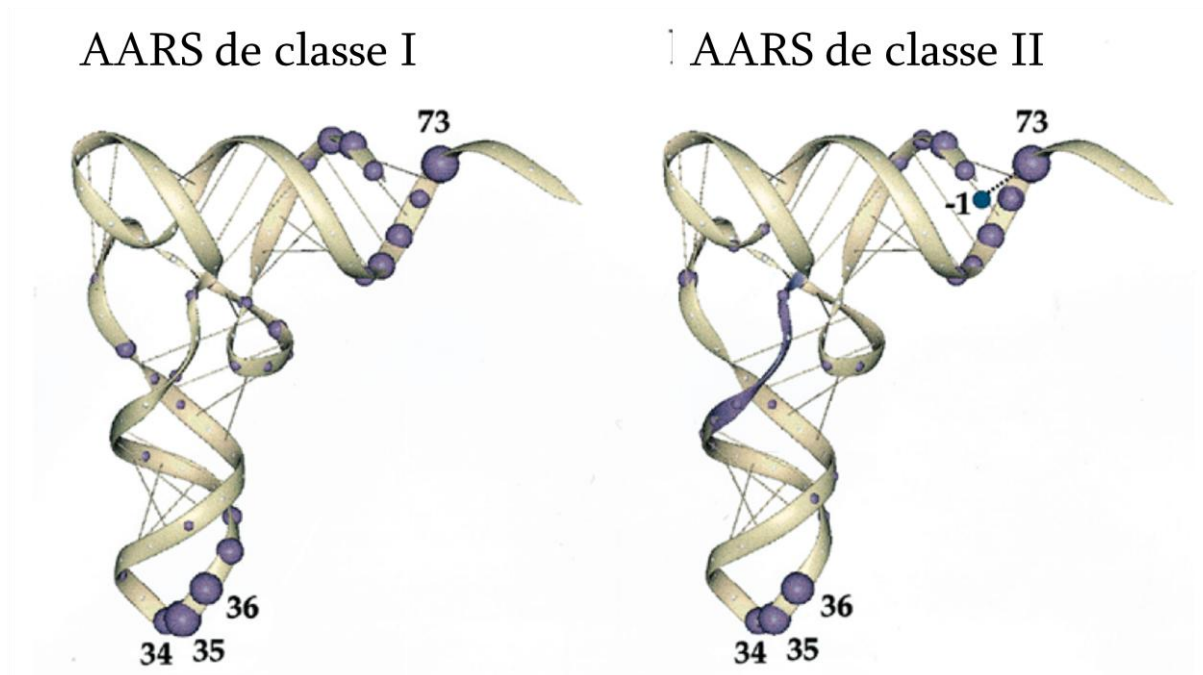


Figure 10 : Distribution des éléments d'identité de l'ARNt. (Giege *et al.* , 1998).

La taille de la sphère est proportionnelle à la fréquence d'utilisation de la position comme déterminant.

La boucle variable, située entre les tiges anticodon et TΨC, peut être très longue. Dans ce cas elle devient un élément d'identité majeur comme dans le cas de l'ARNt^{Tyr} et de l'ARNt^{Ser} (Figure 10).

II. 3. 1. 3. 2 Reconnaissance des déterminants par les AARS

Les AARS de différentes espèces reconnaissent en général les mêmes déterminants sur les ARNt. La spécificité est donc arrivée très tôt au cours de l'évolution. On peut malgré tout noter quelques exceptions comme l'ARNt^{Tyr} de *E. coli* qui est différent de celui de *S. cerevisiae* (Wang and Schultz, 2005; Bonnefond, 2005). Les différences inter-espèces peuvent être exploitées pour la reprogrammation du code génétique, comme cela sera décrit dans la partie IV. 2. 2.

Les membres de chaque sous-classe d'AARS reconnaissent souvent les mêmes déterminants majeurs sur leurs ARNt respectifs. Cette reconnaissance s'effectue grâce à des modules structuraux conservés. Chez les AARS de classe Ia, le module en fagot d'hélices permet de reconnaître l'anticodon sauf dans le cas de la LeuRS qui interagit avec le bras accepteur ainsi que les boucles T et D (PDB : 4CQN). Chez les AARS de classe Ib, l'anticodon et le domaine variable sont deux déterminants majeurs reconnus par trois modules distincts. Chez les AARS de classe Ic, en fonction de l'organisme, un ou deux mo-

dules sont présents pour reconnaître l'ARNt. L'un d'eux toujours présent reconnaît l'anticodon. L'autre module reconnaît la boucle variable mais n'est pas toujours présent. Chez les AARS de classe II, il y a de nombreux cas particuliers. La SerRS, contrairement aux autres AARS de la classe IIa, ne reconnaît pas l'anticodon mais reconnaît le bras variable allongé (PDB : 3W3S (Itoh *et al.* , 2013)). La ThrRS a un domaine supplémentaire de reconnaissance de l'ARNt comparé aux autres AARS de la classe IIa (PDB : 1QG6 (Ward *et al.* , 1999)). La PheRS possède deux domaines de reconnaissances de l'ARNt, elle reconnaît l'anticodon et les boucles T et D (PDB : 2IY5 (Moor *et al.* , 2006)).

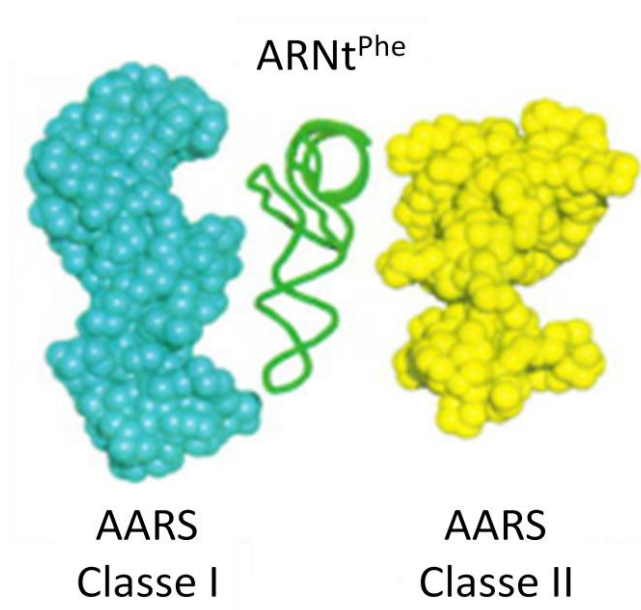


Figure 11 : Symétrie de l'interaction des AARS avec l'ARNt. A gauche une AARS de classe I et à droite une AARS de classe II (Arnez and Moras, 1997)

Il est intéressant de noter que l'ARNt est reconnu par les AARS de classe I et II d'un côté différent, ce qui pourrait justifier que l'acide aminé soit attaché en 3' ou en 2' du ribose en fonction de la classe de l'AARS, comme on peut le voir sur la figure ci-dessus (Arnez and Moras, 1997).

II. 3. 1. 4 L'édition chez les AARS

Malgré la spécificité des AARS pour les acides aminés, certaines AARS peuvent aminoacyler des acides aminés erronés. Bien évidemment, vu la variété chimique des chaînes latérales des acides aminés, toutes les AARS ne sont pas concernées comme par exemple les AARS de classe Ib et Ic. Il a été établi assez tôt qu'une forme d'édition était indispensable pour les acides aminés apolaires et les acides aminés plus petits que l'acide aminé spéci-

fique et ne pouvant donc pas être exclus sur des critères de taille (Pauling, 1957). Ces mécanismes d'édition sont apparus à cause de la pression évolutive née de la toxicité liée à l'incorporation d'acides aminés inappropriés. La liste des AARS capable d'activer et/ou charger de mauvais acides aminés a été regroupée dans le tableau 2 (Giegé and Springer, 2016).

Tableau 2 : Correction d'erreurs chez les AARS (Jakubowski and Goldman, 1992; Jakubowski, 1999; Mascarenhas *et al.* , 2009; Perona and Gruic-Sovulj, 2013)

AARS	Acides aminés activés	Domaine d'édition	Edition pré-transfert	Edition post-transfert	Trans-édition
Classe I					
IleRS	V, T, C	CP1	+	+	
ValRS	I, T, C, S, A	CP1	+	+	
LeuRS	I, M, C, D, N	CP1	+	+	
MetRS	Non protéogène (homocystéine)	Site actif	+	-	
Classe II					
SerRS	T, C	Site actif	+	-	
ThrRS	S, V	Sous domaine	-	+	+
ProRS	A, C	Domaine INS	+	+	+
LysRS	M, L, C, A, T	Site actif	+	-	
AlaRS	G, S, C	Domain en C-ter	+	+	+
PheRS	Y, I, L, M	Domaine B3/B4	+	+	

Pour corriger ces erreurs, il existe des mécanismes d'édition chez chacune de ces AARS (tableau 2). Ces mécanismes impliquent presque à chaque fois un module rattaché, à l'AARS. Ces procédés d'édition sont catégorisés en trois groupes : l'édition pré-transfert, l'édition post-transfert et l'édition *trans* (Figure 12).

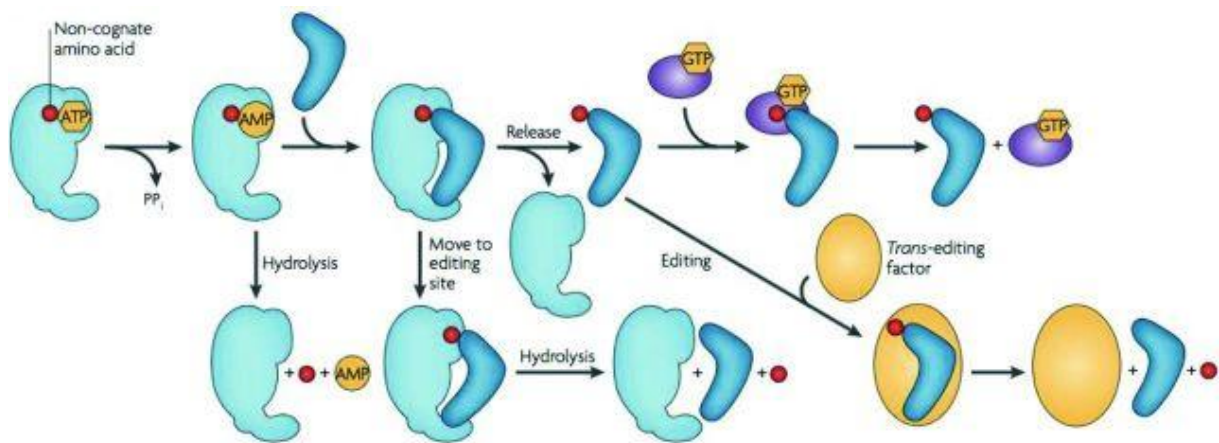


Figure 12 : Schéma représentant les différentes méthodes d'édition chez les AARS (Reynolds *et al.* , 2010)

L'édition pré-transfert se fait après l'activation de l'acide aminé par l'ATP. Trois modèles ont été proposés (Ling *et al.* , 2009) : un premier qui consiste en un changement de site de l'adenylate comme pour l'IleRS (Nomanbhoy *et al.* , 1999; Hendrickson *et al.* , 2002), un deuxième qui consiste en une cyclisation de l'acide aminé puis son relargage dans le milieu (Jakubowski, 2011) et le troisième consiste en une hydrolyse spontanée de l'adenylate comme pour la SerRS (Cusack *et al.* , 1990). L'édition post-transfert se fait après la charge de l'acide aminé sur l'ARNt comme pour la ValRS où une fois l'acide aminé chargé, celui-ci est déplacé dans une seconde poche qui permet de faire une sélection stérique (Fersht and Dingwall, 1979a). Pour l'édition *trans*, l'ARNt aminoacylé peut être édité hors de l'AARS.

Chez les AARS de classe Ia, le système d'édition est toujours le même. Il s'agit d'un double tamisage. Les AARS vont dans un premier temps activer et éventuellement charger sur les ARNt moins efficacement, des acides aminés plus petits que l'acide aminé correct. Dans un second temps, les acides aminés plus petits que le substrat sont amenés à une seconde poche de reconnaissance et, s'ils y entrent, y seront hydrolysés. Le cas de la MetRS est particulier, puisque la reconnaissance du soufre permet d'assurer un taux de discrimination élevé vis-à-vis des autres acides aminés canoniques. Toutefois, il existe un métabolite intermédiaire, l'homocystéine, qui peut être activé par la MetRS. En revanche, après activation, l'homocystéyl-adénylate est efficacement cyclisé en thiolactone, et finalement édité (Jakubowski and Fersht, 1981). Enfin, bien qu'EF-Tu et plus généralement le ribosome, soient capables d'utiliser des ARNt incorrectement aminoacylés (Chapeville *et al.* , 1962),

il est remarquable qu'EF-Tu soit capable d'exercer une certaine discrimination. Il a en effet été observé que des ARNt incorrectement aminoacylés ont une affinité anormalement faible ou anormalement forte pour EF-Tu, ce qui conduit à une moins bonne traductibilité (LaRiviere *et al.* , 2001). Ce phénomène est dû à une compensation thermodynamique des affinités pour le facteur de la partie ARNt et de la partie acide aminé.

III Méthionyl-ARNt synthétase

La méthionyl-ARNt synthétase se distingue des autres aminoacyl ARNt synthétases car elle est responsable de l'aminoacylation de deux ARNt^{Met} très différents : l'initiateur et l'élongateur. Ils ont toutefois le même anticodon. En fonction de l'organisme dont elle provient, cette AARS est monomérique ou dimérique. Toutefois, dans les bactéries où elle est dimérique, on peut par troncature C-terminale obtenir une version d'environ 550 résidus qui a conservé toute son activité enzymatique. La version tronquée de la MetRS d'*E. coli* a été très étudiée d'abord dans les années 1970 en utilisant des méthodes biochimiques et enzymologiques (Cassio *et al.* , 1967; Blanquet *et al.* , 1972; Blanquet *et al.* , 1974; Blanquet *et al.* , 1975b; Blanquet *et al.* , 1976). Ensuite, de nombreux variants de la MetRS et des ARNt^{met} ont été analysés en utilisant différentes techniques biochimiques et biophysiques. Le mécanisme réactionnel a pu être interprété sur ces bases ainsi que sur celles des diverses structures 3D qui ont été résolues sous forme Apo et Holo (Mechulam *et al.* , 1999; Serre *et al.* , 2001; Crepin *et al.* , 2003; Schmitt *et al.* , 2009). Il a fallu attendre 2005 pour la publication d'une structure de la MetRS complexée à l'ARNt qui a permis de visualiser la reconnaissance de l'ARNt en particulier de son anticodon (Nakanishi *et al.* , 2005). Cette structure ne permet toutefois pas de visualiser l'interaction de l'extrémité acceptrice 3' de l'ARNt avec l'enzyme.

III. 1 Structure de la MetRS

Les premières études de la méthionyl-ARNt synthétase d'*E. coli* ont été effectuées dans les années 60, à la suite de la découverte du rôle de la méthionine dans le démarrage de la traduction (Waller, 1963; Cassio *et al.* , 1967; Heinrikson and Hartley, 1967; Cassio, 1968; Bruton and Hartley, 1970). Une version monomérique complètement active de l'enzyme

avait été obtenue par protéolyse ménagée du dimère présent chez *E. coli* (2*676AA). Cette version monomérique est composée de 551 résidus et correspond à l'enzyme native dont la partie C-terminale a été retirée. En 1971, ce monomère a été cristallisé par l'équipe de Waller (Cassio and Waller, 1971). Toutefois, la structure tridimensionnelle n'a été résolue que bien plus tard. De nos jours les structures 3D de MetRS de nombreux organismes ont été résolues (*Trypanosoma brucei*, *Escherichia coli*, *Homo sapiens*, *Thermus thermophilus*, *Aquifex aeolicus*, *Mycobacterium tuberculosis*, *Staphylococcus aureus*, *Brucella suis*, *Brucella abortus*, *Pyrococcus abyssi*, *Mycobacterium smegmatis*, *Leishmania major*, *Saccharomyces cerevisiae*, *Acinetobacter baumannii*).

Dans la structure 3D de la MetRS d'*E. coli*, représentée sur la figure ci-dessous, on retrouve les quatre domaines spécifiques des AARS de classe Ia (Figure 13).

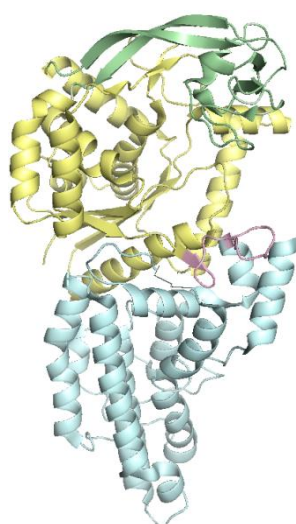


Figure 13 : Structure 3D d'un variant monomérique de la MetRS d'*E. coli* avec le repliement de Rossmann en jaune, le CP (connective peptide) en vert, le domaine KMSKS en rose ainsi que le domaine de reconnaissance de l'ARNt en bleu (3H9C).

III. 1. 1 Les différents modules de la MetRS

Les alignements structuraux et de séquences ont permis de trier les MetRS et de les répartir dans différentes catégories en fonction de l'organisation structurale du domaine CP1. Le domaine CP1 est constitué de zinc « knuckles ». A quelques rares exceptions près (*Hodgkinia cicadicola*), chaque MetRS possède au moins une structure de type « Zinc knuckles », mais le nombre de zincs effectivement liés peut être variable, de 0 à 2.

Cette répartition donne lieu à quatre familles, la première à deux «knuckles» et deux zincs, la deuxième à deux «knuckles» mais un seul est occupé par un ion, la troisième à un « knucle» occupé par un zinc et la quatrième à un « knucle» sans zinc. Cette classification, basée sur l'analyse de la structure de la MetRS d'*E. coli* combinée à des alignements de séquence, a été confirmée par la détermination de structure 3D pour chacun des membres (Mechulam *et al.* , 1999). On trouve en effet dans la famille 1 la MetRS de *P. abyssi* (Crepin *et al.* , 2004), celle de *T. thermophilus* (Sugiura *et al.* , 2000) dans la famille 3 et celle de *Mycobacterium tuberculosis* (Barros-Álvarez *et al.* , 2018) dans la famille quatre. Chez *E. coli* l'interaction avec le zinc est assurée par quatre cystéines dans le domaine CP (Fourmy *et al.* , 1993a; Fourmy *et al.* , 1993b). Il existe un second motif de type zinc knucle qui comprend les séquences ¹²⁹DPEK₁₃₂ et ¹⁷⁵SVVS₁₇₈. Ce repliement est stabilisé non pas par un zinc mais par des liaisons hydrogène entre l'aspartate et les sérines.

III. 1. 2 Rôle de l'ion Zinc

Afin de comprendre le rôle du site zinc, des études de mutations dirigées ont été effectuées sur la version monomérique de la MetRS d'*E. coli*. Il a été montré que le domaine CP1 a un rôle dans la réaction d'activation ainsi que dans la charge de l'acide aminé sur l'ARNt. Les mutations sur les résidus 138 et 139 influent sur la vitesse de formation du méthionyl-adénylate alors que les mutations de la position 135 influent sur la vitesse d'aminacylation (Fourmy *et al.* , 1995). Ces trois résidus sont plutôt bien conservés dans les MetRS qui ont deux domaines et un atome de zinc comme la MetRS d'*E. coli*.

L'étude des MetRS de *T. thermophilus* (Nureki *et al.* , 1993) et *B. stearothermophilus* (Schmitt *et al.* , 1997) montrent d'une part que le domaine qui lie le zinc est assez éloigné du site actif et d'autre part que les domaines CP chez *E. coli* et *T. thermophilus* sont orientés de manière différente par rapport au repliement de Rossmann. Une telle conformation suggère que le domaine est assez flexible pour qu'il interagisse avec le site actif. L'ensemble des données disponibles, et en particulier l'obtention de variants dans lesquels le zinc est faiblement lié, suggère que l'ion zinc lui-même a un rôle essentiellement structural (Fourmy *et al.* , 1993b).

Au-delà des quatre domaines cœur de la MetRS (Repliement de Rossmann, CP1, KMSKS et domaine de reconnaissance de l'anticodon) certaines MetRS possèdent des modules complémentaires (Figure 14).

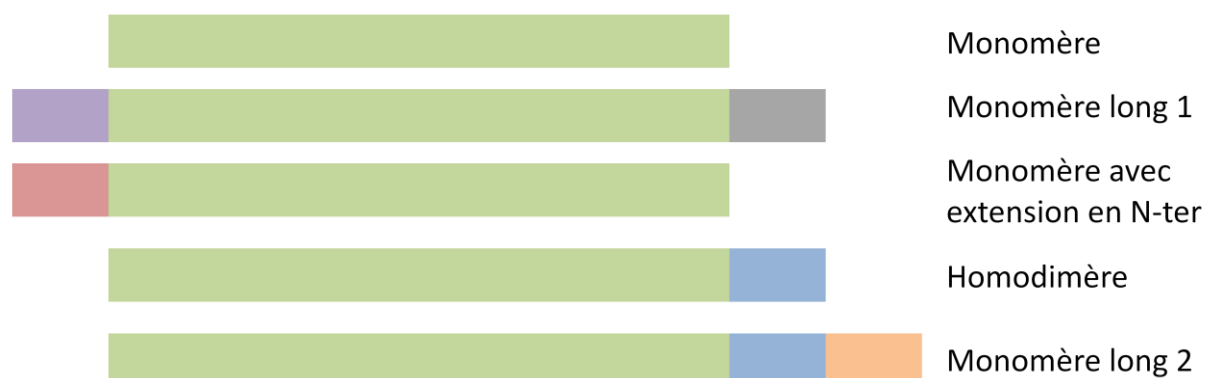


Figure 14 : Organisation modulaire de la MetRS. Le corps de l'enzyme est représenté en vert, le domaine C-terminal responsable de la dimérisation est représenté en bleu, les domaines possibles en N-terminal sont représentés en violet ou en rouge et le domaine situé à l'extrémité C-terminal de l'enzyme de type hélice-tour-hélice est représenté en orange.

Comme le montre la figure précédente, le module additionnel le plus fréquent est positionné en C-terminal. Il est généralement associé à la dimérisation de la MetRS à quelques exceptions près. Chez la MetRS humaine, le domaine C terminal prend une structure particulière (Cahuzac *et al.*, 2000) qui diminue l'activité *in vitro* mais qui permettrait *in vivo* de créer une augmentation locale de la concentration en ARNt aminoacylé en constituant des rassemblements d'enzymes (Kaminska *et al.*, 2001)

III. 1. 3 Domaine de dimérisation

La séquence du domaine de dimérisation s'est révélée proche de celle d'une protéine, Trbp111, présente chez plusieurs organismes dont *Aquifex aeolicus*. Cette petite protéine qui a la même taille que l'extension C-terminale de la MetRS d'*E. coli*, et est aussi dimérique (Morales, 1999). Il a été montré que cette protéine pouvait se lier efficacement aux ARNts. Trbp 111 a des orthologues dans différents organismes chez les bactéries et les archées. Sa structure ainsi que celle de son orthologue chez *E. coli* et *T. thermophilus*, CsaA, (Swairjo, 2000; Kawaguchi, 2001) ont été résolues. Ce sont des dimères symétriques formant des repliements de type OB-fold connectés au milieu par un domaine central de dimérisation. Le site de dimérisation a été cartographié par mutagenèse. Grâce à un modèle d'amarrage, il a été proposé que Trbp 111 se liait à la boucle TΨC (Morales, 1999). Ce modèle a été confirmé par des expériences d'empreintes (Nomanbhoy *et al.*, 2001). De plus,

un complexe IleRS-ARNt^{ile}-Trbp111 a été détecté *in vitro*. Cette affinité de Trbp 111 pour l'ARNt se retrouve de façon intéressante sur le module de dimérisation de la MetRS d'*E. coli*. En effet, le dimère a une affinité plus grande pour l'ARNt^{met} que la version monomérique sans le module de dimérisation (Blanquet *et al.* , 1973; Blanquet *et al.* , 1976; Gale *et al.* , 1996). Il faut toutefois noter que la liaison de deux ARNt sur la MetRS dimérique se fait de façon anti-coopérative, ce qui pourrait faciliter le relargage des ARNt aminoacylés (Blanquet *et al.* , 1976).

III. 1. 4 Réaction d'activation de la méthionine

Les paramètres thermodynamiques relatifs de la réaction d'activation de la méthionine ont été déterminés à l'équilibre et à l'état pré-stationnaire. La liaison de certains substrats (méthionine et ARNt), la formation du méthionyl-adénylate ainsi que sa pyrophosphorolyse s'accompagnent d'une variation de fluorescence intrinsèque (Blanquet *et al.* 1972). Ces variations, reflètent la formation des différents complexes. C'est grâce à ces variations que l'ensemble des paramètres de la réaction a pu être mesuré à l'équilibre et à l'état pré-stationnaire (Blanquet *et al.* , 1972; Hyafil *et al.* , 1976) (Figure 15).

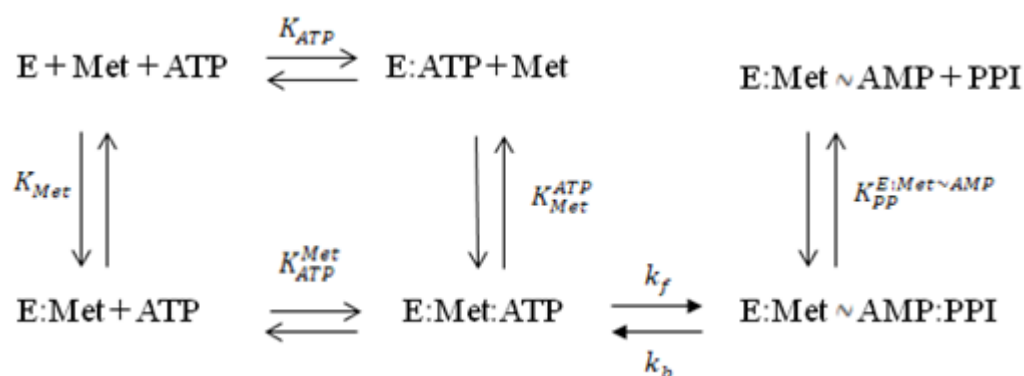


Figure 15 : Etapes élémentaires de l'activation de la méthionine par l'ATP

D'un point de vue mécanistique, un des résultats importants est la mise en évidence de couplage synergique entre le site de liaison de la méthionine et celui de l'ATP. Le couplage se traduit par un gain d'affinité pour l'un des ligands quand l'autre est déjà lié (Blanquet *et al.* , 1975b; Fayat *et al.* , 1977). Il est toutefois compensé par la répulsion électrostatique entre le carboxylate de la Met et le phosphoryle α de l'ATP. Ce phénomène a été mis en évidence grâce à l'étude de l'interaction entre l'enzyme et des analogues de la méthionine et de l'ATP. Une altération de ce couplage a été observée lors de mutations dans les

motifs HIGH et KMSKS (Mechulam *et al.* , 1991; Schmitt *et al.* , 1994; Fourmy *et al.* , 1995).

Ces deux motifs jouent un rôle clé dans la catalyse de la réaction d'activation. Il a été montré en mutant les glycines flanquant le motif KMSKS que cette région est mobile et que cette mobilité est importante pour la réaction d'activation (Schmitt *et al.* , 1994). Les mutations des lysines du motif KMSKS ont montré qu'elles avaient un rôle crucial dans la stabilisation de l'état de transition précédant la formation du methionyl adenylate. Leur fonction est assurée via leur interaction avec la partie pyrophosphate de l'ATP (Mechulam *et al.* , 1991; Schmitt *et al.* , 1994).

III. 2 Reconnaissance des substrats

Plusieurs MetRS ont été cocrystallisées avec un ou plusieurs des substrats, ce qui a permis de mieux comprendre les interactions impliquées.

III. 2. 1 La liaison de la méthionine

La MetRS d'*E. coli* a été cocrystallisée avec de la méthionine, par micro-ensemencement avec des cristaux de MetRS apo. La structure de ce complexe a été affinée à 1,8 Å de résolution (Serre *et al.* , 2001). Cette structure a révélé une poche hydrophobe qui est délimitée par les résidus A12, L13, Y15, W253, A256, P257, Y260, I297, H301, W305 et la boucle ²⁹⁴GKD²⁹⁶. Dans le cas de l'enzyme Apo, il y a une molécule d'eau maintenue par la L13 et la Y260. Lorsque la méthionine vient se positionner dans la poche, elle chasse la molécule d'eau et l'atome de soufre vient prendre la place de celle-ci. Celui-ci vient former des liaisons hydrogène avec le groupement -OH en *para* de l'Y260 ainsi qu'avec l'amine de la chaîne principale de la L13. L'amine de la méthionine interagit avec la chaîne latérale de D52 ainsi qu'avec l'oxygène de la chaîne principale de la L13. Le carboxylate fait une interaction indirecte avec l'H24 du motif HIGH via une molécule d'eau.

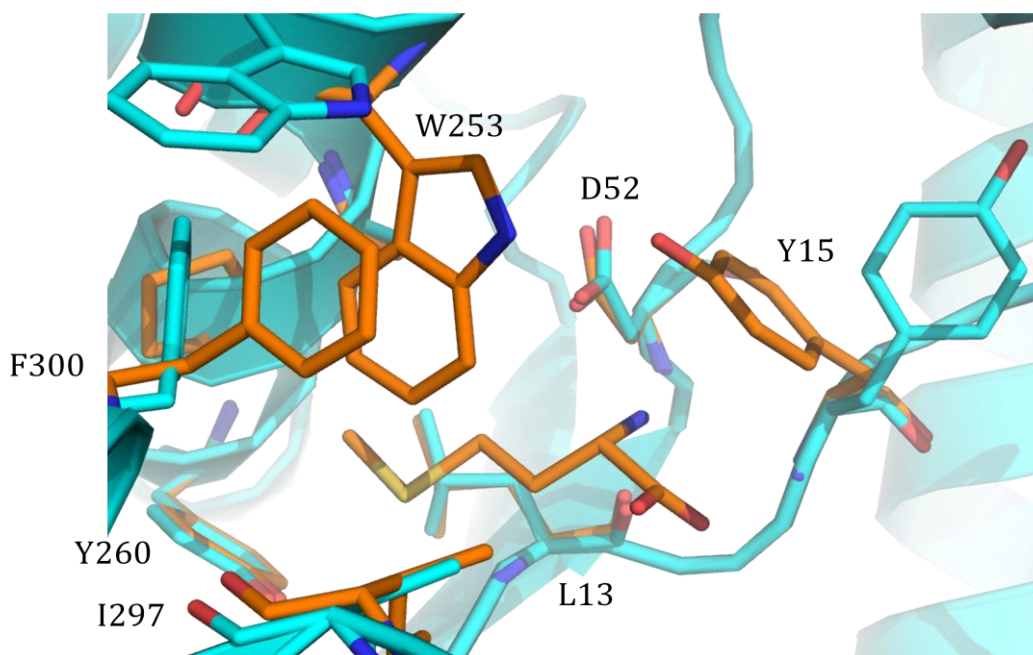


Figure 16 : Site de reconnaissance de la méthionine de la MetRS d'*E. coli*.

La structure de la MetRS de *E. coli* apo (cyan) est superposée à la structure du complexe MetRS:Met (orange). Les acides aminés importants sont représentés en bâton.

L'arrivée de la méthionine provoque le basculement de plusieurs chaînes latérales et ainsi la diminution du volume de la poche (Figure 16). D'un côté on observe le basculement du couple W253-F300 et du couple F304-W229. Ce réarrangement permet à la chaîne latérale de la Q225 de basculer. De l'autre côté on observe le basculement de la Y15 qui entraîne les mouvements en série de l'N17 et de la L60. Le groupement -OH de la Y15 stabilise tous ces changements de conformation en formant une liaison hydrogène avec l'amine de la chaîne latérale du W253. Le changement de conformation des aromatiques pourrait expliquer la variation de fluorescence observée lors de la formation du complexe enzyme:méthionine (Blanquet *et al.*, 1972). Ainsi, la poche de liaison de la chaîne latérale ne se ferme de manière stable que lors de la liaison du substrat. La liaison de la méthionine à la MetRS implique donc un mécanisme d'ajustement induit ou «induced fit »

III. 2. 2 La liaison du méthionyl-adénylate

La MetRS d'*E. coli* a été également cocrystallisée en présence d'un analogue stable du méthionyl-adénylate (Crepin *et al.*, 2002). La partie méthionine de celui-ci se positionne exactement comme la méthionine. Le ribose interagit avec le E27 ainsi qu'avec le G293 et

l'D296. L'adénine interagit avec la chaîne principale des résidus 24 et 25 comme évoqué dans la partie II. 3. 1. 2. 1. 1 (Figure 17).

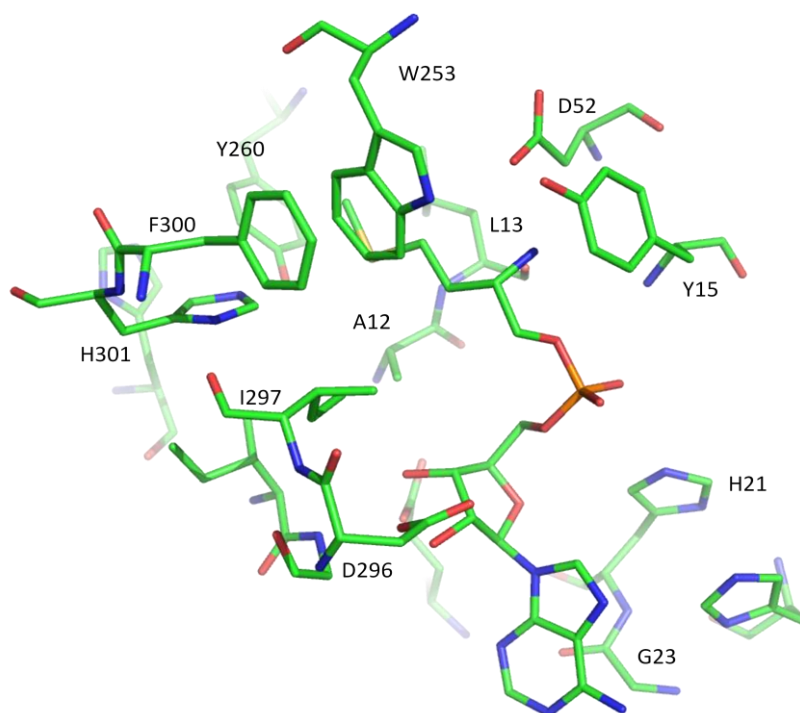


Figure 17 : Structure de la MetRS d'*E. coli* liée au méthioninyl-adénylate (analogue du méthionyl-adénylate, PDB : 1PG0)

III. 2. 3 La fixation de l'ATP

Aucune structure de MetRS complexée uniquement à l'ATP n'est disponible. Egalement, aucune structure de MetRS n'a été obtenue en présence de PPi ou de méthionyl-adénylate. En revanche, les structures de la MetRS de *S. aureus* liée à l'ATP-Mg²⁺ et à des inhibiteurs (4QRE et 4QRD) ainsi que celle de la MetRS de *Leishmania major* liée au méthionyl adénylate et au PPi-Mg²⁺ (3KFL ;(Larson *et al.* , 2011)) ont été déterminées. Dans ces structures, en plus des interactions citées pour la méthionine et l'AMP, on observe les interactions avec les groupements phosphates β et γ ainsi qu'avec un ion magnésium pontant les deux phosphates. On observe le changement de conformation du motif KMSKS prédit par les résultats biochimiques. Les lysines viennent interagir avec les phosphates et le Mg²⁺. Cette observation structurale renforce le modèle qui propose que l'interaction des lysines avec la partie pyrophosphate permet de dissiper la charge du phosphate α et de faciliter son attaque par le carboxylate de la Met (Figure 18).

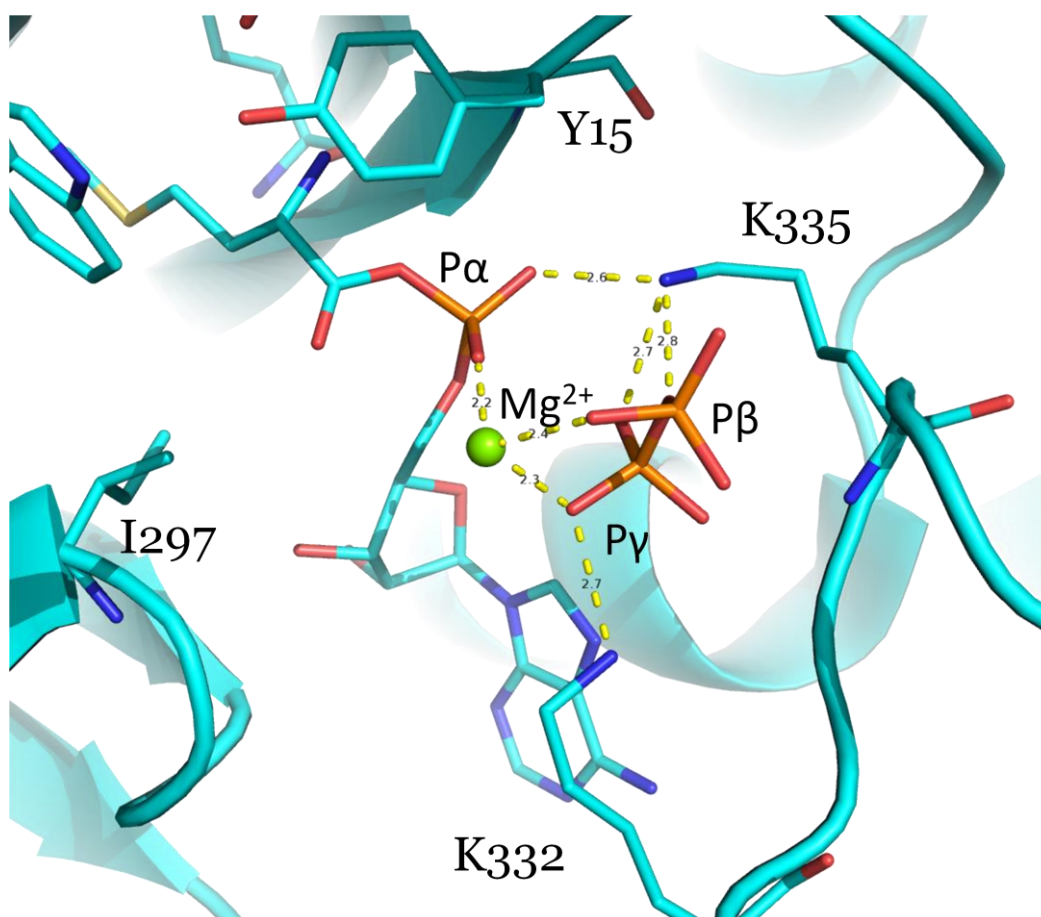


Figure 18 : Structure du complexe MetRS:méthionyl adénylate:Ppi:Mg²⁺ de *Leishmania major* (PDB : 3KFL). Cette figure illustre les interactions entre les phosphates, le magnésium et les lysines (332 et 335). Ces dernières permettent la délocalisation des charges négatives des phosphates.

III. 2. 4 Interaction avec l'ARNt

Les interactions entre les ARNt^{Met} et la MetRS d'*E. coli* ont été beaucoup étudiées par mutagenèse dirigée (Ghosh *et al.* , 1990; Ghosh *et al.* , 1991; Meinnel *et al.* , 1991a; Meinnel *et al.* , 1991b; Schmitt *et al.* , 1993; Kim *et al.* , 1993) et marquage d'affinité (Hountondji *et al.* , 1985; Leon and Schulman, 1987). Ces études ont permis de trouver des régions spécifiques dans le domaine C-ter qui interagissent avec l'ARNt et qui sont responsables de la spécificité de l'enzyme pour l'anticodon CAU. Ces deux régions correspondent aux résidus 380-395 et 450-465. Certains résidus sont spécifiquement en interaction avec l'ARNt, comme P460, W461, N452, N391 et R495 qui permettent la discrimination positive d'ARNt^{Met} portant l'anticodon CAU. Les résidus D449 et D456 permettent une discrimination négative. La suppression de ces charges négatives permet d'aminoacyler des ARNt^{Met} portant un anticodon non CAU (Schmitt *et al.* , 1993). Tous ces résidus sont portés par les hélices α9 et α11, qui interagissent par des liaisons électrostatiques (Figure 19).

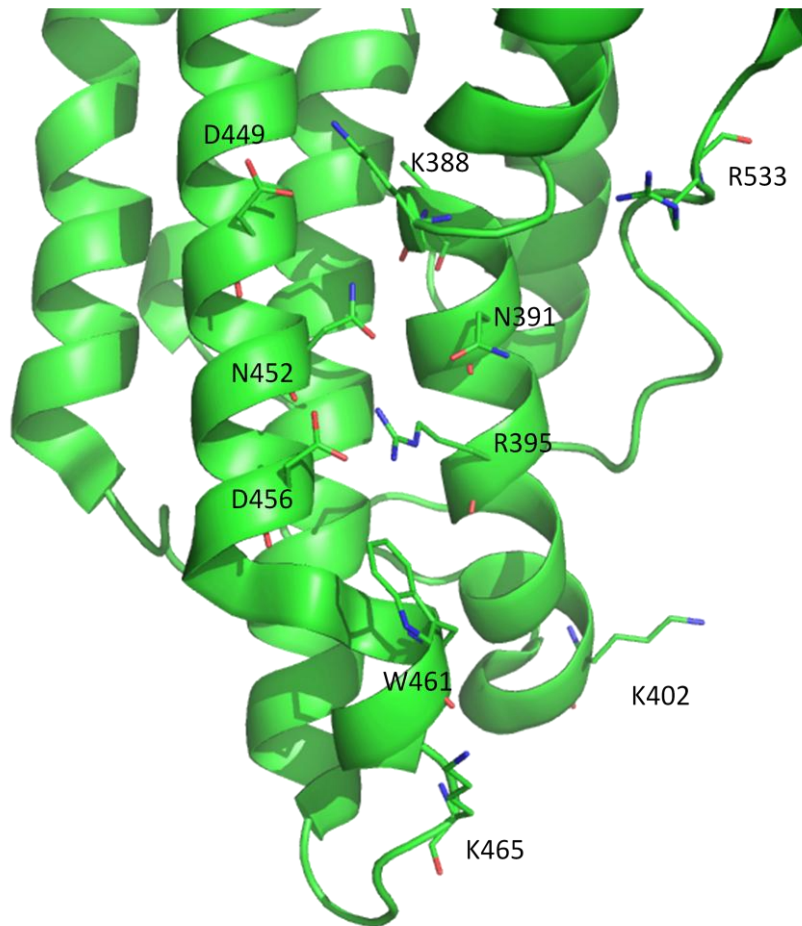


Figure 19 : Vue du domaine C-terminal en hélice α de la MetRS d'*E. coli*. Les principaux résidus impliqués dans la liaison de l'anticodon CAU des ARNt^{Met} sont représentés en bâtons.

D'autres acides aminés qui ne sont pas dans ces régions interagissent avec l'ARNt. Une mutation de R533, qui est à 20 Å du W461 réduit l'affinité pour la tige anticodon de l'ARNt^{met} (Gale *et al.*, 1996). La structure de la MetRS d'*E. coli* complexée à l'ARNt n'a pas été résolue mais celle de ce complexe d'*Aquifex aeolicus* l'a été (Figure 20).

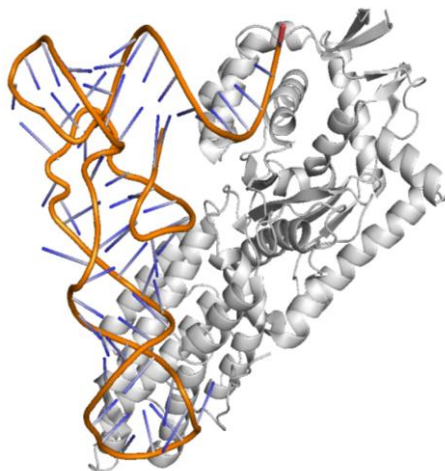


Figure 20 : Structure de la MetRS monomérique d'*Aquifex aeolicus* complexée avec l'ARNt^{Met} (PDB : 2CT8). Dans cette structure l'extrémité CCA de l'ARNt n'est pas visible.

Cette structure confirme par homologie le rôle des résidus cités précédemment et permet de voir leur mode d'interaction avec les nucléotides (Nakanishi *et al.* , 2005). Le modèle permet de visualiser les détails de la reconnaissance de l'anticodon par l'enzyme en particulier pour la base centrale (U36). En effet, on observe une poche de reconnaissance spécifique au niveau de l'hélice 349-365 (390-405 chez *E. coli*) et le début de l'extension C-terminale. L'uracile 36 forme une série de liaisons hydrogènes avec la chaîne principale de F492 (Y530 chez *E. coli*) et de L494 (R533 chez *E. coli*, (Nakanishi *et al.* , 2005); 2CSX et 2CT8). De plus, le modèle met en évidence l'interaction entre les nucléotides 60 à 72 et les résidus des zones 185-198 (216-225 chez *E. coli*) et 258-262 (295-299 chez *E. coli*). En particulier R262 (Y299 chez *E. coli*) interagit avec le phosphate de C72 et Q98 (Ser 224 chez *E. coli*) interagit avec le ribose de C72. De plus, le ribose de G69 interagit avec R189 (pas d'équivalent chez *E. coli*) et E192 (L217 chez *E. coli*). Chez *E. coli*, bien que les acides aminés correspondant à ces positions soient aussi chargés, ils ne sont pas conservés. Il est donc possible que le bras accepteur se positionne de façon différente. Enfin, il faut souligner que la structure de ce complexe a été résolue avec et sans analogue du méthionyl-adénylate. En plus des changements de conformation observés dans les autres structures, le tryptophane 288 (Y325 chez *E. coli*) bascule et vient interagir avec l'adénine du méthionyl-adénylate.

III. 3 Misacylation et édition

Il a été montré assez tôt que les MetRS d'*E. coli* ou de *B. stearothermophilus* étaient capables d'aminocyler des acides aminés erronés, tels que la norleucine et l'homocysteine, mais environ 1000 fois moins efficacement (Fersht and Dingwall, 1979a; Jakubowski and Goldman, 1992).

La norleucine est un intermédiaire dans la voie de synthèse de certains acides aminés (Kisumi *et al.* , 1976a; Kisumi *et al.* , 1976b). Aucun mécanisme de correction n'a été mis en évidence. Ainsi il semble qu'il n'y ait pas eu de pression évolutive pour créer un système d'édition efficace pour empêcher la charge de la norleucine. Cela peut être dû au fait que la norleucine remplace suffisamment bien la méthionine pour que les enzymes restent actives (Jakubowski and Goldman, 1992).

L'homocystéine est un précurseur de la méthionine. De plus, l'homocystéine est présente dans une autre voie métabolique chez les mammifères, comme produit intermédiaire d'une voie de trans-sulfuration qui permet de convertir la méthionine en cystéine. Les vitesses d'activation de cet acide aminé sont plus faibles avec les MetRS d'*E. coli* et *B. stearothermophilus* mais ne sont que deux à trois fois inférieures à celles mesurées avec la méthionine (Kim et al. 1993; Fersht et Dingwall 1979; H. Jakubowski et Fersht 1981). L'homocystéine est éditée par un mécanisme pré-transfert.

Le mécanisme illustré sur la figure 21 se résume par, la transformation de l'homocystéinyl adénylate réagissant sur lui-même pour former un thiolactone libérant de l'AMP et du PPi. Il a été montré chez les MetRS de *E. coli*, de *B. stearothermophilus* (Jakubowski and Fersht, 1981), de *S. cerevisiae* (mitochondriale et cytoplasmique (Jakubowski, 1991; Senger et al. , 2001)).

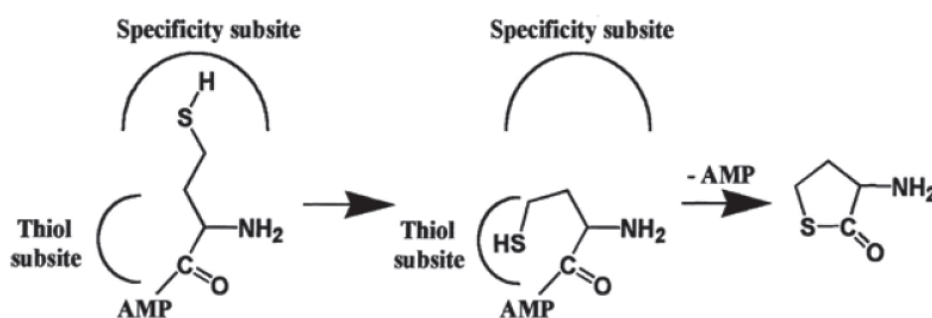


Figure 21 : Mécanisme de cyclisation de l'homocystéine (Jakubowski, 2011)

De façon intéressante, la méthionine peut elle-même être éditée selon un mécanisme similaire à celui de l'édition de l'homocystéine mais cette fois avec un départ du groupement méthyle. La réaction de cyclisation avec la méthionine est toutefois 50 fois plus lente que celle mesurée avec l'homocystéine (Kim et al. 1993). Des mutations sur les acides aminés Y15 et W305 ont montré une diminution de la spécificité de la cyclisation de l'homocystéine par rapport à la méthionine, ce qui montre l'importance de la reconnaissance du groupement méthyle dans le mécanisme d'édition.

Un autre acide aminé qui peut être utilisé par la MetRS est la Sélénométhionine. La SeMet n'étant pas présente naturellement chez *E. coli* il n'y a aucun mécanisme d'édition. Il n'y en a de toute façon pas besoin car l'incorporation de SeMet n'est pas toxique pour *E. coli* (Frost and Lish, 1975). Cette incorporation de SeMet est d'ailleurs utilisée pour produire

des protéines contenant de la SeMet à la place de la Met, pour le phasage des données de diffraction aux rayons X (Hendrickson *et al.* , 1990; Hendrickson, 1991).

IV Incorporation d'acides aminés non canoniques

IV. 1. Les acides aminés non canoniques

Les acides aminés non canoniques (AANC) peuvent être définis comme des acides aminés qui ne font pas partie du code génétique naturel. Pour cette raison, ils sont aussi appelés acides aminés non protéogènes (Hunt, 1985) ou non naturels (Wang and Schultz, 2005). Toutefois, l'appellation AANC est plus générique car elle comprend les acides aminés présents dans le vivant mais ne participant pas à la synthèse protéique et aussi les acides aminés synthétisés en laboratoire.

L'incorporation d'AANC permet l'incorporation de nouvelles fonctionnalités chimiques dans les protéines et ce avec une grande précision. Il est possible de modifier un seul atome par un fluor, d'ajouter une fonctionnalité chimique en remplaçant un hydrogène par une fonction alcool ou même de supprimer un groupement comme en remplaçant un soufre par un CH₂. Il est donc possible d'ajouter des acides aminés avec des atomes qui ne sont normalement pas présents chez les acides aminés classiques comme du Se à la place du soufre ou des halogènes à la place d'hydrogène (Budisa, 2004; Young and Schultz, 2010). L'ajout de groupements chimiques tels que des groupements nitro ou méthoxy permet d'améliorer les réactions redox (Amir *et al.* , 2013). L'ajout d'un alkyle ou d'un azido permet ensuite de faire réagir cet acide aminé avec une autre molécule (Prescher and Bertozzi, 2005). Cela peut permettre d'immobiliser la protéine, ou d'y fixer spécifiquement un sucre, un oligonucléotide ou un fluorophore (Gautier and Hinner, 2015). Certains AANC sont photosensibles et peuvent être réactifs aux UV, fluorophores ou encore porter des groupements photo réactifs (Neumann, 2012; Spicer and Davis, 2014).

IV. 2. Incorporation *in vivo*

Les protéines contenant des AANC sont appelées alloprotéines ou xéno-protéines. Il existe différentes stratégies pour les produire. *In vivo*, il en existe deux qui dépendent de la

structure des acides aminés à incorporer. Une des méthodes utilise la capacité du système de traduction à se tromper dans l'incorporation des acides aminés. L'autre utilise des paires orthogonales d'ARNt et d'AARS modifiées. *In vitro*, une seule méthode est très largement utilisée.

IV. 2. 1 Méthode utilisant le système de traduction sans modification

Cette méthode consiste à incorporer des AANC en utilisant la perméabilité du système de traduction. En utilisant une souche auxotrophe pour un acide aminé, il devient dans certains cas particuliers, possible de remplacer celui-ci par un acide aminé non canonique.

Il a été montré *in vitro* que le système de traduction était capable d'incorporer de nombreux AANC (Josephson *et al.* , 2005; Hartman *et al.* , 2007). Initialement la perméabilité du système de traduction avait été mise en évidence en incorporant de la SeMet (Cowie and Cohen, 1957; Richmond, 1962). L'incorporation de la SeMet a été utilisée à partir des années 90 (Hendrickson *et al.* , 1990) pour faciliter la résolution du problème des phases en cristallographie. Le potentiel de l'incorporation d'acides aminés non canoniques a été ensuite étudié plus en profondeur par Budisa (Budisa, 2004) et Tirrell (Link *et al.* , 2003). Dans cette méthode, on utilise simplement une souche auxotrophe pour l'acide aminé que l'on veut remplacer. L'AANC à incorporer est alors ajouté comme seule source d'un acide aminé proche. A part le cas classique de la SeMet on peut ainsi incorporer de la Trifluoroisoleucine (Wang *et al.* , 2003) ou des aminotryptophanes (Budisa *et al.* , 2002)

IV. 2. 2 Méthode utilisant un système de traduction modifié

La seconde méthode *in vivo* est beaucoup plus complexe. Elle implique de rendre des AARS spécifiques des AANC, et de les coupler à un ARNt comportant un anticodon réservé. Cette méthode permet d'étendre le code génétique car elle réassigne le codon réservé à un AANC. L'extension du code génétique implique d'ajouter un ARNt et une AARS exogène à un organisme donné.

IV. 2. 2. 1 Paires orthogonales AARS-ARNt

Le codon le plus réassigné est le codon stop Ambre (Liu and Schultz, 2010; Chin, 2017). On utilise des paires AARS-ARNt dites orthogonales. En effet, l'AARS ne doit pas recon-

naitre les ARNt de l'organisme et doit aminoacyler spécifiquement l'ARNt supplémentaire par un AANC (figure 22). Cet ARNt orthogonal ne doit pas être un bon substrat pour les AARS endogènes. La sélectivité des systèmes orthogonaux n'est pas absolue. Les ARNt orthogonaux reconnaissant le codon ambre sont généralement utilisés 10 à 1000 fois plus efficacement que les ARNt compétiteurs endogènes (Zhang *et al.* , 2017a). Des cas où la sélectivité est supérieure à 1000 ont été trouvés et améliorés par évolution dirigée (M. S. Zhang *et al.* 2017; Liu *et al.* Schultz 2010; Ellefson *et al.* 2014). Une paire orthogonale doit être définie en fonction de l'organisme dans lequel elle est destinée à être ajoutée. Par exemple pour *E. coli*, il y a majoritairement deux couples orthogonaux : TyrRS-ARNt_{CUA} de *Methanocaldococcus jannaschii* et PylRS-ARNt_{CUA} de *Methanosarcina mazei*. L'orthogonalité des couples AARS/ARNt est basée sur les différences des déterminants d'identité et de leurs mécanismes de reconnaissance d'un organisme à l'autre (Figure 22). Par exemple, l'ARNt^{Tyr} est reconnu par le bras variable chez *M. jannaschii* alors qu'il est reconnu par l'anticodon chez *E. coli*. Le couple PylRS-ARNt_{CUA} de *M. mazei* est orthogonal à des bactéries, des levures et des cellules de mammifères (Neumann *et al.* , 2008; Chen *et al.* , 2009; Hancock *et al.* , 2010; Gautier *et al.* , 2011). Ceci est dû à la structure spécifique de l'ARNt^{Pyl} et à l'unique mode de reconnaissance de la PylRS (Nozawa *et al.* , 2009; Suzuki *et al.* , 2017) qui reconnaît le bras variable de l'ARNt.

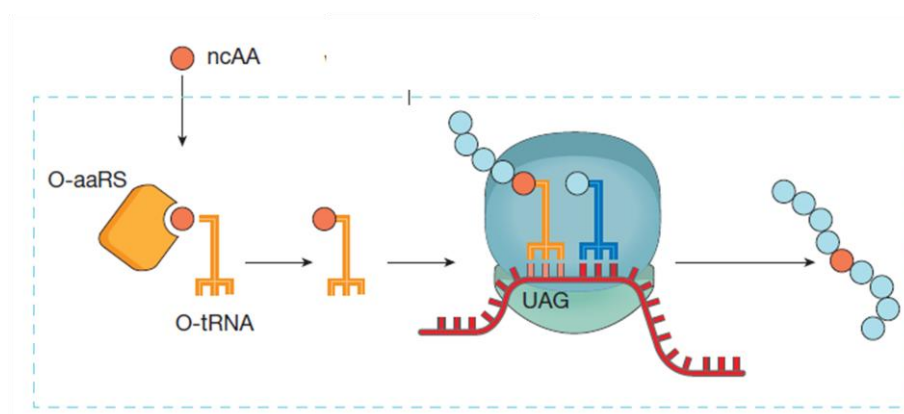


Figure 22 : Système de traduction avec le codon UAG réassigné. Figure tirée de (Chin, 2017).

IV. 2. 2. 2 ARNt et ARNr orthogonaux

Les 64 codons possibles limitent le nombre d'AANC incorporables. En conséquence, une autre stratégie a été développée pour augmenter le nombre d'AANC dans le code gén-

tique. L'utilisation de quadruplets permet d'avoir jusqu'à 256 codons possibles. Un ribosome orthogonal permettant de lire à la fois les codons ambre et les quadruplets a été créé (Anderson *et al.* , 2004; Neumann *et al.* , 2010). Un système incluant 12 paires orthogonales et le ribosome modifié a ainsi été construit. Il a permis d'incorporer des acides aminés capables d'être fonctionnalisés chimiquement. (Wang *et al.* , 2014) (Figure 23).

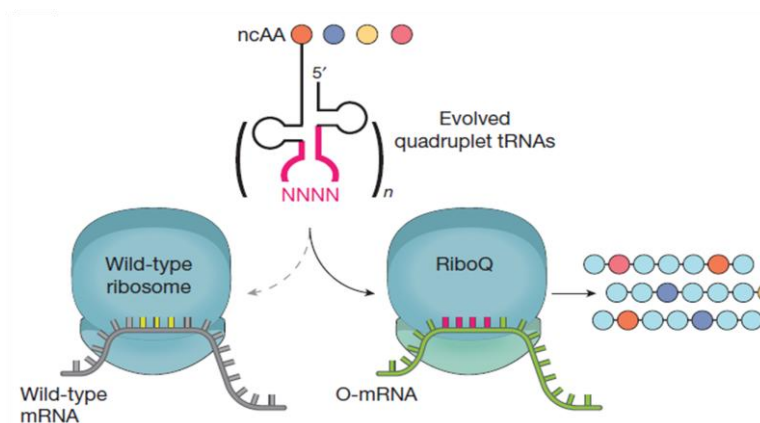


Figure 23 : Système de traduction faisant intervenir un codon quadruplet. Figure tirée de (Chin, 2017)

D'autres approches utilisant des ribosomes orthogonaux ont été utilisées. Récemment, une équipe a mis au point un ribosome où les sous-unités sont attachées, via les ARNr 16S et 23S, afin de séparer la traduction native de la traduction orthogonale. On peut ainsi éviter un mélange de sous-unités entre le système orthogonal et le système résident. Le ribosome orthogonal peut ensuite être adressé à un ARNm particulier en jouant sur les séquences SD et anti-SD. Enfin, ce ribosome a également été rendu résistant à un antibiotique, ce qui permet de maximiser la traduction orthogonale à un moment donné de la croissance bactérienne (Orelle *et al.* , 2015).

Une autre approche pour incorporer des AANC serait de réassigner un codon redondant à un AANC. Cela a été décrit *in vitro* (Iwane *et al.* , 2016) mais pas *in vivo*. *In vitro*, il est possible de n'utiliser que 20 codons élongateurs et le codon initiateur pour effectuer la traduction des 20 acides aminés canoniques. Il reste 41 codons qui peuvent être en principe utilisés pour des AANC.

Plus récemment, l'émergence de systèmes viables comportant des nucléotides non naturels a permis de développer une nouvelle façon d'étendre le code génétique (Zhang *et al.* , 2017b). En effet, en associant l'utilisation d'un nucléotide à une paire AARS-ARNt ortho-

gonale il a été possible d'incorporer *in vivo* de la p-azidophénylalanine dans des peptides à une position spécifique.

IV. 2. 2. 3 Les AARS orthogonales

Les paires orthogonales permettent d'incorporer des AANC mais l'incorporation de ceux-ci est limitée par l'affinité des AARS orthogonales pour les acides aminés non canoniques et la capacité de la paire orthogonale à fonctionner efficacement dans l'organisme hôte. Pour améliorer cette affinité, plusieurs méthodes ont été mises au point. Deux axes importants ont été développés, d'une part la génération de variants et d'autre part l'obtention d'un crible efficace pour trouver les meilleurs variants.

Une méthode d'évolution dirigée pour les AARS-orthogonales (AARS-O) doit tenir compte de cinq paramètres (Varga Rodriguez et al 2018) :

- Le niveau d'expression des paires orthogonales AARS/ARNt
- La concentration intracellulaire en AANC
- L'orthogonalité des AANC, c'est-à-dire que l'AANC ne doit pas être reconnu par le système de traduction résident ni interférer avec le métabolisme hôte.
- Les facteurs de traduction et les mécanismes d'édition
- La sensibilité de la méthode de détection de l'incorporation de l'AANC

Nous décrivons trois méthodes centrales. Les deux premières méthodes nécessitent une étude structurale des AARS pour sélectionner les positions à muter. La troisième en revanche est moins orientée.

La première méthode est spécifique à l'acide aminé incorporé. En effet, elle nécessite que celui-ci ait un groupement azide sur la chaîne latérale (Link *et al.* , 2006). Après avoir sélectionné les positions à muter, une banque de mutants est constituée par mutagenèse de saturation. Le crible consiste à produire une protéine membranaire dans laquelle l'AANC doit être incorporé. Celui-ci est ensuite fonctionnalisé chimiquement pour créer une fluorescence. Les variants adéquats sont alors sélectionnés par un tri cellulaire associant la mesure de fluorescence à la cytométrie en flux (FACS). Cette méthode a permis de trouver des variants de la MetRS plus spécifiques pour l'azidonorleucine que pour la méthionine (Tanrikulu *et al.* , 2009) (Figure 24).

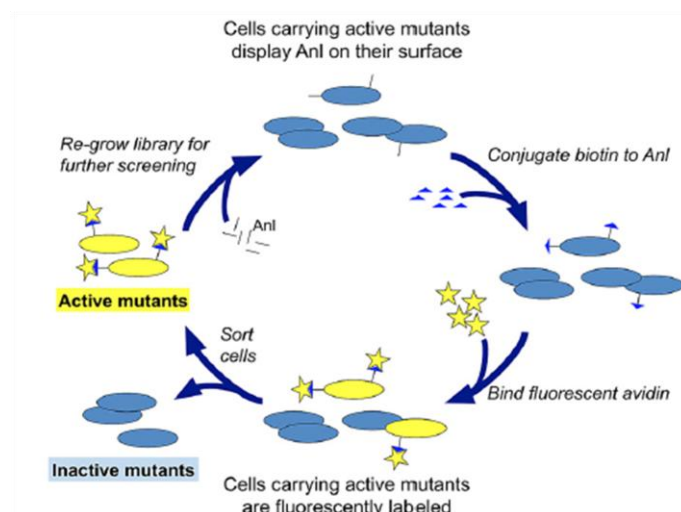


Figure 24 : Représentation schématique du mécanisme d'évolution dirigée spécifique aux AANC portant un groupement azido, ici l'azidonorleucine. Figure tirée de (Tanrikulu *et al.*, 2009)

La deuxième méthode d'évolution dirigée utilisée pour modifier des AARS-O s'appuie sur le « Multiplexe Automated Genome Engineering » (MAGE) pour générer une banque de mutants. Cette méthode consiste à faire *in vivo* des cycles de croissances qui comportent séquentiellement :

- L'insertion par électroporation d'un mélange oligonucleotides qui visent un gène spécifique dans les cellules, pour muter plusieurs positions données.
- La croissance cellulaire
- L'induction de la protéine qui permet l'incorporation des oligonucléotides
- L'interruption de croissance

Ce cycle permet d'introduire jusqu'à 30 % de modification toutes les 2h-2h30 (Wang *et al.*, 2009) et d'obtenir une bibliothèque de gènes avec des mutations à des positions spécifiques (figure 25). Cette méthode a été utilisée pour créer des variants de la TyrRS de *M. jannaschii* reconnaissant un anticodon différent (Amiram *et al.*, 2015). La souche utilisée est modifiée pour faciliter l'incorporation des oligonucléotides dans le plasmide.

La sélection dans ce cas ci a été faite par FACS. La fluorescence diminue en fonction du nombre d'AANC incorporés dans des variants de la GFP ou encore des variants de protéines interagissant avec la GFP. Cette méthode de crible dépend ici aussi des acides aminés incorporés (Figure 25).

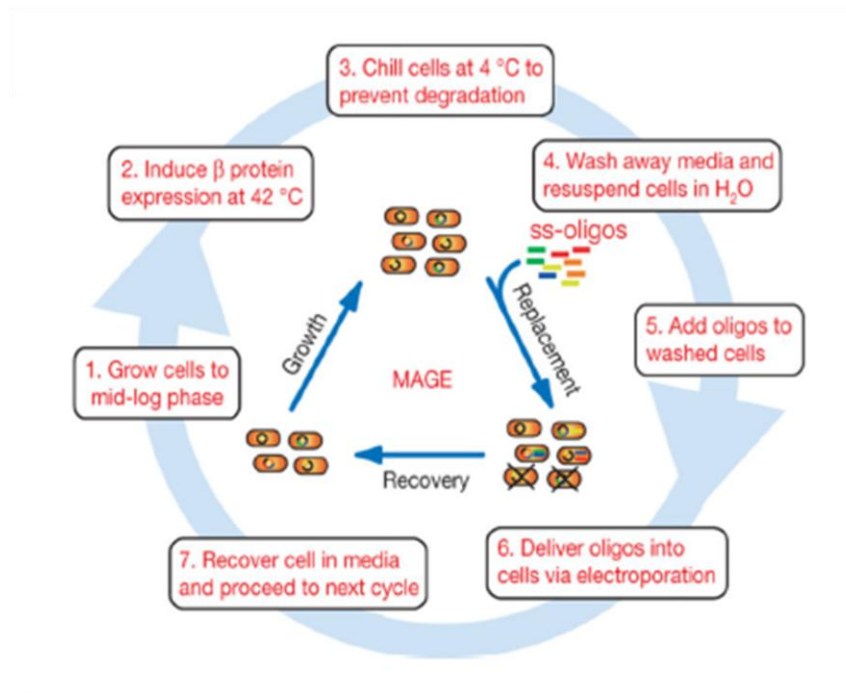


Figure 25 : Représentation schématique des étapes du cycle de « Multiplexe Automated Genome Engineering » (MAGE).
Figure tirée de (Wang *et al.*, 2009).

Une troisième méthode est appelée « Phage Assisted Continuous Evolution » (PACE). Elle permet de lier l'évolution d'un gène à la propagation des bactériophages (Esvelt *et al.*, 2011). Pour cette méthode, le gène à muter est inséré dans le génome d'un bactériophage à la position d'un gène essentiel pour la survie du phage, tandis que ce gène est exprimé par un plasmide. L'activité et/ou l'expression de la protéine codée par le plasmide est rendue conditionnelle à la propriété recherchée pour le gène inséré dans le bactériophage. Le gène qui doit évoluer sera donc transféré de cellule hôte à cellule hôte grâce au cycle de vie du bactériophage. Seuls les phages avec le gène essentiel actif pourront se propager et continuer à diffuser le gène mutant dans les cellules. A ce système est ajoutée une mutation dans le système de traduction qui augmente le taux d'erreurs lors de la traduction. Ce système permet d'obtenir en cycle fermé une génération de variants. Il est ainsi possible de faire apparaître des premiers variants de PylRS après 48 h et d'obtenir des variants efficaces après 20 jours (Bryson *et al.*, 2017). Dans cette étude, la méthode de criblage a consisté à introduire des codons ambre dans le gène III, un gène qui code pour une protéine à l'origine des filaments des bactériophages. Si la PylRS incorpore des acides aminés, le gène est fonctionnel et dans le cas contraire, le phage ne prolifère pas.

Une variante de cette méthode, « Phage-Assisted NonContinuous Evolution » (PANCE), a été développée. Elle diffère dans le sens où les différentes étapes sélectives et mutagènes se font en changeant le milieu de culture dans le réacteur (Suzuki *et al.* , 2017). La méthode de sélection utilisée pour les AARS orthogonales est la même que pour le PACE (Figure 26).

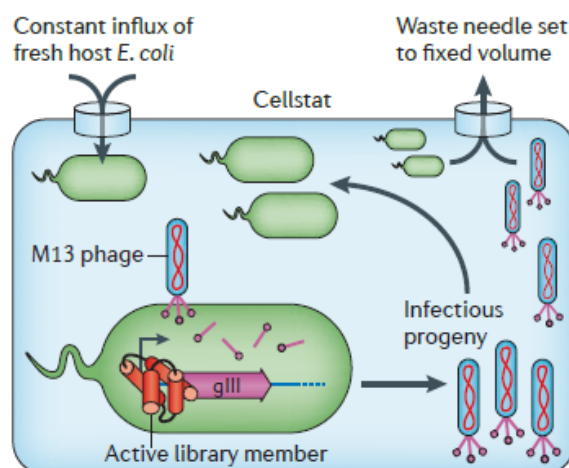


Figure 26 : Représentation schématique du processus de «Phage Assisted Continuous Evolution » PACE. Figure tirée de (Packer and Liu, 2015). Dans un réacteur (bleu), les cellules hôtes d'*E. coli* (vert) permettent les développements et l'évolution d'une population de bactériophages (bleu).

Dans le cadre de l'extension du code génétique, il y a peu d'études sur la modification des sites d'édition des AARS pour améliorer l'incorporation d'AANC. Un exemple intéressant est toutefois le domaine d'édition de la PheRS de *P. horihoshii* qui a été ajouté à la TyrRS d'*E. coli*. Ce domaine éditant la tyrosine, son ajout a permis de rendre la TyrRS spécifique de la 3-iodo-L-tyrosine (Oki *et al.* , 2008).

Toutes ces méthodes ont permis de créer des couples AARS-ARNt orthogonaux plus spécifiques pour certains acides aminés non canoniques, mais il restait encore à optimiser les autres acteurs du système de traduction.

IV. 2. 3 Modification des facteurs d'élongation et de terminaison

Par la lecture du codon ambre, les ARNt orthogonaux entrent en compétition avec les facteurs de terminaison qui ont pour rôle de terminer la synthèse de la protéine à la lecture d'un codon stop (RF1 pour le codon ambre chez *E. coli*). Plusieurs souches d'*E. coli* ont été créées où le gène de RF1 est supprimé (Mukai *et al.* , 2010; Johnson *et al.* , 2011; Wu *et al.* , 2013; Lajoie *et al.* , 2013). Pour rendre la souche viable, il a été nécessaire de

supprimer des codons ambre. Dans la première souche qui a été créée seuls les codons présents dans les sept gènes essentiels utilisant ce codon sont mutés (Mukai *et al.* , 2010). Dans une autre souche, tous les codons ambre ont été remplacés par des codons UAA (ocre) (Lajoie *et al.* , 2013; Mukai *et al.* , 2015).

EF-Tu est une protéine qui reconnaît les ARNt aminoacylés et les emmène jusqu'au ribosome comme il est montré sur la figure 27. En moyenne, son affinité pour les différents AA-ARNt est du même ordre de grandeur. EF-Tu reconnaît les ARNt par deux sites, le bras accepteur de l'ARNt et l'acide aminé. Les affinités pour chacun des deux sites sont variables et se compensent. Si l'interaction avec l'acide aminé est faible, le bras accepteur de l'ARNt aura une meilleure affinité pour EF-Tu. D'une façon générale les petits acides aminés ont une moins bonne affinité pour EF-Tu et donc l'affinité de leurs ARNt est haute (LaRiviere *et al.* , 2001) (Figure 27).

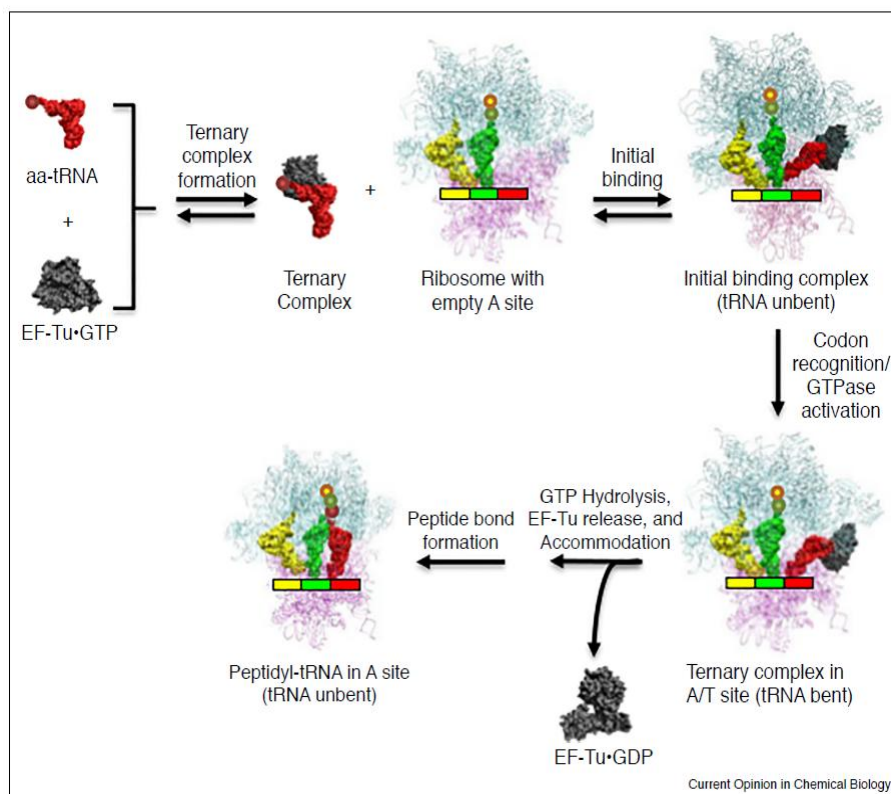


Figure 27 : Rôle d'EF-Tu dans les différentes étapes du mécanisme bactérien de traduction. Figure tirée de (Uhlenbeck and Schrader, 2018)

Lorsque les ARNt sont chargés avec des AANC, leur incorporation dans les protéines est faible. Chez *E. coli*, il a été montré que la concentration en EF-Tu et en ARNt aminoacylés sont semblables (Dennis and Bremer, 1974; Young and Furano, 1981). Il était donc raison-

nable de penser que les ARNt endogènes rentreraient en compétition avec les exogènes. La façon la plus simple de résoudre ce problème est d'augmenter la concentration en EF-Tu (Schrader *et al.* , 2011). Les affinités des ARNt pour EF-Tu augmentent lorsque la température baisse, ce qui peut être mis à profit (LaRiviere *et al.* , 2001). La dernière méthode utilisée pour augmenter l'affinité d'EF-Tu pour les AA-ARNt est de muter soit l'ARNt soit EF-Tu (Guo *et al.* , 2009; Ohtsuki *et al.* , 2010; Fan *et al.* , 2015; Katoh *et al.* , 2016; Gan *et al.* , 2017). En 2009, Guo et al ont effectué des modifications de l'ARNt^{Tyr} pour augmenter son affinité pour EF-Tu.

L'incorporation de D-AA fut difficile. D'une part, les ARNt chargés avec des D-AA ont une affinité faible pour EF-Tu. D'autre part, la vitesse de transfert du peptide entre deux D-AA est très faible, ce qui arrête la synthèse du peptide dans le site P via l'intervention d'EF-G. Ce problème a été résolu en mutant l'ARNt portant le D-AA. Le motif de la boucle D d'un ARNt^{Pro1} ainsi que le bras T de l'ARNt^{Glu} ont été transférés sur l'ARNt utilisé, ce qui a permis d'augmenter son efficacité (Gopalan *et al.* , 2015). Les acides aminés Glu et Asp ont une faible affinité pour EF-Tu comparés aux autres acides aminés canoniques. Le bras T de leurs ARNt compensent ce manque d'affinité (Dale *et al.* , 2004; Schrader *et al.* , 2011). L'utilisation de la boucle D de l'ARNt^{Pro} permet de recruter EF-P ce qui évite l'arrêt de la traduction du peptide. En effet la vitesse de synthèse du peptide est diminuée lorsqu'il y a plusieurs prolines à incorporer et l'intervention d'EF-P permet d'éviter l'interruption de la traduction (figure 25). L'utilisation des bras T et D des ARNt^{Glu} et des ARNt^{Pro} a permis d'améliorer l'incorporation des D-AA (Katoh *et al.* , 2016; Katoh *et al.* , 2017) (Figure 28).

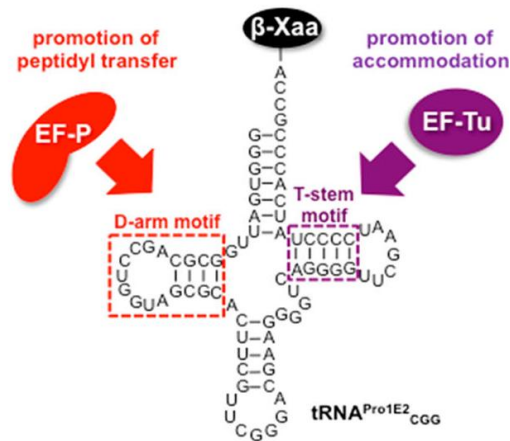


Figure 28 : ARNt^{Pro} modifié pour faciliter l'incorporation des acides aminés D ou β. Figure tirée de (Kato and Suga, 2018)

Ces différentes méthodes pour étendre le code génétique *in vivo* sont efficaces. En revanche, beaucoup de paramètres ne sont pas complètement contrôlables. C'est pourquoi certaines équipes de chercheurs se sont penchées sur la traduction *in vitro*.

IV. 3. Incorporation d'acides aminés non canoniques *in vitro*

IV. 3. 1 Système de traduction *in vitro*

En 2001 a été décrit le premier système de traduction modifiable ; le « Protein Synthesis Using Recombinant Element », aussi appelé système PURE (Shimizu *et al.*, 2001) (Figure 29).

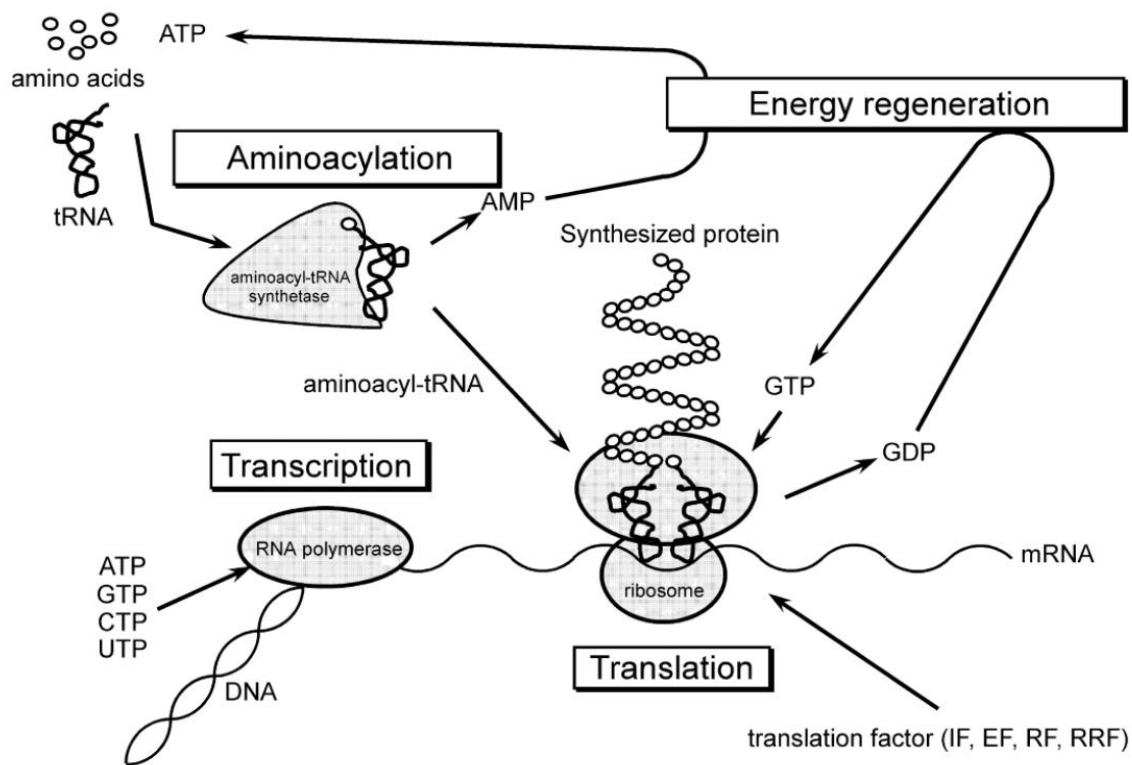


Figure 29 : Représentation schématique du système de traduction *in vitro* (PURE system. Figure tirée de (Shimizu *et al.* , 2001))

Le système PURE consiste à purifier les éléments composant le système de traduction, en incluant tous les éléments essentiels à la traduction : les facteurs de traduction (IF1, IF2, IF3, EF-Tu, EF-G et EF-Ts), les ARNt, les AARS, les ribosomes, les facteurs de terminaison (RF), les acides aminés et les facteurs de recyclage du ribosome (RRF) de *E. coli*. Ce système peut être utilisé pour synthétiser simultanément jusqu'à cinq peptides ou protéines différents. De cette manière il a été possible d'incorporer une valine sur un codon ambre en utilisant un Val-ARNt_{supp} synthétisé chimiquement et en omettant RF1 (Shimizu *et al.* , 2001). Le système PURE a plusieurs avantages comparé au système de traduction *in vitro* à partir d'extrait cellulaires. Ce système permet en effet de contrôler tous les éléments ajoutés dans la réaction, ce qui permet d'avoir un meilleur rendement et d'incorporer un AANC sur le codon UAG. De plus, il ne contient pas de protéases ni de nucléases. Ainsi, les protéines produites sont stables et les ARNm ne sont pas dégradés. Les nucléotides et acides aminés sont uniquement utilisés pour la traduction, ce qui améliore le rendement. La possibilité d'exclure certaines AARS et d'utiliser des ARNt aminoacylés à l'avance, permet d'incorporer des AANC (Hartman *et al.* 2007 ; Josephson, Hartman, et Szostak

2005). Malgré le fait que le système PURE soit un outil utile et efficace pour la synthèse de peptides *in vitro*, la préparation de tous les éléments de ce système ainsi que leur purification prennent du temps. La présence des AARS et des RF complique la réassignation des codons.

Un autre système contenant moins d'éléments a été développé (Forster *et al.*, 2003). Ce système ne contient pas d'AARS ni de facteur de terminaison ni de RRF. Les ARNt ajoutés doivent être aminoacylés au préalable. Ce type de système permet de réassigner chaque codon à un nouvel acide aminé. Tout comme *in vivo*, la disponibilité des ARNt aminoacylés est primordiale. Il faut donc être capable de les préparer (figure 30).

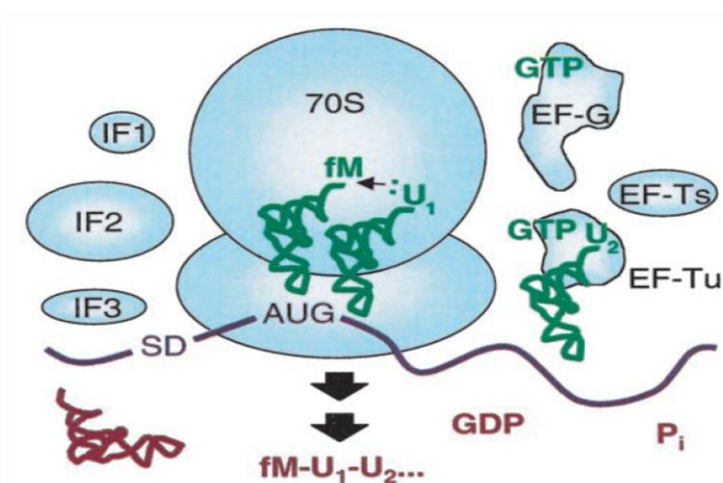


Figure 30 : Représentation schématique du système de traduction *in vitro* réduit. Figure tirée de (Forster *et al.*, 2003)

IV. 3. 2 Méthodes de préparation des ARNt aminoacylés

Initialement, les aa-ARNt étaient préparés par synthèse chimique (Hecht *et al.*, 1978). L'utilisation de flexizyme a permis de simplifier la production d'AA-ARNt. Les flexizymes sont des ribozymes capables de charger des AA sur les ARNt (Murakami *et al.*, 2003b; Murakami *et al.*, 2006). La première génération permettait uniquement de charger des acides aminés aromatiques (Murakami *et al.*, 2003a). La seconde génération a permis d' aminoacyler une plus grande variété d'acides aminés, en fonctionnalisant les acides aminés avec des groupements aromatiques partants (Ohuchi *et al.*, 2007). Les flexizymes de seconde génération ont une extrémité 3' qui vient reconnaître celle de l'ARNt permettant à l'acide aminé complexé au flexizyme d'être ensuite chargé sur l'ARNt (Figure 31).

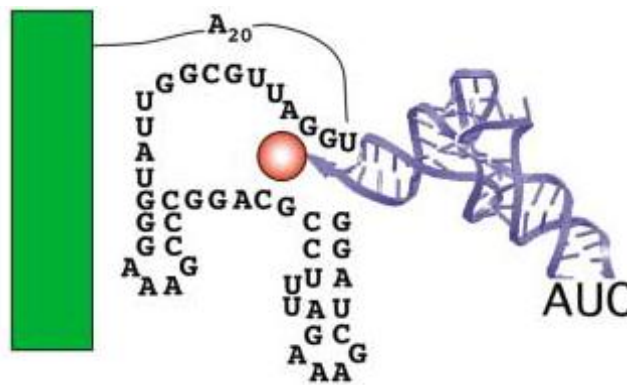


Figure 31 : Schéma montrant la synthèse d'ARNt-aminoacylé catalysée par un flexizyme fixé sur résine. Figure tirée de (Murakami *et al.* , 2003b). Le flexizyme reconnaît d'une part l'acide aminé et d'autre part la tige acceptrice de l'ARNt.

V Les acides aminés β

Les acides aminés non canoniques cités dans la partie précédente sont en grande partie des acides aminés α qui ont la même chaîne principale que les acides aminés canoniques mais avec une chaîne latérale différente. Cependant il existe un autre type d'acides aminés, les acides aminés β . Ceux-ci ont un carbone en plus dans la chaîne principale et ont aussi une chaîne latérale variable. On distingue les AA β^2 et les β^3 . Les premiers ont leur chaîne latérale en β du groupement aminé et les seconds en β du carboxyle (figure 32). Les acides aminés β sont souvent catégorisés comme acides aminés non naturels mais ils sont présents chez certains micro-organismes, plantes (Yasumoto and Satake, 1998), et chez les mammifères (Griffith, 1986). De plus, on les retrouve assez régulièrement dans des molécules naturelles (Kudo *et al.* , 2014). Ils sont souvent utilisés comme précurseurs de molécules plus importantes comme lors de la biosynthèse de la kirromycine. La β -Ala, par exemple, permet le développement musculaire. Il y a plusieurs types d'AA- β utilisés pour la synthèse de peptides bioactifs. Les premiers sont les homologues des acides aminés canoniques α avec simplement un méthylène en plus (Seebach *et al.* , 2004). Le second type d'acides aminés a une structure contrainte, souvent cyclique (figure 32).

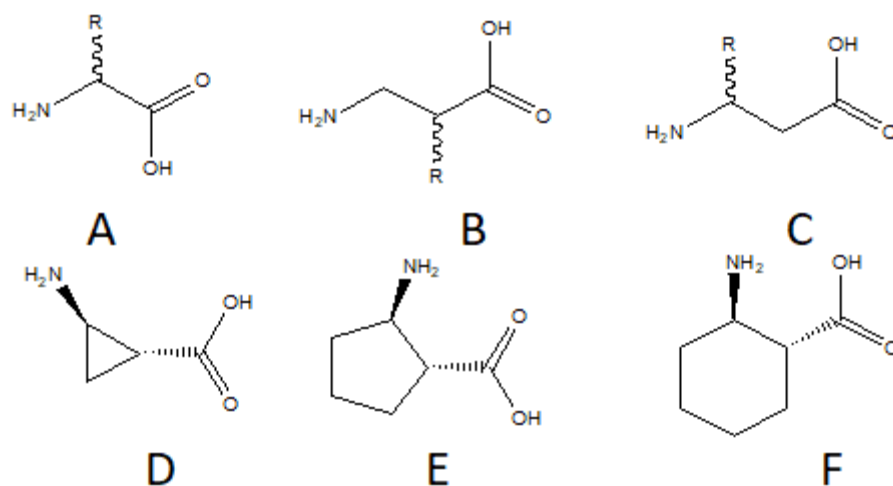


Figure 32 : Acides aminés α (A), β^2 (B), β^3 (C), Acide trans-2-amino cyclopropanoïque (D), Acide trans-2-amino cyclopentanoïque (E), Acide trans-2-amino cyclohexanoïque (F)

V. 1. Intérêt des AA- β

V. 1. 1. Spécificité des AA- β

La présence d'un groupement méthylène en plus dans la chaîne principale augmente la flexibilité de la chaîne principale, les possibilités de rotation ainsi que les distances entre les C_α (figure 33).

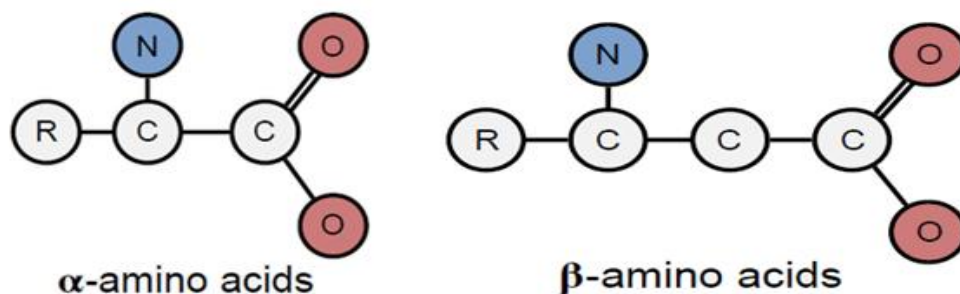


Figure 33 : Structure chimique des acides aminés α et β^3

Une fois incorporé dans une chaîne peptidique, un AA- β change les propriétés structurales du peptide. En effet, cette incorporation change l'homogénéité de la répartition des chaînes latérales ainsi que la répartition des charges (figure 34).

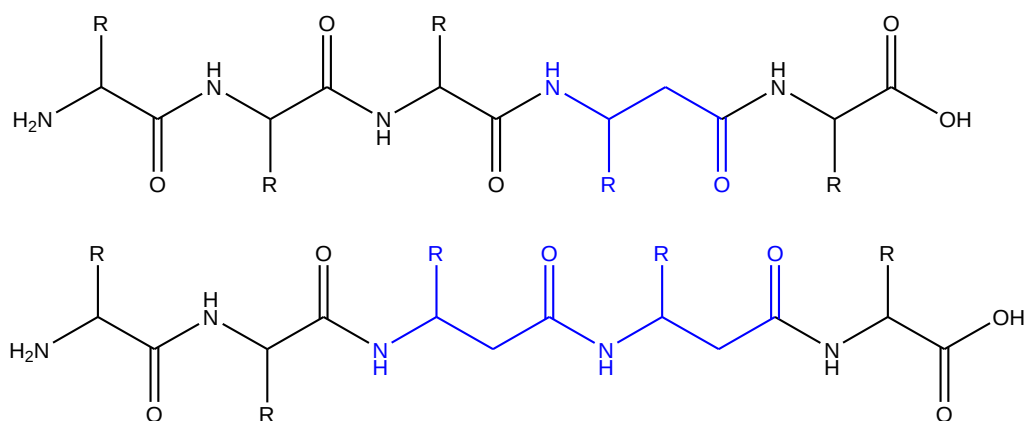


Figure 34 : Structure chimique de peptides contenant un et deux acides aminés β

Comparés à des peptides composés uniquement d'acides aminés α , les peptides β créent des motifs de structures secondaires différents et provoquent une résistance à l'hydrolyse par les enzymes (Daura *et al.*, 1997).

Des études sur la structure tridimensionnelle des peptides composés d'acides aminés α , β^2 et β^3 ont montré que les peptides contenant des acides aminés β avaient une grande tendance à former des hélices (Seebach *et al.*, 1999; Daura *et al.*, 2001; Seebach *et al.*, 2006). Les structures de celles-ci sont modulables en fonction du type d'acide d'aminé qu'elles contiennent. Les hélices sont souvent décrites avec la nomenclature de Gellman (Appella *et al.*, 1996) « hélice-X », où X est le nombre d'atomes constituant le cycle formé par la liaison hydrogène mise en jeu (figure 35). Les hélices contenant des AA- β sont le plus souvent des hélices 10, 12, et 14, la rotation de l'hélice varie en fonction du type d'acide aminé (β^2 , β^3 ou β cyclique) et de sa stéréo-isométrie (R ou S). Il est possible de contrôler le type d'hélice en modulant ces variables. En alternant les L-AA- β et les D-AA- β , il est possible d'éviter la formation d'une hélice (Seebach and Gardiner, 2008) et, au contraire, en utilisant uniquement des AA- β^3 le β peptide forme une « hélice-14 » (Seebach *et al.*, 1996). Une autre différence entre les peptides α et les peptides β , est que bien qu'il soit possible de former des hélices avec uniquement des peptides α , on aura plus de contraintes sur les chaînes latérales des acides aminés qui la composent.

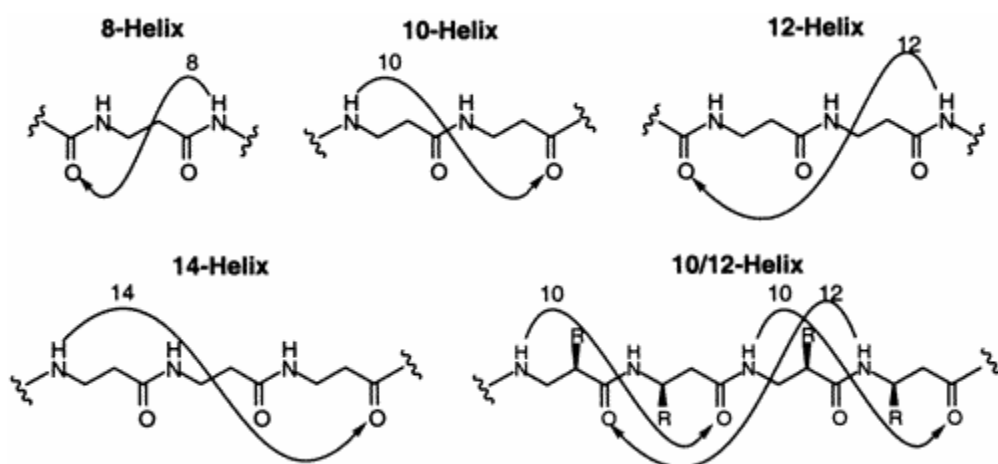


Figure 35 : Représentation des différents types d'hélices obtenus pour des β -peptides (Raynal, 2002)

V. 1. 2. Applications

Les propriétés géométriques des peptides β combinées à leur résistance à la protéolyse en font une très bonne base pour créer de nouvelles structures peptidiques biocompatibles.

A partir de là, des modifications ont été faites sur des peptides anti-tumoraux, en remplaçant des acides aminés α par des β , pour améliorer l'efficacité du peptide (Gademann *et al.*, 1999; Gademann *et al.*, 2001; Seebach *et al.*, 2001; Seebach *et al.*, 2008). D'autre part, les peptides β forment des foldamères. Ceux-ci sont classés en trois catégories hiérarchisées par structure. La première catégorie est constituée des peptides monomères utilisés dans la conception de molécules bioactives, d'antimicrobiens peptidiques et d'inhibiteurs spécifiques pour les interactions protéine-protéine.

L'une des applications dans le domaine pharmaceutique utilise la proximité structurale des AA- β avec les AA- α pour créer des peptides qui imitent des structures naturelles afin de copier des interactions peptides-protéines (Koglin *et al.*, 2003; Seebach and Gardiner, 2008) ou des interactions protéine-ADN (Gelman *et al.*, 2003). Par exemple, Seebach et al ont créé un peptide contenant des AA- β^3 , capable d'imiter un peptide responsable de la reconnaissance du non soi par les cellules T. Cela permet ainsi d'augmenter l'inhibition de la cytotoxicité, en inhibant les protéines responsables de la reconnaissance du non soi des cellules T (Seebach and Gardiner, 2008). En synthétisant des peptides contenant des acides aminés β cycliques (figure 36), il est possible de synthétiser des peptides qui sont analogues de neuropeptides. Bien qu'ils soient plus rigides que les neuropeptides Y, il a été montré que leurs capacités à inhiber les protéines reconnaissant ces peptides est bonne. Ils offrent

donc une bonne option en chimie médicinale pour développer des inhibiteurs. (Koglin *et al.* , 2003; Berlicki *et al.* , 2013).

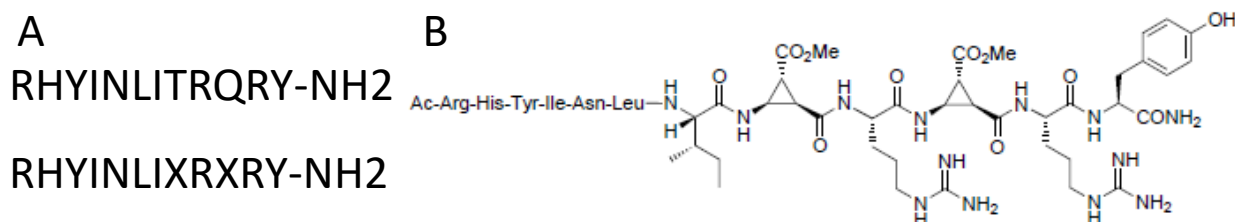


Figure 36 A : Séquence peptidique du neuropeptide Y et d'un de ses analogues contenant des AA- β écrits X. B : Formule développée de l'analogue du neuropeptide Y contenant deux acides aminés β (Koglin *et al.* , 2003)

La deuxième catégorie de structure est basée sur la construction de foldamères de peptides β auto-assemblés. Ceux-ci permettent la formation de canaux ioniques artificiels (Clark *et al.* , 1998), de fuseaux d'hélices (Daniels *et al.* , 2007; Horne *et al.* , 2007) ou encore d'imiter l'ADN et ce en utilisant des AANC comportant des chaînes latérales imitant les nucléosides. (Brückner *et al.* , 2003b; Brückner *et al.* , 2003a).

La capacité de peptides β à former des structures secondaires a été beaucoup étudiée depuis 2001, le premier exemple étant l'assemblage de peptides β^3 formant des hélices 14 (Raguse *et al.* , 2001). Il a été montré qu'en sélectionnant les AA formant le peptide, il était possible d'orienter la formation de fagots d'hélices, avec un cœur hydrophobe et des chaînes latérales extérieures hydrophiles (Horne *et al.* , 2007; Giuliano *et al.* , 2009). La structure de certains de ces fagots d'hélices a ainsi été résolue comme on peut le voir sur la figure suivante.

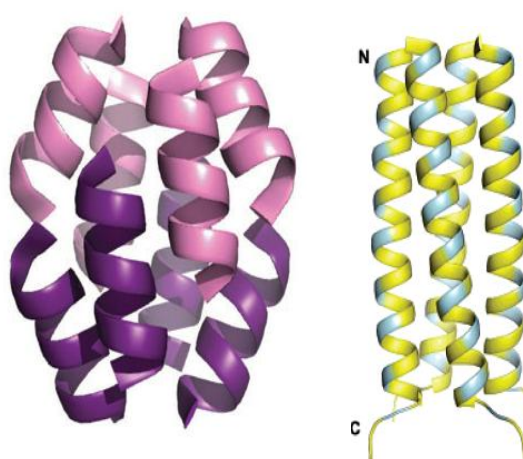


Figure 37 : Fagot de peptides β , octamérique et tetramérique (Daniels *et al.* , 2007; Horne *et al.* , 2007)

Le fagot d'hélices créé par le laboratoire de Schepartz a été fonctionnalisé, dans le sens où il a une activité. En fonction des AA qui le constituent, il peut catalyser des réactions chimiques telles qu'une saponification ou encore lier un métal (Wang and Schepartz, 2016).

Les canaux ioniques artificiels sont formés par une série de peptides cycliques s'insérant les uns au-dessus des autres dans la membrane lipidique (Clark *et al.*, 1998).

Les similitudes existantes entre des peptides β et l'ADN permettent de mimer l'ADN. En effet, deux acides aminés proches dans une « hélice 14 » sont espacés de manière similaires à deux nucléotides l'un au dessus de l'autre au sein d'une hélice d'ADN. (Brückner *et al.*, 2003b; Brückner *et al.*, 2003a).

La troisième catégorie de foldamères est plus récente. L'assemblage supramoléculaire de peptides β a permis de créer de nouveaux nanomatériaux. Plusieurs structures macroscopiques ont été formées à base de peptides β . Les avantages de ces matériaux sont nombreux. Ils sont biocompatibles, peu toxiques, simples à synthétiser et fonctionnalisables (Hauser *et al.*, 2011; Mahmoud *et al.*, 2011). Ces structures supramoléculaires sont formées par des interactions non covalentes permettant de construire des nouveaux bio-nanomatériaux (Pizzey *et al.*, 2008; Pomerantz *et al.*, 2008; Kwon *et al.*, 2010; Kim *et al.*, 2012; Del Borgo *et al.*, 2013).

Ces créations utilisant des acides aminés β montrent l'importance de développer les méthodes de préparation et d'incorporation de ces acides aminés.

V. 2 Incorporation d'acides aminés β dans des peptides

Bien que de nombreux peptides contenant des acides aminés β aient été synthétisés par synthèse peptidique, de nombreux efforts ont été faits pour avoir un système de traduction ribosomal capable d'incorporer des acides aminés β dans des peptides et des protéines *in vitro* ou *in vivo*. Le plus gros avantage de la synthèse ribosomale par rapport à la synthèse chimique est la vitesse, permettant par exemple de fabriquer des bibliothèques de peptides comportant plus de 10^{12} éléments différents à partir de bibliothèques d'ARNm aléatoire. De plus, les peptides peuvent être cyclisés chimiquement, ce qui permet d'avoir des bibliothèques de peptides linéaires et cycliques. La macro-cyclisation peut améliorer la résistance aux protéases et l'affinité du peptide (Driggers *et al.*, 2008).

En 2006, une première étude a montré la perméabilité d'une partie du système de traduction pour certains acides aminés β . L'équipe de Szostak a montré que certaines AARS étaient capables d'utiliser l'homologue β^3 de l'acide aminé qu'ils utilisent naturellement. Dans cette étude, des mélanges d'acides aminés ont été chargés sur les ARNt, ces derniers ont été digérés avant d'être analysés en LC-MS. L'analyse en spectrométrie de masse a permis d'identifier quels AA étaient chargés. L'affinité des AARS pour ces acides aminés n'a toutefois pas été mesurée, la spectrométrie de masse étant uniquement qualitative (Hartman *et al.*, 2006). Il a ensuite été montré par la même équipe que pour introduire ces AA *in vivo* dans des peptides, le ribosome était limitant (Hartman *et al.*, 2007). En 2012, l'équipe de Hecht a travaillé sur la modification du ribosome pour améliorer l'incorporation des AA- β (Dedkova *et al.*, 2012). Pour ce faire, une banque de mutants de l'ARNr 23S a été créée par PCR. Cette banque a d'abord été criblée négativement pour la résistance à l'érythromycine afin de sélectionner des variants ayant une architecture du centre peptidyltransférase permettant le passage d'acides aminés β . Les mutants obtenus ont été criblés positivement pour la sensibilité à la β -puromycin afin d'avoir des ribosomes susceptibles d'incorporer des acides aminés β (Maini *et al.*, 2015). Cette étude a permis de trouver des ribosomes mutants capables de catalyser la synthèse peptidique à partir d'acides aminés β .

En 2016, l'équipe de Suga a combiné les flexizymes permettant de produire les AA-ARNt plus facilement et le système de traduction *in vitro* PURE pour incorporer des acides aminés β (Fujino *et al.*, 2016). Ils ont montré que, en variant la concentration de EF-Tu, il était possible d'incorporer la plupart des AA- β . Au final, la β^3 -L-Pro, le β^3 -L-Trp et le β^3 -L-Glu ne sont pas incorporés tandis que la β^3 -L-Phe et la β^3 -L-Met sont les acides aminés les mieux incorporés. Ils définissent grâce à ces résultats quatre points qui sont susceptibles d'être limitants pour l'incorporation d'AA- β :

- La formation du complexe entre EF-Tu et les AA-ARNt
- L'arrivée de l'AA-ARNt au site A du ribosome
- La formation de la liaison peptidique entre le peptide et l'AA- β
- La formation de la liaison entre l'AA- β attaché au peptide et l'AA suivant.

De plus, malgré le fait que l'incorporation de plusieurs AA- β dans un même peptide soit possible, l'incorporation de plusieurs AA- β en séquence n'a été obtenue que plus tard (voir plus loin).

En parallèle, une étude a été menée sur l'incorporation d'acides aminés β *in vivo* (Melo Czekster *et al.* , 2016). Il a été montré que certaines AARS d'*E. coli* étaient capables d'aminoacyler les AA- β . De façon surprenante, les paramètres michaeliens mesurés pour l'activation de certains AA- α sont très proches de ceux mesurés pour les AA- β . Du plus, il a été montré que certains AA- β -ARNt avaient plus d'affinité pour EF-Tu que d'autre, en particulier la Met et la Phe, ce qui peut expliquer que ces AA- β sont mieux incorporés dans des peptides *in vitro* (Fujino *et al.* , 2016). L'incorporation *in vivo* de la β^3 - (p-Br)Phe a aussi été montrée, en utilisant le ribosome modifié développé par l'équipe de Hecht (Dedkova *et al.* , 2012; Maini *et al.* , 2015). Enfin, en 2018, l'équipe de Suga a optimisé son procédé d'incorporation pour introduire plusieurs AA β à la suite (Katoh and Suga, 2018). Il a été fait en sorte d'améliorer les points limitants qui avaient été définis auparavant. L'ARNt a été modifié comme pour incorporer des acides aminés D (IV. 2. 3). Ceci a permis d'augmenter l'affinité pour EF-Tu et pour EF-P et ainsi d'améliorer la prise en charge par le ribosome et la formation de la liaison peptidique. Ainsi, jusqu'à sept acides aminés β ont pu être incorporés à la suite.

Dès lors que, *in vitro*, beaucoup de problèmes qui auraient pu ralentir l'incorporation d'AA- β dans des peptides ont été résolus, nous sommes confrontés aux limitations de ce système. L'autre alternative, pour produire des peptides β , la synthèse *in vivo*, n'est toutefois pas sans limitations. On ne connaît en effet pas de couple orthogonal ARNt :AARS spécifique d'un AA- β .

Objectifs

Le but de ce projet est de créer une aminoacyl-ARNt synthétase spécifique d'un AA- β . Comme cela se fait habituellement, la cible sera un ARNt portant un anticodon complémentaire du codon Stop ambre (UAG). L'aminacyl-ARNt synthétase point de départ de cette étude, sera la méthionyl-ARNt synthétase d'*Escherichia coli*, puisque la β -Met fait partie des AA correctement incorporés par le ribosome (Fujino *et al.* , 2016). Ce choix a, au début de cette thèse, été guidé par la disponibilité de très nombreuses données sur cette enzyme. En particulier, la structure 3D de cette protéine est connue à haute résolution, y compris sous forme complexée à la méthionine. Cette enzyme a été un objet d'étude de longue date au laboratoire, et nous maîtrisons de nombreux outils biochimiques, enzymatiques, biophysiques, structuraux et génétiques qui facilitent l'obtention et la caractérisation de variants. Nous disposons de très nombreux variants, qui ont permis d'étudier de façon approfondie le mécanisme catalytique. Enfin, et surtout, nous disposons d'une part de variants capables d'aminacyler un ARNt portant un anticodon CUA, complémentaire du codon ambre, et d'autre part de variants dont la spécificité pour la méthionine est altérée. Pour atteindre notre objectif, nous utiliserons de manière conjointe l'ensemble des outils ci-dessus. Nous y ajouterons des outils de modélisation moléculaire *in silico*, basés sur des simulations utilisant la dynamique moléculaire et la mécanique quantique. Cette partie du travail se fera en collaboration avec Vaitea Opuu et Thomas Simonson au laboratoire. Nous avons à l'origine de cette étude une collaboration avec les Drs Valérie Pezo et Philippe Marlière, à l'Institut de Biologie Systémique et Synthétique (Évry), destinée à construire des systèmes microbiens permettant la sélection/amélioration *in vivo* d'enzymes ayant les propriétés recherchées. Les objectifs seront donc premièrement d'étudier l'interaction entre la MetRS et la β -Met et dans un second temps, d'améliorer cette affinité et la spécificité pour la β -Met. Enfin, se basant sur le point de départ ainsi défini, nous souhaitons créer une méthode d'évolution dirigée *in vivo* permettant de faire évoluer le système de traduction pour améliorer l'incorporation des AA- β .

Partie I

I Introduction

L'émergence des peptides composés d'AA- β a suscité divers développements dans les nanomatériaux (Gopalan *et al.*, 2015) ou dans le domaine pharmaceutique (Cabrele *et al.*, 2014). La production de peptides β constitue donc un enjeu important. La synthèse chimique est toutefois longue et le rendement décroît rapidement lorsque le peptide s'allonge. Il est donc nécessaire de trouver une alternative. Une synthèse ribosomale permettrait de construire de larges banques de peptides rapidement et efficacement. C'est pourquoi l'incorporation d'acides aminés β dans des peptides grâce au système de traduction ribosomale est de plus en plus étudiée depuis les quinze dernières années. En 2006, il a été montré que certains AA- β pouvaient être chargés sur les ARNt grâce au système de traduction ribosomale *in vitro* (Hartman *et al.*, 2006; Hartman *et al.*, 2007). Des études postérieures ont essayé d'améliorer les divers éléments pour optimiser l'incorporation des AA- β avec succès (Fujino *et al.* 2016; Melo Czekster *et al.* 2016; Katoh et Suga 2018). L'incorporation *in vitro* est devenue faisable. Toutefois, la lourdeur du protocole de synthèse ribosomale *in vitro* a poussé à mettre en place des systèmes *in vivo*. La plupart des améliorations obtenues pour les systèmes *in vitro* sont transposables aux systèmes *in vivo*, excepté le fait que la préparation des AA-ARNt se fait en amont dans les systèmes *in vitro*. *In vivo*, les AARS catalysent cette réaction. Pour certains AA- β (Melo Czekster *et al.*, 2016), les constantes cinétiques mesurées pour leur estérification sont, de façon surprenante, proches de celles de leurs homologues α . Cette étude a été publiée au moment où nous commençons ce travail de thèse, mais il nous a semblé indispensable d'approfondir le sujet afin de mieux comprendre la faible différence entre les vitesses d'aminacylation de l'acide aminé α et de l'acide aminé β^3 , dans le cas de la MetRS.

La Méthionyl-ARNt synthétase a été étudiée au laboratoire depuis de nombreuses années (Blanquet *et al.*, 2005). Différentes études par mutagenèse dirigée ont été menées sur des AA du site actif, du site de reconnaissance de l'anticodon, du site zinc ainsi que ceux des motifs spécifiques de la classe I. De plus, la structure de l'enzyme a été résolue sous sa

forme apo ou holo. L'ensemble des données obtenues fait de la MetRS une enzyme de choix pour l'ingénierie moléculaire. La MetRS a d'ailleurs déjà fait l'objet d'un criblage pour la rendre spécifique d'un AANC, l'isonorleucine.

La première partie de ce travail consiste en l'évaluation de l'interaction entre la β -méthionine et la méthionyl-ARNt synthétase. De façon assez surprenante, les travaux publiés juste avant le début de ma thèse (Melo Czekster *et al.*, 2016) suggéraient une affinité pour la β -Met équivalente à celle pour la méthionine. Dans cette partie, nous avons mesuré les constantes cinétiques avec la méthionine et la β -méthionine. Dans un second temps, nous nous sommes intéressés au positionnement de la β -méthionine dans la MetRS grâce à la détermination de la structure 3D résolue par RX à une résolution de 1,45 Å (Figure 38).

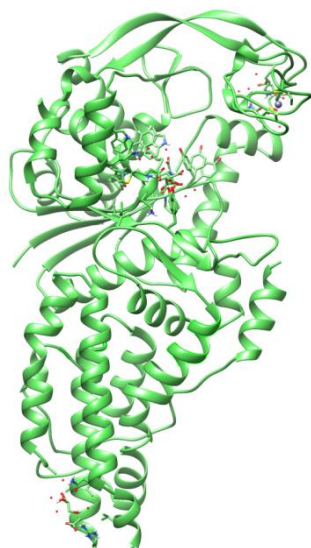


Figure 38 : Structure de la MetRS monomérique cocrystallisée avec la β^3 -L- méthionine. Les environnements de la β^3 -L- méthionine et ceux des molécules de citrate et glycérol observées dans le cristal sont montrés.

La première partie de notre étude a révélé la présence de méthionine contaminante dans la β -méthionine. Ainsi, nous démontrons que les mesures d'échange isotopique ATP-pyrophosphate (PPi) ne reflètent pas l'activation de la β -met mais celle de la Met contaminante. Dans un second temps, nous avons mis au point une méthode de digestion sélective de la méthionine pour obtenir de la β -méthionine à contamination très réduite.

Une méthode basée sur des analyses LC-MS et LC-MS/MS a ensuite été utilisée pour confirmer que chaque étape de la réaction d'aminacylation avait effectivement lieu. Les constantes cinétiques à l'état pré-stationnaire ont ensuite été mesurées en utilisant la variation de fluorescence intrinsèque de l'enzyme. Grâce à toutes ces méthodes, nous avons

comparé les deux substrats, et montré que si l'affinité de la β -Met pour la MetRS n'est que vingt-quatre fois plus faible que celle de l' α -Met, son activation en β -Met-adénylate est 5000 fois plus lente que l'activation de la Met en Met-AMP. La structure du complexe MetRS: β -Met permet de proposer des bases structurales pour ces résultats et de donner des pistes pour améliorer l'activité de l'enzyme vis-à-vis de la β -Met.

Enfin, de façon surprenante, la structure de la MetRS β -méthionine a révélé une seconde poche de liaison de la β -méthionine. Cette poche étant physiquement connectée au site actif, nous avons envisagé la possibilité que son occupation puisse influencer la catalyse. Nous avons donc dirigé une mutation pour détruire ce second site. Toutefois, cette mutation n'a qu'un effet modéré sur la catalyse.

Ces résultats nous informe donc que contrairement à ce qui avait été publié jusque là, la capacité de la MetRS à activer la β -Met est très faible. Il est donc envisageable de tenter de l'améliorer. L'étude structurale du complexe MetRS: β -Met est une bonne base pour trouver des positions à muter afin d'améliorer la capacité de l'enzyme à activer la β -Met de l'enzyme.

Ces travaux sont présentés dans l'article inclus ci-après que je signe en tant que premier auteur.



Contents lists available at ScienceDirect

Journal of Structural Biology

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

Use of β^3 -methionine as an amino acid substrate of *Escherichia coli* methionyl-tRNA synthetase

Giuliano Nigro^a, Sophie Bourcier^b, Christine Lazennec-Schurdevin^a, Emmanuelle Schmitt^{a,*}, Philippe Marlière^c, Yves Mechulam^{a,*}

^a Laboratoire de Biochimie, BIOC, Ecole polytechnique, CNRS, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France

^b Laboratoire de Chimie Moléculaire, LCM, Ecole polytechnique, CNRS, Institut Polytechnique de Paris, 91128 Palaiseau Cedex, France

^c Xenome Research Group, Institute of Systems and Synthetic Biology, CNRS, UMR8030, UEVE, CEA, GENOPOLE, 5 rue Henri Desbrières, 91030 Evry Cedex, France

ARTICLE INFO

Keywords:

Non-standard amino acid
Beta amino acid
Translation
X-ray crystallography
tRNA aminoacylation

ABSTRACT

Polypeptides containing β -amino acids are attractive tools for the design of novel proteins having unique properties of medical or industrial interest. Incorporation of β -amino acids *in vivo* requires the development of efficient aminoacyl-tRNA synthetases specific of these non-canonical amino acids. Here, we have performed a detailed structural and biochemical study of the recognition and use of β^3 -Met by *Escherichia coli* methionyl-tRNA synthetase (MetRS). We show that MetRS binds β^3 -Met with a 24-fold lower affinity but catalyzes the esterification of the non-canonical amino acid onto tRNA with a rate lowered by three orders of magnitude. Accurate measurements of the catalytic parameters required careful consideration of the presence of contaminating α -Met in β^3 -Met commercial samples. The 1.45 Å crystal structure of the MetRS: β^3 -Met complex shows that β^3 -Met binds the enzyme essentially like α -Met, but the carboxylate moiety is mobile and not adequately positioned to react with ATP for aminoacyl adenylate formation. This study provides structural and biochemical bases for engineering MetRS with improved β^3 -Met aminoacylation capabilities.

1. Introduction

The design of polypeptides having enhanced or novel properties is a major challenge for synthetic biology. Indeed, although natural proteins mostly rely on the set of 20 canonical amino acids to perform their functions, additional chemical groups are required for some of the biological activities. These groups are brought by rare amino acids such as selenocysteine or pyrrolysine, by post-translational modifications or by cofactors. Considerable work has been made for expanding the genetic code in order to direct the incorporation of non-standard amino acids at desired positions in polypeptides. Two types of approaches were used, using either living cells or *in vitro* reconstituted ribosomal translation systems.

A key requirement for *in vivo* approaches is to ensure the availability of tRNA aminoacylated with the non-standard amino acid. In some cases, this has been done using mutated aminoacyl-tRNA synthetases (aaRS) able to aminoacylate a wild-type tRNA. For instance, *E. coli* methionyl-tRNA synthetase mutants were selected for aminoacylation of tRNA^{Met} with the methionine surrogate azidonorleucine, resulting in efficient protein labelling *in vivo* in either bacterial or eukaryotic cells (Mahdavi et al., 2016; Ngo et al., 2009; Tanrikulu et al., 2009). The

early studies led to the development of more general residue-specific methods that allow partial to quantitative replacement of canonical amino acids by non-canonical analogues (Johnson et al., 2010). For site-specific incorporation, the non-standard amino acid is generally encoded by a non-sense or a frameshift codon (Chin, 2017; Liu and Schultz, 2010). In this case, an orthogonal tRNA, not recognized by any endogenous aaRS, is used together with an orthogonal aaRS, unable to aminoacylate any endogenous tRNA (Wang et al., 2006). *In vitro* approaches generally use cell-free systems reconstituted from purified components (Shimizu et al., 2001). One important advantage of such systems is that aminoacyl-tRNAs can if necessary be prepared using methods that do not require aminoacylation by aaRSs. Indeed, highly flexible methods for *in vitro* aminoacylation of tRNAs with a wide variety of amino acids are available. These methods are based on RNA catalysts named flexizymes (Murakami et al., 2003, 2006; Ohuchi et al., 2007) and have virtually no limitation of acid donor substrates.

Most non-standard amino acids successfully used for *in vivo* incorporation were L- α -amino acids, resulting in the same polypeptide backbone as the canonical amino acids. β -amino acids bear an additional methylene group either between the nitrogen atom and the α -carbon (β^2 type) or between the α -carbon and the carboxylate carbon

* Corresponding authors.

E-mail addresses: emmanuelle.schmitt@polytechnique.edu (E. Schmitt), yves.mechulam@polytechnique.edu (Y. Mechulam).

<https://doi.org/10.1016/j.jsb.2019.107435>

Received 4 September 2019; Received in revised form 15 November 2019; Accepted 13 December 2019

1047-8477/© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

(β^3 type) that results, when incorporated in a polypeptide, in a non-standard main chain. These non-canonical amino acids have deserved increasing interest because they give polypeptides unique properties, such as increased resistance to enzymatic hydrolysis (Heck et al., 2006; Webb et al., 2005) or the capacity to form peculiar helical structures (Petersson et al., 2007; Petersson and Schepartz, 2008; Seebach et al., 2006). Actually, β -amino acids were used to create peptides with interesting biological activities (Cheloha et al., 2015; Seebach and Gardiner, 2008).

Ribosomes could be modified to allow incorporation of β amino acids in cell free systems (Dedkova et al., 2011; Maini et al., 2013). It was also shown that various β -amino acids were compatible with ribosomal translation (Fujino et al., 2016). Moreover, elongation of peptides with successive β -amino acids was more recently rendered possible thanks to the use of engineered tRNAs having enhanced affinities for elongation factors EF-Tu and EF-P (Kato and Suga, 2018). These studies in cell-free systems required the supply of β -aminoacyl-tRNAs synthesized using flexizymes or chemical tools. Overall, the availability of aaRSs able to specifically aminoacylate tRNAs with β^3 -amino acids remains a bottleneck for their incorporation into polypeptides *in vivo*. Indeed, only few studies are available. For instance, mass spectrometry-based approaches have demonstrated the potential of aaRSs to use a wide variety of non-natural amino acids, including β -amino acids (Hartman et al., 2007, 2006). More recently, the capacity of some aaRS to use β^3 -amino acids was studied kinetically and β^3 -Phe analogues were successfully incorporated into dihydrofolate reductase in *E. coli* (Melo Czekster et al., 2016).

Here, we report a detailed study of the recognition of β^3 -Met by *E. coli* methionyl-tRNA synthetase (MetRS). In particular, we show that the interpretation of kinetic studies can be obscured by the presence of contaminating α -Met in commercial β^3 -Met samples and that the enzyme has limited capacities to activate β^3 -Met. However, MetRS binds β^3 -Met rather efficiently. We also describe a high-resolution crystallographic structure of the MetRS: β^3 -Met complex that allows understanding why β^3 -Met activation is poorly efficient. This structure will serve as a basis for engineering MetRS with improved β^3 -Met aminoacylation capabilities.

2. Results and discussion

2.1. Evidence for the presence of contaminating α -Met in β^3 -Met samples

It is well established that interpretation of [32 P]ATP-PPI exchange experiments can be obscured by the presence of methionine traces in commercial amino acids samples (Fersht and Dingwall, 1979). Prior to studying activation of β^3 -Met by MetRS, we used LC-HRMS experiments to search for the presence of α -Met in the β^3 -Met sample. A protonated molecule at m/z 164.074 (corresponding to β^3 -MetH⁺; retention time of 2.4 min) was mainly observed on the ESI-P mass spectrum for the β^3 -Met sample (Fig. 1). However, the results showed the presence of a compound having the exact molecular mass of α -MetH⁺ (m/z 150.058) at 2.2 min in β^3 -Met suggesting the occurrence of contaminating α -Met. Accordingly, the peak of contaminating α -Met corresponded to the major peak observed in an α -Met sample analyzed as a reference (data not shown). From the relative responses of the β^3 -Met and α -Met samples, we evaluated the contamination of α -Met in the β^3 -Met sample of about 0.1%. Actually, BOC-L-methionine was used as a reactant during the synthesis of β^3 -Met, as communicated by the company from which the compound was purchased. As an attempt to remove contamination, we treated β^3 -Met samples with methionine gamma lyase (MGL), an enzyme that degrades L-methionine into methanethiol, ammonia and 2-oxobutanate. This enzyme was likely to be specific of α -Met as compared to β^3 -Met since it binds both the amino and carboxylate groups of methionine (Fukumoto et al., 2014). The treated sample was then analyzed by mass spectrometry (Fig. 1). As expected, the molecular ion at m/z 150.058 [α -MetH]⁺ became undetectable in the

MGL sample (reduced by a factor of at least ten), whereas the peak corresponding to β^3 -Met remained present.

2.2. MetRS binds β^3 -Met

Because of this contamination, we investigated binding of β^3 -Met using enzyme intrinsic fluorescence titration experiments. This method is much less sensitive than ATP-PPI exchange to trace contaminations. Indeed, enzyme fluorescence varies measurably only if a significant proportion of the enzyme is bound to the ligand whereas even a very small fraction of enzyme bound to α -Met can catalytically give detectable accumulation of the [32 P]PPI reaction product. It is known that MetRS displays large variations of intrinsic fluorescence when binding methionine or methionyl adenylate from which useful thermodynamic parameters can be derived. Throughout this study, we used a His-tagged M547 monomeric version of *E. coli* MetRS, fully active, both *in vitro* and *in vivo* (Schmitt et al., 2009). Addition of β^3 -Met to MetRS caused an increase of intrinsic tryptophan fluorescence, similar to the increase caused by titration with α -Met. A dissociation constant of 1.2 ± 0.2 mM for the MetRS: β^3 -Met complex was derived from the titration data (Fig. 2 and Table 1). This value was only 24-fold higher than the K_d of MetRS: α -Met. According to the 0.1% contamination level of α -Met in our β^3 -Met sample, the concentration of α -Met in the titration experiments did not exceed 4 μ M. Because this value was low as compared to the K_d of α -Met for MetRS (50 μ M, Table 1), we concluded that the contaminating α -Met could not influence the measured K_d value of β^3 -Met. Consistent with this conclusion, no detectable effect on the dissociation constant of the MetRS: β^3 -Met complex was observed if β^3 -Met was treated with MGL prior to the measurement (Fig. 2). This latter result also showed that the non-canonical amino acid was not affected by MGL treatment.

2.3. Interpretation of the [32 P]ATP-PPI exchange assay is obscured by amino acid contamination

Measurement of the apparent catalytic parameters ($k_{cat-app}$ and K_{m-app} for β^3 -Met) of the [32 P]ATP-PPI isotopic exchange reaction using the commercial β^3 -Met sample showed that as compared to α -Met, $k_{cat-app}$ was lowered by a factor of 42 (1.2 s⁻¹ as compared to 50 s⁻¹) whereas K_{m-app} was increased by a factor of 35 (630 μ M as compared to 18 μ M). Overall, the catalytic efficiency of the isotopic exchange reaction was reduced by an apparent factor of ca 1500 when α -Met was replaced by β^3 -Met. Notably, these results were at odd with those of a study reporting that MetRS was very tolerant to the substitution of α -Met by β^3 -Met, with only a modest (2-fold) preference for Met in ATP-PPI exchange (Melo Czekster et al., 2016). Moreover, in the presence of 2 mM β^3 -Met treated with MGL, the isotopic exchange rate dropped to 0.07 s⁻¹ as compared to 0.9 s⁻¹ in the presence of 2 mM untreated β^3 -Met (a value in agreement with $k_{cat-app}$ of 1.2 s⁻¹ and K_{m-app} of 0.63 mM). This result showed that, at least with the commercial sample, the β^3 -Met-dependent [32 P]ATP-PPI exchange rate mostly reflected activation of contaminating α -Met. The exchange rates measured when varying the concentration of untreated β^3 -Met were therefore interpreted using a competitive inhibition scheme (Fig. 3). In this scheme, β^3 -Met binds the enzyme at a site overlapping the α -Met site but cannot be activated. Fitting showed that the experimental results were consistent with activation of β^3 -Met occurring at a negligible rate and β^3 -Met being contaminated with 0.09% α -Met (mol/mol; Fig. 3). Notably, this level of contamination was compatible with the results of mass spectrometry analyses of β^3 -Met. Finally, traces of α -Met were also evidenced in enzyme preparations. Attempts to diminish the contamination to a level where β^3 -Met-dependent [32 P]ATP-PPI exchange rates can be accurately measured were unsuccessful. Because of the high sensitivity of ATP-PPI isotopic exchange to the presence of traces of α -Met, we concluded that this method was not suitable to detect a possible activation of β^3 -Met by MetRS.

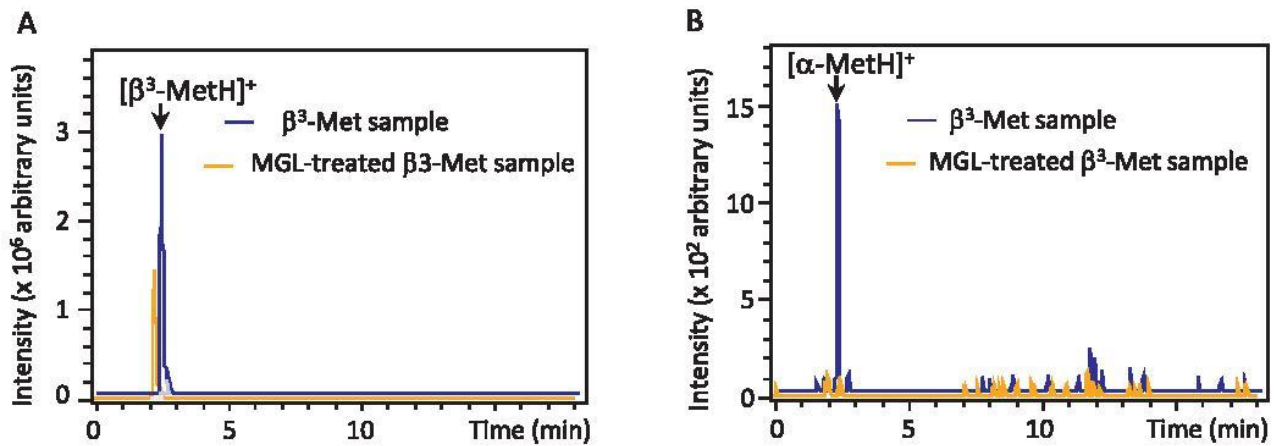


Fig. 1. LC-HRMS analysis of β^3 -Met (blue) and MGL-treated β^3 -Met (yellow) samples. Extracted ion current (EIC) chromatograms of (A) m/z 164.074 $\pm$ 0.001 corresponding to $[\beta^3\text{-MetH}]^+$ ($\text{C}_6\text{H}_{14}\text{NO}_2\text{S}$) and (B) m/z 150.058 $\pm$ 0.001 corresponding to $[\alpha\text{-MetH}]^+$ ($\text{C}_5\text{H}_{12}\text{NO}_2\text{S}$). The blue chromatograms were shifted to the right for clarity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

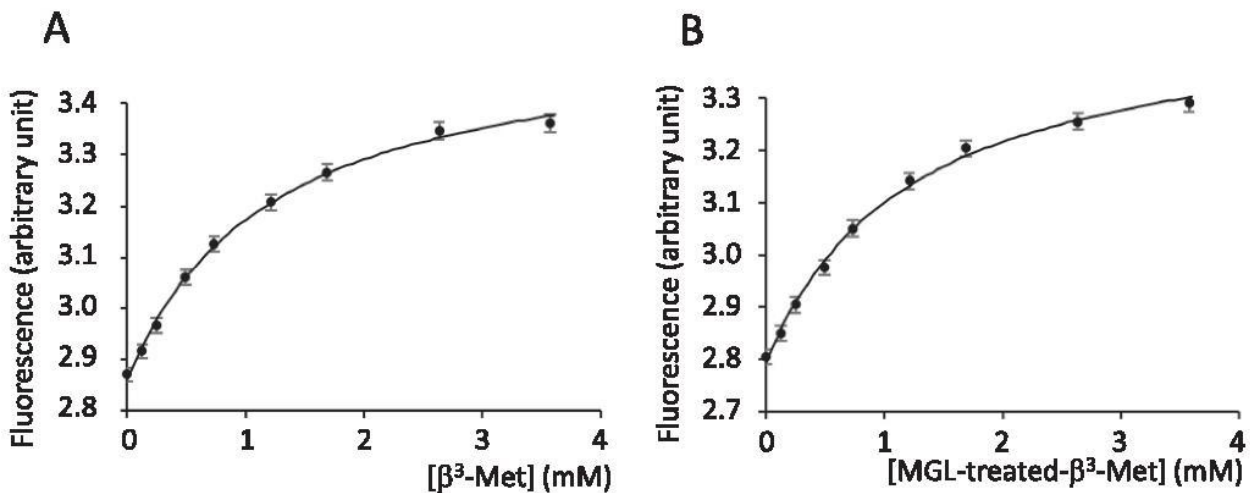


Fig. 2. Intrinsic enzyme fluorescence titration of MetRS with β^3 -Met. Panel A: MetRS (0.5 μM) was titrated with increasing amounts of commercial β^3 -Met, as described in Materials and Methods. The experimental points show the measured fluorescence as a function of the concentration of β^3 -Met. Data were fitted to a simple equilibrium model using standard least-square procedures. The curve shows the resulting fit ($K_d = 1.2$ mM). Panel B: Same experiment but β^3 -Met was treated with MGL before titration ($K_d = 1.2$ mM).

Table 1

Kinetic and thermodynamic constants of α -Met or β^3 -Met binding, activation and aminoacylation by wild-type MetRS and the V298I variant. Dissociation and rate constants are defined in Scheme 1. Each experiment was performed at least twice independently. Results are expressed as mean $\pm$ either standard deviation from the independent experiments or standard error from the fitting procedure, whichever the greater.

Enzyme/Substrate	WT/ α -Met	WT/ β^3 -Met	V298I/ α -Met	V298I/ β^3 -Met
K_{Met} (mM)	0.05 $\pm$ 0.01	1.2 $\pm$ 0.2	0.029 $\pm$ 0.001	0.56 $\pm$ 0.01
$K_{\text{Met}}^{\text{ATP}}$ (mM)	0.070 $\pm$ 0.004	0.30 $\pm$ 0.04	0.063 $\pm$ 0.004	0.15 $\pm$ 0.02
k_f (s^{-1})	260 $\pm$ 10	0.051 $\pm$ 0.003	262 $\pm$ 12	0.057 $\pm$ 0.005
K_{PP} (mM)	0.075 $\pm$ 0.025	1.0 $\pm$ 0.25	0.06 $\pm$ 0.01	0.8 $\pm$ 0.2
k_b (s^{-1})	189 $\pm$ 23	1.2 $\pm$ 0.3	195 $\pm$ 30	0.95 $\pm$ 0.2
Aminoacylation rate (s^{-1})	0.52 $\pm$ 0.07	(5.6 $\pm$ 0.3) 10^{-4}	n.d.	n.d.

Aminoacylation rates were measured in the presence of 2 μM tRNA, 2 mM ATP and 2 mM of either α -Met or β^3 -Met. Results are the mean $\pm$ sd from two independent experiments.

2.4. β^3 -Met is a poor substrate of MetRS

As described above, fluorescence titration can be used to follow the bindings of α -Met or β^3 -Met to MetRS. It is also known that addition of ATP in the presence of magnesium ions to the MetRS:Met complex leads to a large decrease of fluorescence intensity reflecting synthesis of

methionyl adenylate. Hence, a MetRS: β^3 -Met complex formed in the presence of MGL-treated β^3 -Met was mixed with ATP- Mg^{2+} . We observed a slow increase in fluorescence (rate constant of the order of 0.05 s^{-1}), contrasting with the fast decrease observed in the case of α -Met adenylate formation. However, importantly, this increase in fluorescence was reversed by the subsequent addition of

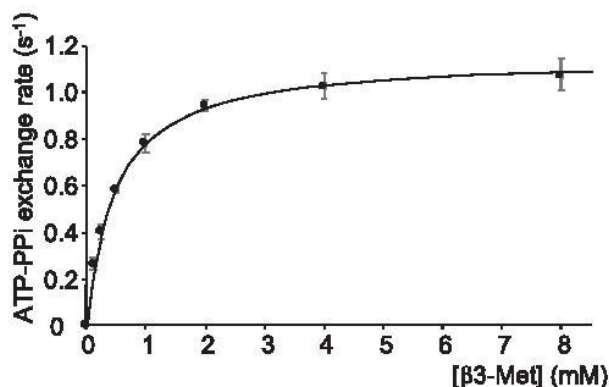


Fig. 3. Fitting the β^3 -Met-dependent [32 P]ATP-PPi exchange with a competition model. Initial rates of [32 P]ATP-PPi exchange catalyzed by MetRS were measured in the presence of 2 mM ATP and various concentrations of commercial β^3 -Met. Data (average from three independent experiments) were fitted to a classical competitive inhibition scheme, where isotopic exchange depends on contaminating α -Met. Adjusted parameters for fitting were the contamination rate λ ($[\alpha\text{-Met}]/[\beta^3\text{-Met}]$) and the inhibition constant K_i . The curve shows the best fit to the experimental data, with a contamination rate of $(8.9 \pm 0.3) \cdot 10^{-4}$ and an inhibition constant of $477 \pm 21 \mu\text{M}$. Michaelian parameters of the α -Met-dependent exchange were measured with the same batch of enzyme ($k_{\text{cat}} = 50 \text{ s}^{-1}$ and $K_m = 18 \mu\text{M}$) and kept constant in the fitting procedure. The theoretical equation giving the initial rate (v_i) as a function of the concentration of β^3 -Met ($[S]$) was: $v_i = k_{\text{cat}} [E] \frac{[S]}{[S] + \frac{K_m}{\lambda} \left(1 + \frac{[S]}{K_i}\right)}$.

pyrophosphate, suggesting that it could be due to the formation of β^3 -Met adenylate. In order to unambiguously confirm the formation of β^3 -Met adenylate, we analyzed the product of the reaction of ATP-Mg $^{2+}$ with MetRS: β^3 -Met using LC-HRMS (Fig. 4A). These experiments clearly identified an ion at m/z 493.127 with elemental composition of $\text{C}_{16}\text{H}_{26}\text{N}_6\text{O}_8\text{PS}$ corresponding to protonated β^3 -Met adenylate. This product was not present in the control sample where the reaction was stopped immediately after addition of MetRS (Fig. 4A). As expected, the corresponding ion for α -Met (m/z 479) was not observed. To obtain structural characterization of this compound, LC-HRMS/MS experiments have been carried out on the molecular ion (Fig. 4B). Three major ions were obtained. The ion at m/z 348.07 was formed from m/z 493.127 by elimination of $\text{C}_6\text{H}_{15}\text{NO}$ according with its elemental composition $\text{C}_{10}\text{H}_{15}\text{N}_5\text{O}_7\text{P}$ and corresponded to protonated AMP. Ions at m/z 164.074 ($\text{C}_{10}\text{H}_{15}\text{N}_5\text{O}_7\text{P}$) and m/z 136.062 ($\text{C}_5\text{H}_6\text{N}_5$) were formed from the molecular ion and correspond to the protonated β^3 -Met and adenine, respectively. A fragmentation mechanism of the m/z 493 ion is proposed in the Supplementary Scheme S1. These results clearly confirmed the identification of the reaction product as β^3 -Met adenylate.

The experiments described above showed that the increase in intrinsic enzyme fluorescence observed during reaction of MetRS with MGL-treated β^3 -Met and ATP-Mg $^{2+}$ indeed reflected the synthesis of β^3 -methionyl adenylate. We therefore measured the rates of fluorescence variation in the presence of 2 mM ATP-Mg $^{2+}$ and increasing amounts of MGL-treated β^3 -Met in order to derive the thermodynamic constants of the reaction, $K_{\beta^3\text{-Met}}^{\text{ATP}}$, the dissociation constant of β^3 -Met from the MetRS:ATP-Mg $^{2+}$ complex and $k_f \beta$, the rate of β^3 -methionyl adenylate synthesis (Scheme 1, Supplementary Fig. S1). The results (Table 1) showed that affinity of β^3 -Met for the MetRS:ATP-Mg $^{2+}$ complex was reduced by one order of magnitude as compared to that of α -Met, and that the rate of activation of the β amino acid was 5200-fold lower than that of canonical methionine. Similarly, the enzyme remained able to catalyze the reversion of enzyme-bound β^3 -methionyl adenylate by PPi-Mg $^{2+}$, though at a rate reduced by a factor of 157 as compared to the reversion of α -methionyl adenylate. Overall, these experiments showed that MetRS was indeed able to activate β^3 -Met, but with a catalytic efficiency ($k_f \beta / K_{\beta^3\text{-Met}}^{\text{ATP}}$) reduced by four orders of magnitude as compared to canonical methionine. This low catalytic efficiency was

mainly caused by small $k_f \beta$ and $k_b \beta$ kinetic constants, reflecting a non-optimal stabilization of the transition state of the activation reaction according to the Arrhenius law.

Finally, we examined the ability of MetRS to transfer β^3 -Met onto tRNA $^{\text{Met}}$. This was first done by analyzing the products of aminoacylation reactions (see methods) using mass spectrometry after digestion with S1 nuclease (Hartman et al., 2007, 2006). The results showed the presence of a product corresponding to AMP esterified with β^3 -methionyl-AMP (molecular ion at m/z 493.127) whereas no α -methionyl-AMP was detected (Fig. 4C). In order to prevent deacylation during sample preparation and to have a better signal in MS/MS, an aliquot of the aminoacyl-tRNA was N-acetylated using acetic anhydride prior to S1 nuclease digestion. Consistent with our expectations, N-acetyl- β^3 -methionyl-AMP (MH^+ m/z 535.138) was found in the acetylated sample whereas N-acetyl- α -methionyl-AMP (MH^+ m/z 521.121) was absent (Fig. 4C). Reference experiments with α -Met instead of β^3 -Met gave the expected results (Fig. 4D). Identifications of N-acetyl- β^3 -methionyl-AMP and N-acetyl- α -methionyl-AMP were further confirmed using HR-MS/MS of protonated compounds. Supplementary Table S1 brings together the detected fragments ions for both compounds. These results allowed us to propose the fragmentation scheme of protonated molecules reported in Supplementary Scheme 2.

We finally measured the rate of tRNA aminoacylation with β^3 -Met using the assay based on [$3'$ - 32 P]-labelled tRNA (Ledoux and Uhlenbeck, 2008). In the presence of 2 mM MGL-treated β^3 -Met and 2 mM ATP, MetRS catalyzed the aminoacylation of tRNA $^{\text{Met}}$ (2 μM) at a rate three orders of magnitude lower than the rate measured in the control experiment with α -Met (Fig. 5 and Table 1). It should be underlined that the rate value measured with β^3 -Met must be taken cautiously since, even after treatment with MGL, traces of α -Met remain. Nevertheless, taking into account the mass spectrometry experiments, the results show that MetRS can transfer β^3 -Met onto tRNA though at a reduced rate, in agreement with the reduced efficiency of the β^3 -Met activation step as compared to the Met activation step.

2.5. Structural study of β^3 -Met binding to MetRS

The results described above raise the question of how β^3 -Met is recognized by MetRS. Understanding this may also guide efforts aimed at engineering the enzyme for more efficiency and specificity towards the β amino acid. We therefore crystallized the enzyme in the presence of β^3 -Met. This was done by microseeding a supersaturated (1.08 M ammonium citrate, 6 °C) enzyme solution containing 20 mM β^3 -Met with crystals of the unliganded enzyme. Using such crystals, a complete dataset was obtained at 1.45 Å resolution (Table 2). For better comparison of α -Met and β^3 -Met binding sites, we also grew crystals of the MetRS: α -Met complex and obtained an X-ray diffraction dataset at 1.2 Å resolution. The previously published structure of the MetRS: α -Met complex was indeed determined at a lower 1.85 Å resolution (PDB id 1F4L, (Serre et al., 2001)). To avoid any bias, the two structures were solved by molecular replacement using the structure of unliganded MetRS as a search model (PDB id 1QQT, (Mechulam et al., 1999a)). In both cases, the presence of the ligand was obvious in the molecular replacement density map. The structures were then refined using successive rounds of manual rebuilding in Coot and energy minimization in Phenix. Final R/Free-R values were 0.092/0.126 (MetRS: β^3 -Met) complex and 0.145/0.171 (MetRS: α -Met complex; Table 2). The enzyme is made up of an N-terminal Rossmann fold containing the catalytic centre (Mechulam et al., 1991; Schmitt et al., 1994) and a C-terminal helix-bundle domain responsible for recognition of the CAU anticodon of tRNAs $^{\text{Met}}$.

As previously described (Crepin et al., 2003; Serre et al., 2001; Tanrikulu et al., 2009), binding of α -Met caused a dramatic rearrangement of aromatic residues. The side chains of W229, W253, F300 and F304 rotate such that W229 stacks on F304 and W253 stacks on F300. Furthermore, Y15 rotates, acting as a lid that locks the amino

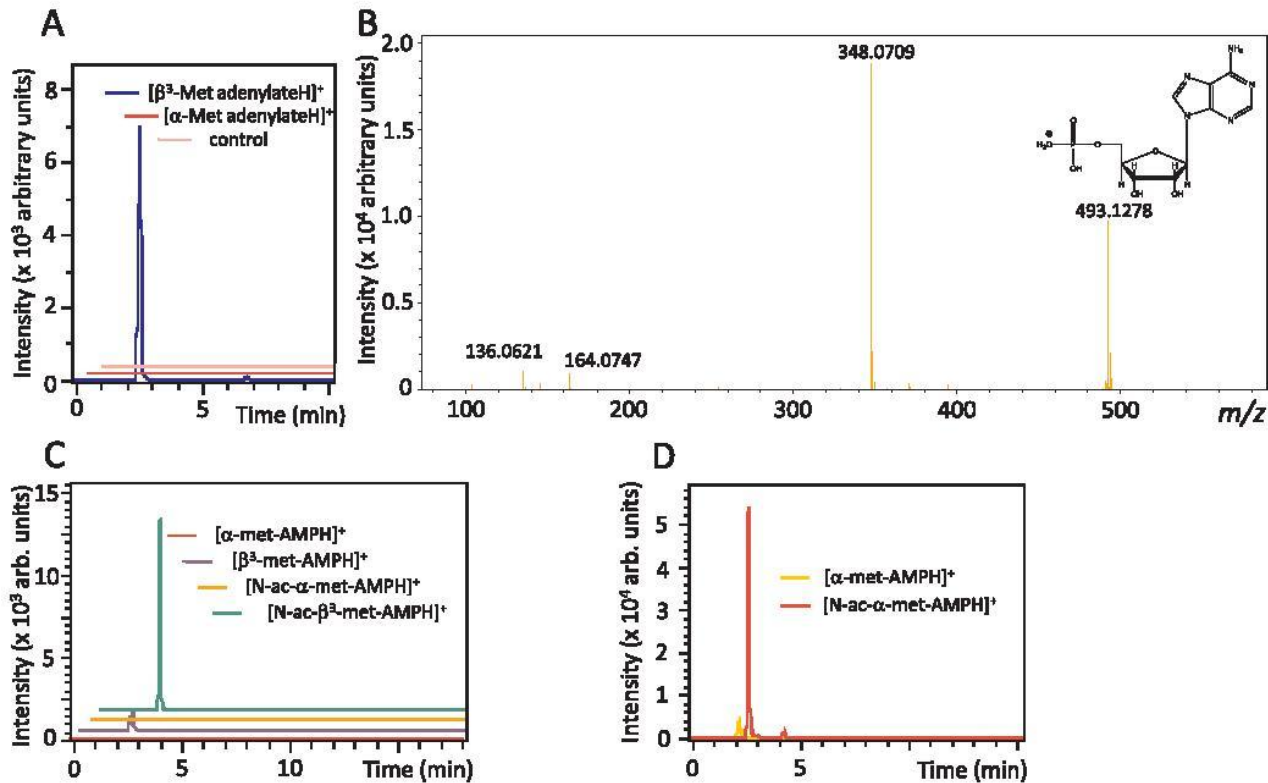
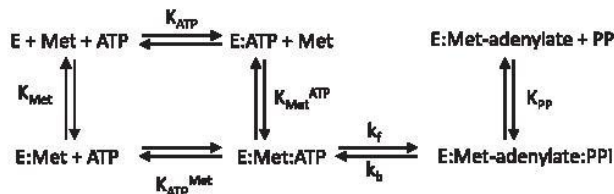


Fig. 4. LC-HRMS analyses of MGL-treated β^3 -Met activation and transfer onto tRNA^{Met} . Panel A: Analysis of the products of the activation reaction. The red and blue chromatograms correspond to the sample in which MetRS was reacted with ATP and MGL-treated β^3 -Met. Extracted ion current (EIC) chromatograms of m/z 479.111 $\pm$ 0.001 (red, corresponding to $[\alpha\text{-Met adenylateH}]^+$, $\text{C}_{15}\text{H}_{24}\text{N}_6\text{O}_8\text{PS}$) and m/z 493.126 $\pm$ 0.001 (blue, corresponding to $[\beta^3\text{-Met adenylateH}]^+$, $\text{C}_{16}\text{H}_{26}\text{N}_6\text{O}_8\text{PS}$). The pink chromatogram corresponds to m/z 493.126 $\pm$ 0.001 ($[\beta^3\text{-Met adenylateH}]^+$) for the zero reaction time control sample. Panel B: HR-MS/MS ($E_{\text{col}} = 15$ eV) of the m/z 493.1263 ion obtained at 2.6 min retention time (blue chromatogram in Panel A). The proposed mechanism for the decomposition of the molecular ion is indicated on the Supplementary Scheme 1. Panel C: LC-HRMS analysis of the aminoacylation reaction in the presence of MGL-treated β^3 -Met. The samples were submitted to S1 nuclease cleavage as indicated in Materials and Methods. The panel shows the extracted ion current chromatograms (EIC) of m/z 479.111 $\pm$ 0.001 (corresponding to $[\alpha\text{-methionyl-AMPH}]^+$, $\text{C}_{15}\text{H}_{24}\text{N}_6\text{O}_8\text{PS}$; red) and m/z 493.126 $\pm$ 0.001 (corresponding to $[\beta^3\text{-methionyl-AMPH}]^+$, $\text{C}_{16}\text{H}_{26}\text{N}_6\text{O}_8\text{PS}$; purple). The two other chromatograms correspond to the samples acetylated prior to S1 cleavage, showing the EIC of m/z 521.121 $\pm$ 0.001 (corresponding to $[\text{N-acetyl-}\alpha\text{-methionyl-AMPH}]^+$, $\text{C}_{17}\text{H}_{26}\text{N}_6\text{O}_9\text{PS}$; orange) and m/z 535.138 (corresponding to $[\text{N-acetyl-}\beta^3\text{-methionyl-AMPH}]^+$, $\text{C}_{18}\text{H}_{28}\text{N}_6\text{O}_9\text{PS}$; green). The panel shows the presence of the compounds expected from the transfer of β^3 -Met onto tRNA and the absence of detectable aminoacylation with α -Met. Panel D: Reference aminoacylation experiment using α -Met instead of β^3 -Met, with (red) or without (yellow) acetylation. The panel shows the EIC of m/z 479.111 $\pm$ 0.001 (corresponding to $[\alpha\text{-methionyl-AMPH}]^+$, $\text{C}_{15}\text{H}_{24}\text{N}_6\text{O}_8\text{PS}$; yellow) and m/z 521.121 $\pm$ 0.001 (corresponding to $[\text{N-acetyl-}\alpha\text{-methionyl-AMPH}]^+$, $\text{C}_{17}\text{H}_{26}\text{N}_6\text{O}_9\text{PS}$; red). The ions at m/z 535.138 ($[\text{N-acetyl-}\beta^3\text{-methionyl-AMPH}]^+$) and m/z 521.121 ($[\text{N-acetyl-}\alpha\text{-methionyl-AMPH}]^+$) were further characterized by HR-MS/MS (Supplementary Table 1 and Supplementary Scheme 2). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Scheme 1. Scheme for methionyl adenylate synthesis and reversion by MetRS. The scheme defines the five dissociation constants and the two rate constants k_f and k_b .

acid in its binding cavity (Fig. 6). These induced rearrangements reduce the size of the cavity and allow accommodation of the methionine side-chain. The α -Met cavity is bordered by L13, W253, P257, Y260 and I297. The amino group of the substrate is held by the side chain of D52 and by the main chain carbonyl of L13. The carboxylate is adequately positioned for reacting with the α -phosphoryl group of ATP, as suggested from comparison with MetRS structures bound to methionyl adenylate or analogues (Fig. 6, Crepin et al., 2003; Larson et al., 2011).

Such a precise positioning of the carboxylate group with respect to ATP was also observed in class 2 aminoacyl-tRNA synthetases (Schmitt et al., 1998).

Comparison of the MetRS: β^3 -Met structure with the MetRS: α -Met one showed very high similarity. The side chain of β^3 -Met was bound in the same induced pocket as that of α -Met, without any significant positional difference (Fig. 6 and Supplementary Fig. S2). In particular, the rearrangement of aromatic residues (W229, W253, F300 and F304) occurred identically. Also, the amino group was located at the same position as that of α -Met, bound to D52 and L13. However, because of the insertion of a methylene group, the carboxylate of β^3 -Met is positioned differently as compared to α -Met (Fig. 6). More precisely, two alternative positions differing by a rotation of about 87° around the C α -C β bond were observed for the carboxylate of β^3 -Met (Fig. 6 and Fig. S2). In the major conformation (70% occupancy), the carboxylate oxygens were oriented away from the active site. In this conformation, Y15 cannot reach the position where it locks the active site. Accordingly, Y15 was observed at several alternative positions, ranging from the open position seen in the unliganded enzyme to the closed position seen in the MetRS: α -Met complex (Fig. 6 and Fig. S2). In contrast, in

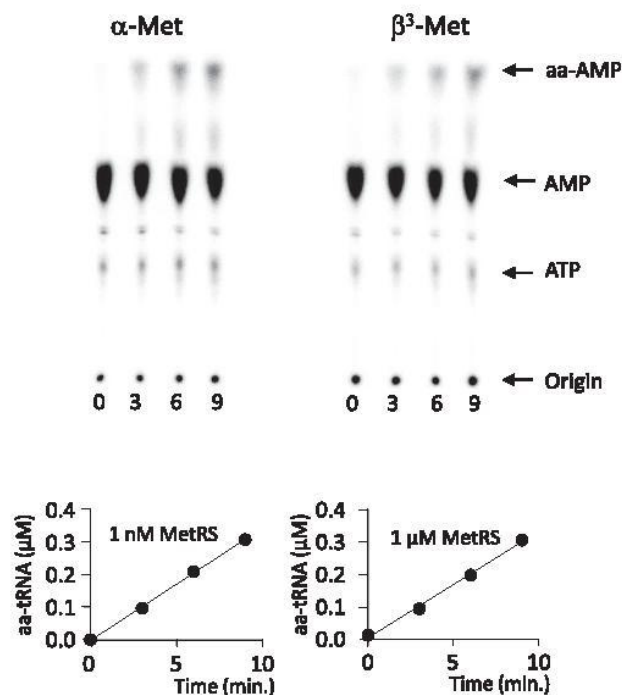


Fig. 5. Aminoacylation of tRNA^{Met} by MGL-treated β³-Met. [³²P]-labelled tRNA^{Met} was aminoacylated with α-Met in the presence of 1 nM MetRS or with MGL-treated β³-Met in the presence of 1 μM MetRS (see Materials and Methods). At the indicated times, aliquots were withdrawn and digested with S1 nuclease. The resulting digests were analyzed by TLC. The figure shows the results of a typical experiment with the fluorogram in the upper part and the corresponding image analysis in the bottom part. The figure shows that the rates differ by three orders of magnitude. The average rates from two independent experiments are reported in Table 1.

the minor conformation, one carboxylate oxygen of β³-Met is located only at 1.3 Å from the ideal position observed in α-Met and full rotation of Y15 is possible.

With α-Met, the electrostatic hindrance between the carboxylate group and the α-phosphoryl of ATP is known to be compensated by strong synergistic binding such that K_{α-Met} and K_{α-Met^{ATP}} have similar values (Table 1; Blanquet et al., 1975; Fayat et al., 1977). With β³-Met, the affinity of the amino acid was 4-fold higher in the presence of ATP, showing that the electrostatic repulsion was less intense. This agrees with the structural observation that the carboxylate of β³-Met is not adequately positioned to react with the α-phosphoryl of ATP.

Furthermore, in the β³-Met complex, the D296-Y325 couple has two alternative conformations. In the first conformation (hereafter termed “OUT”), the Y325 hydroxyl group formed a hydrogen bond with the side chain of D296. In the other conformation (hereafter termed “IN”), Y325 was shifted, causing a rotation of D296 towards the active site. The “OUT” conformation is found in the structures of unliganded MetRS and of the MetRS: α-Met complex, whereas the “IN” conformation is found in the structures of MetRS:methionyl adenylate analogues complexes (Crepin et al., 2003; Mechulam et al., 1999b; Schmitt et al., 2009; Serre et al., 2001). Notably, D296 is conserved in class 1 aaRS and plays an important role in the catalysis of aminoacyl adenylate synthesis (Brick et al., 1989; Crepin et al., 2003; Fersht, 1987). These constraints on D296 may therefore contribute to a reduced efficiency for β³-Met activation.

Overall, the structural results illustrate how the formation of aminoacyl adenylate from β³-Met and ATP can occur, though at a reduced rate, as observed in the kinetic measurements (Table 1).

2.6. Attempt to enhance activity of MetRS towards β³-Met

Interestingly, on the side of the amino acid-binding pocket opposite to the active centre, a supplementary β³-Met minor site with an occupancy of 0.7 was observed. The second β³-Met molecule (hereafter termed β³-Met₂) was bound between loop 295–297 connecting β D to αD on the one hand and helix α4 in the connective polypeptide (CP) on the other hand (Fig. 7). The secondary binding cavity is bordered by V214, K217, W221, K295, V298 and Y325. As described above, Y325 is connected to D296, thereby showing a physical link between the active site and the secondary β³-Met site. We therefore wondered whether occupation of the β³-Met₂ site may affect the activity of MetRS towards β³-Met. In order to impair β³-Met₂ binding, we mutated V298 into I. Measurement of the catalytic parameters of β³-Met activation showed a two-fold increase in affinity of β³-Met for the V298I mutant, both in the presence and in the absence of ATP (Table 1). This effect was however paralleled by a two-fold increase in α-Met affinity, at least in the absence of ATP. Therefore, selectivity of the V298I mutant towards β³-Met as compared to α-Met was not enhanced by a factor greater than 2.

The structure of the V298I mutant was solved in the apo form and bound to either α-Met or β³-Met (Fig. 7B,C,D). As expected, the β amino acid was found only in the main site. However, in the V298I: β³-Met structure, D296 was still observed in the “IN” and “OUT” conformations. Moreover, both conformations of D296 were also observed in the apo and α-Met-bound forms of the V298I mutant. Hence, occupation of the β³-Met₂ cavity by the Ile side-chain was unfortunately sufficient to cause destabilization of the D296 position, meaning that the mutation did not have the desired structural effect. However, in the apo form, alternative conformations for W229 and W253, two residues involved in the conformational rearrangement of aromatic residues induced by α-Met or β³-Met binding were observed (Supplementary Fig. S3). Therefore, the enhanced affinity of the V298I mutant for both β³-Met and α-Met may have been caused by a better propensity to adopt a conformation corresponding to the conformation induced by Met binding.

2.7. Concluding remarks

This study shows that *E. coli* MetRS is able to use β³-Met as a substrate in both the amino acid activation and tRNA^{Met} aminoacylation reactions. These results are in keeping with previous studies showing that several β amino acids, including β³-Met, could be transferred onto tRNA by the corresponding aminoacyl-tRNA synthetase (Hartman et al., 2007, 2006). The efficiency of the reaction appears however very low as compared to what was observed in a recent study (Melo Czekster et al., 2016). This discrepancy might be due to the presence of contaminating α-Met in commercial β³-Met samples. The high-resolution structure of the MetRS: β³-Met complex explains why the non-canonical amino acid has a rather good affinity for the enzyme but is activated *ca* 5000-fold more slowly. This is due to a non-optimal positioning of the carboxylate moiety, rendering reaction with the α-phosphoryl group of ATP difficult. β³-Met can also bind in a secondary site, though less efficiently as suggested by partial occupancy in the structure in spite of the high concentration of amino acid used for crystallization (20 mM). Disrupting the secondary site had little, if any, influence on the efficiency of MetRS towards β³-Met. Nevertheless, this study provides structural and functional foundations for MetRS engineering to obtain a more active and specific enzyme towards β³-Met. Such an enzyme would open the possibility of directing site-specific incorporation of β³-Met *in vivo*.

3. Materials and methods

3.1. Site-directed mutagenesis and enzyme purification

The gene encoding M547 from pBSM547+ (Mellot et al., 1989;

Table 2

Data collection and refinement statistics.

	MetRS: β^3 -Met	MetRS: α -Met	VI298	VI298: β^3 -Met	VI298: α -Met
Data collection					
Space group	P2 ₁	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Cell dimensions					
a, b, c (Å)	78.23,45.23,86.25	78.28,45.19,85.89	78.51,45.53,86.31	78.31,45.11,86.21	78.47,45.15,86.23
α , β , γ (°)	90,107.36,90	90,107.39,90	90,107.38,90	90,107.39,90	90,107.48,90
Resolution (Å)	1.45	1.2	1.5	1.48	1.38
R _{sym} (%) ^a	3.8 (16.0) ^b	10.5 (164)	7.6 (96.4)	11.7 (103)	6.8 (114)
I/ σ (I)	20.9 (7.9)	10.0 (1.1)	11.2(1.2)	9.2 (1.3)	15.9 (1.4)
Completeness (%)	98.4 (95.1)	99.0 (94.8)	99.2 (97.4)	99.6 (98.3)	99.1 (95.6)
Redundancy	3.3 (3.1)	6.4 (6.2)	4.7 (4.6)	6.8 (6.6)	6.8 (6.4)
CC _{1/2} (%) ^c	99.9 (96.9)	99.8 (32.6)	99.8 (95.2)	99.8 (59.7)	99.9 (52.3)
Refinement					
Resolution (Å)	41.7–1.45	33.1–1.2	38.9–1.5	48.6–1.48	48.5–1.38
No reflections	101,094	178,894	93,868	96,417	116,836
R _{work} /R _{free}	0.093/0.126 (0.092/0.136)	0.136/0.167 (0.328/0.354)	0.134/0.176 (0.282/0.337)	0.133/0.168 (0.319/0.353)	0.127/0.162 (0.289/0.323)
No atoms					
Protein	4696	5057	4692	4709	4638
Water	810	746	679	740	685
Non-protein ^e	50	41	40	75	60
B-factors (Å ²)					
Protein	13.1	16.7	22.2	17.7	19.2
Water Non-protein	33.2	35.6	39.5	35.8	36.1
	22.2	34.7	48.2	37.4	41.4
r.m.s. deviations					
Bond lengths (Å)	0.011	0.008	0.008	0.008	0.008
Bond angles (°)	1.304	0.949	0.994	0.947	0.940

A single crystal was used for data collection.

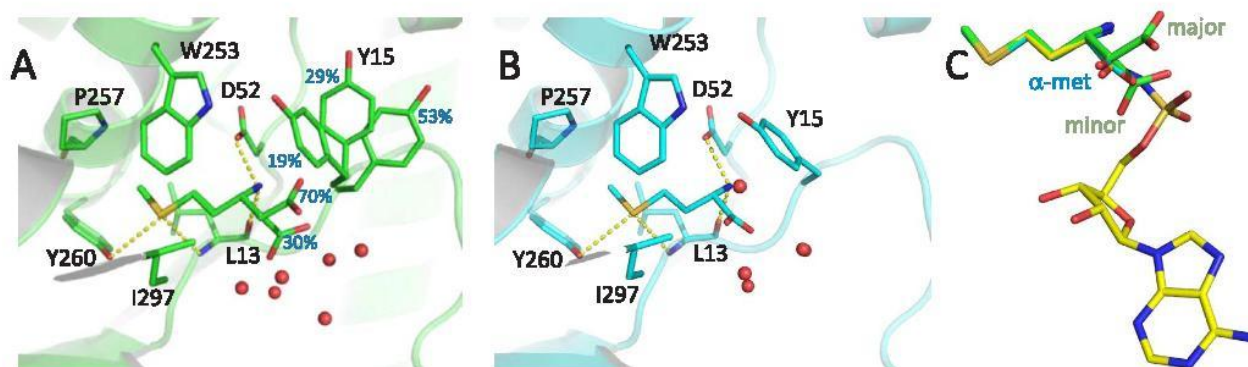
^a $R_{sym}(I) = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_i I_{hkl,i}}$, where i is the number of reflections hkl.^b Values in parentheses are for highest-resolution shell.^c CC(1/2) is the correlation coefficient between two random half data sets (Karplus and Diederichs, 2012).^d $R_{work} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$; R_{free} is calculated with 5% of the reflections.^e ligands and solvent molecules other than water (citrate and glycerol).

Fig. 6. Binding of β^3 -Met to MetRS (panel A) compared to the binding of α -Met (panel B). Relevant side chains are shown in sticks and nearby water molecules in red spheres. Occupancies of alternative conformations (%) are indicated in blue. MetRS is drawn as a ribbon. The structures of MetRS: β^3 -Met, MetRS: α -Met (this study) and MetRS:methionyl-sulfamoyl-adenosine (PDBid 1PFY; (Crepin et al., 2003; Larson et al., 2011)) were superimposed. Panel C compares the conformations of the ligands after superimposition. Major and minor conformations of the carboxylate of β^3 -Met are indicated. Hydrogen bonds are shown with yellow dots. Figs. 5 and 6 were drawn with PyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Schmitt et al., 1993) was subcloned into pET15b1pa (Guillon et al., 2005) in order to add an N-terminal His-tag to the MetRS product. The QuickChange method (Braman et al., 1996) was used to generate site-directed mutations. The M547 variants were produced in BLR(DE3) cells. Cultures were grown overnight at 37 °C in autoinducible TBAI medium containing 50 µg/mL of ampicillin. The cell pellet was suspended in purification buffer (10 mM Hepes-NaOH pH 7.5, 500 mM NaCl, 3 mM 2-mercaptoethanol) and cells were disrupted by ultrasonic disintegration. After centrifugation, the extract was loaded onto Talon affinity resin (Clontech). The resin was washed with purification buffer

containing 10 mM imidazole and the protein eluted by increasing imidazole concentration to 125 mM. The eluate was diluted ten-fold before further purification on a Q-Hiload (GE-Healthcare) ion exchange column. The procedure yielded 15–20 mg of homogeneous protein for a 250 mL culture. Purified proteins were stored either at 4 °C in 10 mM Hepes-NaOH, 10 mM 2-mercaptoethanol or at –20 °C in the same buffer plus 55% glycerol.

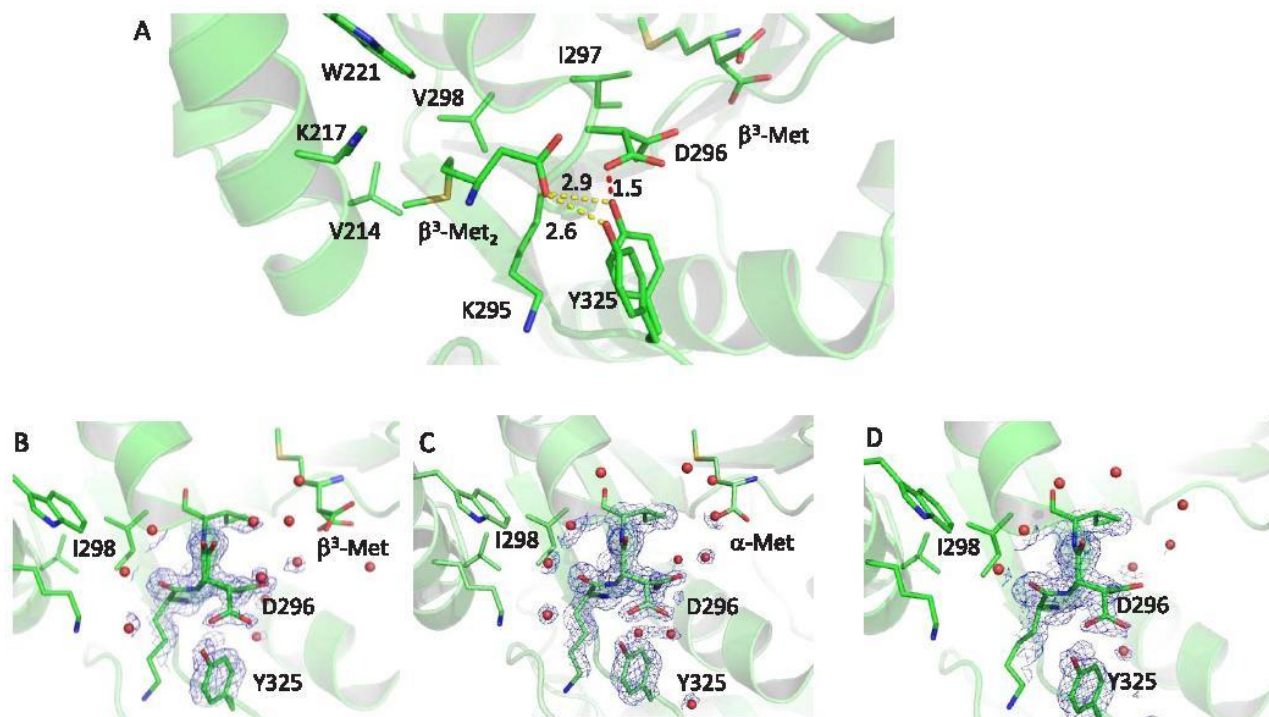


Fig. 7. Secondary β^3 -Met binding site (β^3 -Met₂) in MetRS. MetRS is shown as a transparent ribbon. Panel A shows the environment of β^3 -Met₂. β^3 -Met in the active site is also represented. Relevant side chains are shown in sticks and nearby water molecules in red spheres. The two alternative conformations of D296 and Y325 (see text) are shown. Several distances represented by yellow or red dots are indicated in Å. The bottom panels show the region of D296 in the V298I mutant as a complex with β^3 -Met (panel B), α -Met (panel C) or in the apo form (panel D). The relevant part of the electron density (2mFo-DFc map contoured at 1.0 σ) is represented in blue using the carve command in PyMOL. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Isotopic [32 P]PPI-ATP exchange

Amino acid-dependent [32 P]PPI-ATP exchange activity was measured at 25 °C in standard buffer (20 mM Tris-HCl pH 7.6, 7 mM MgCl₂, 10 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 2 mM ATP and 2 mM [32 P] PPI as described (Blanquet et al., 1974; Schmitt et al., 1994). For Km measurements, concentrations of α -Met or β^3 -Met were varied between 8 μ M and 2 mM or 0.125 mM to 8 mM, respectively.

3.3. Treatment of β^3 -Met with methionine gamma lyase (MGL)

L- β^3 -Met was purchased from Fluorochem Ltd. (UK), dissolved in standard buffer and adjusted to pH 7.5 with NaOH. For α,γ -elimination of contaminating methionine to α -ketobutyrate, methanethiol and ammonia, 100 μ L batches of β^3 -Met (160 mM) were incubated overnight at 37 °C with 1 unit of *Pseudomonas putida* MGL recombinantly expressed in *E. coli* (Biovision Inc., Milpitas, USA) in the presence of 150 μ M phosphoenolpyruvate. A control sample was made by omitting MGL; this control treatment had no effect on the rates of [32 P]PPI-ATP isotopic exchange.

3.4. tRNA aminoacylation assay

Aminoacylation of *E. coli* tRNA^{Met} produced and purified as described (Mechulam et al., 2007) was assayed using the [32 P]-labelling method (Ledoux and Uhlenbeck, 2008). tRNA^{Met} was labelled by exchange of the 3' terminal adenosine (Ledoux and Uhlenbeck, 2008) catalyzed by *E. coli* tRNA nucleotidyltransferase overexpressed from a recombinant plasmid (kind gift from Dr. Kozo Tomita, University of Tokyo, Japan). The His-tagged nucleotidyltransferase was purified by affinity on Talon resin. The labelling reaction (50 μ L) contained tRNA (1 μ M), sodium pyrophosphate (50 μ M), MgCl₂ (10 mM), glycine (pH 9;

50 mM), [α - 32 P] ATP (0.3 μ M, 111 TBq/mmol) and 0.02 μ M nucleotidyltransferase. After 5 min at 37 °C, 5 μ L of a solution containing 1 μ M CTP and 0.1 U of pyrophosphatase were added. After 2 min at 37 °C, the reaction was phenol and then ether extracted. The aqueous phase was run over a Micro Bio-Spin P-6 molecular sieving column (Bio-Rad) to remove nucleotides and in particular unreacted ATP. After ethanol precipitation, labelled tRNA was redissolved in standard buffer containing 7 mM MgCl₂ and 20 μ M unlabelled tRNA.

Aminoacylation activity was measured in 20 μ L reactions at 25 °C in 20 mM Tris-HCl (pH 7.6), 7 mM MgCl₂, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 7 mM MgCl₂, 150 mM KCl, in the presence of 2 mM ATP, 2 μ M [32 P] tRNA^{Met} and 2 mM of either α -Met or MGL-treated β^3 -Met. M547 enzyme diluted in reaction buffer containing 200 μ g/mL BSA was used at a concentration of 1 nM in the case of α -Met and 1 μ M in the case of β^3 -Met. At desired incubation times, 3 μ L of the reaction mixture were withdrawn and mixed with an equal volume of S1 nuclease (NEN from laboratory stocks, 2 units/ μ L in 0.3 M Na-acetate, pH 4.8). The digestion products were run on a PEI cellulose TLC plate (Bio-Rad) in glacial acetic acid/1 M NH₄Cl/water (5:10:85). Plates were phosphorimaged with a Typhoon scanner (GE Healthcare). Images were processed with ImageJ (Schneider et al., 2012) in order to determine for each incubation time the fraction of aminoacylated tRNA, from which aminoacylation rates were derived. A typical experiment is shown in Fig. 5.

3.5. Fluorescence at equilibrium

Variations of the intrinsic fluorescence of M547 and its variants (0.5 μ M) upon titration with substrates were followed at 25 °C in 20 mM Tris-HCl (pH 7.6), 10 mM 2-mercaptoethanol, 2 mM MgCl₂ and 0.1 mM EDTA as described (Mechulam et al., 1991; Schmitt et al., 1994). Measurements were done in a Hellma 1 cm $\times$ 0.4 cm cuvette with

an FP-8300 JASCO spectrofluorometer ($\lambda_{exc} = 295$ nm, $\lambda_{em} = 340$ nm). All titration curves were corrected for dilution by multiplying the measured fluorescence by the ratio of the current volume in the cuvette over the initial volume before titration. Concentrations of α -Met or β^3 -Met (untreated or MGL-treated) were varied from 3 μ M to 1 mM and from 0.06 mM to 4 mM, respectively. Data were fitted to simple saturation curves from which the corresponding dissociation constants were derived using the Origin software (OriginLab Corp.; see Fig. 2).

3.6. Fluorescence at the pre-steady state

Fluorescence measurements at the pre-steady-state were performed as described (Hyafil et al., 1976; Schmitt et al., 1994) using an SX20 stopped flow apparatus (Applied Photophysics, UK). All experiments were performed in 20 mM Tris-HCl (pH 7.6), 10 mM 2-mercaptoethanol, 2 mM $MgCl_2$ and 0.1 mM EDTA. The formation of α - or β^3 -methionyl adenylate was initiated by mixing 1:1 (v/v) an enzyme solution (1 μ M for α -Met or 2 μ M for β^3 -Met) containing ATP- Mg^{2+} (2 mM) and PPi (10 μ M) with a solution containing the same concentrations of ATP- Mg^{2+} plus variable amounts of the amino acid (10 μ M to 640 μ M for α -Met or 25 μ M to 2 mM for β^3 -Met). For the reverse reaction, enzyme:adenylate complexes were pre-formed by incubating the enzyme with ATP- Mg^{2+} and the amino acid for a few minutes at 25 °C (1 μ M enzyme, 0.1 mM ATP- Mg^{2+} and 500 μ M α -Met or 2 μ M enzyme, 2 mM ATP- Mg^{2+} and 25 μ M β^3 -Met). The solution was then mixed 1:1 (v/v) with a solution containing the same concentrations of ATP- Mg^{2+} and amino acid plus variable amounts of pyrophosphate (20 μ M to 1.28 mM for α -Met or 20 μ M to 2 mM for β^3 -Met). After mixing, fluorescence was recorded and fitted to single exponentials from which the rate constants were derived (Supplementary Fig. S1). Each rate was determined three times. Kinetic parameters (k_f and k_b , see Scheme 1) and equilibrium parameters (K_{Met}^{ATP} and K_{PP}) were deduced from the fit of the measured rate constants to the theoretical saturation curves (Hyafil et al., 1976; Mechulam et al., 1991; Schmitt et al., 1994) using Origin. Each experiment was performed at least twice independently. Results in Table 1 are expressed as mean $\pm$ either standard deviation from the independent experiments or standard error from the fitting procedure, whichever the greater. All experiments were performed using MGL-treated β^3 -Met.

3.7. Preparation of samples for mass spectrometry experiments

Amino acid samples (160 mM α -Met, β -Met or MGL-treated β -Met) were acidified by adding formic acid (10% v/v final concentration). Samples were then diluted 1000-fold in 0.1% formic acid prior to analysis. Aminoacyl adenylate samples were prepared in 0.8 mL solutions (20 mM Tris-HCl pH 7.6, 2 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.1 mM EDTA) containing 0.5 μ M MetRS, 1 mM MGL-treated β -Met and 2 mM ATP- Mg^{2+} . The reaction at 25 °C was monitored by following the fluorescence as described above. After ca 15 min, reaction was nearly complete and 10% (v/v) formic acid and precipitated material was removed by centrifugation prior to analysis. Analysis of tRNA aminoacylation after S1 nuclease digestion was performed essentially as described (Hartman et al., 2007, 2006). tRNA^{Met} was first aminocyclated in 1 mL samples containing 20 mM Tris-HCl (pH 7.6), 7 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 2 mM ATP, 1 μ M MetRS and either 0.4 mM of α -Met or 1 mM of β -Met. After incubation for 30 min at 25 °C, the reaction was divided into 4 samples each of which was precipitated with ethanol. One sample was dissolved in 100 μ L of 3 mM Na-acetate (pH 5.5) containing 1 mM zinc acetate and tRNA was digested with 10 units of S1 nuclease for 20 min at 37 °C. Formic acid (10% v/v final concentration) was then added before mass spectrometry analysis. Another sample was submitted to acetylation as follows. After dissolution of the pellet in 100 μ L of 5 mM sodium acetate (pH 5.5), 100 μ L DMSO, 20 μ L of glacial acetic acid and 20 μ L of acetic

anhydride were added. After a 30 min incubation on ice, the solution was ethanol precipitated and tRNA was digested with S1 nuclease, as described above.

3.8. Mass spectrometry

Chromatographic grade solvents (99.99% purity), acetonitrile (ACN) and formic acid (FA), were purchased from Sigma Aldrich. Liquid chromatography/ high-resolution mass spectrometry (LC-HRMS) analyses were performed with the timsTOF mass spectrometer coupled with an Elute HPLC system (Bruker Daltonics, Bremen, Germany).

For all experiments, 10 μ L of the sample were injected and separated on an Atlantis T3 column (3 μ m, 150 $\times$ 2.1 mm; Waters, Saint Quentin, France). The effluent was introduced at a flow rate of 0.2 mL min⁻¹ into the interface with a gradient increasing from 10% of solvent B to 90% in 8 min to achieve 100% at 10 min (A: water with 0.1% formic acid (FA); B: acetonitrile with 0.1% FA). From 10 min to 12 min, the percentage of solvent increased until 90%. The gradient was then set at 10% of B for the last 6 min.

Electrospray ionization was operated in the positive ion mode. Capillary and end plate voltages were set at -4.5 kV and -0.5 kV, respectively. Nitrogen was used as the nebulizer and drying gas at 2 bar and 8 L/min, respectively, with a drying temperature of 220 °C.

In MS/MS experiments, the precursor ion was selected with an isolation window of 1 Da and the collision induced dissociation was performed using collision energies (E_{col}) ranging from 7 to 25 eV. Tuning mix (Agilent, France) was used for calibration. The elemental compositions of all ions were determined with the instrument software Data Analysis. The precision of mass measurement was better than 5 ppm.

3.9. Crystallization and structure determination

Crystals of M547 or of its variants were obtained in a solution containing 30 mM KPO₄, 4 mM 2-mercaptoethanol, 1.08 mM ammonium citrate (pH 7.0) and 3.6 mg/mL of protein by microseeding with crystals of M547 (Mechulam et al., 1999b). For ligand binding studies, the amino acid was added to the crystallization medium prior to microseeding. Final concentrations during crystallization were 2 mM for α -Met or 20 mM for β^3 -Met. For data collection, crystals were quickly soaked in a solution containing 1.4 M ammonium citrate, 30 mM potassium phosphate (pH 7.0) and 25% v/v of glycerol before flash-cooling in liquid nitrogen. Data were collected at the Proxima 1 and Proxima 2 beamlines at SOLEIL synchrotron (Gif sur Yvette, France). Diffraction images were analyzed with the XDS program (Kabsch, 1988; Collaborative Computational Project n°4, 1994) and the data further processed using programs from the CCP4 package (Collaborative Computational Project n°4, 1994). Data statistics are summarized in Table 1. Each structure was solved by rigid body refinement of the M551 model (PDB id 1QQT), using PHENIX (Adams et al., 2010). Coordinates and associated B factors were refined through several cycles of manual adjustments with Coot (Emsley et al., 2010) and positional refinement with PHENIX. Refinement statistics are summarized in Table 1.

4. Accession numbers

Models coordinates and reflection files were deposited at the PDB with identification numbers 6SPN (MetRS- β^3 Met), 6SPO (MetRS- α Met), 6SPP (MetRS_V298I), 6SPQ (MetRS_V298I- α Met), 6SPR (MetRS_V298I- β^3 Met).

CRediT authorship contribution statement

Giuliano Nigro: Investigation, Methodology, Writing - review & editing. Sophie Bourcier: Investigation, Methodology, Writing - review

& editing. **Christine Lazennec-Schurdevin**: Investigation, Writing - review & editing. **Emmanuelle Schmitt**: Conceptualization, Methodology, Supervision, Writing - original draft, Writing - review & editing. **Philippe Marlière**: Conceptualization, Methodology, Supervision, Writing - review & editing. **Yves Mechulam**: Conceptualization, Methodology, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by grants from the Centre National de la Recherche Scientifique (France) and by Ecole polytechnique (France) to Unité Mixte de Recherche n°7654. GN was a recipient of a PhD scholarship from Ecole polytechnique. We thank the staff of the macromolecular crystallography PX1 and PX2 beamlines at the SOLEIL synchrotron (Saclay, France) for expert assistance during data collection. Kozo Tomita (University of Tokyo, Japan) is acknowledged for the kind gift of the *E. coli* tRNA nucleotidyltransferase-overproducing plasmid.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jsb.2019.107435>.

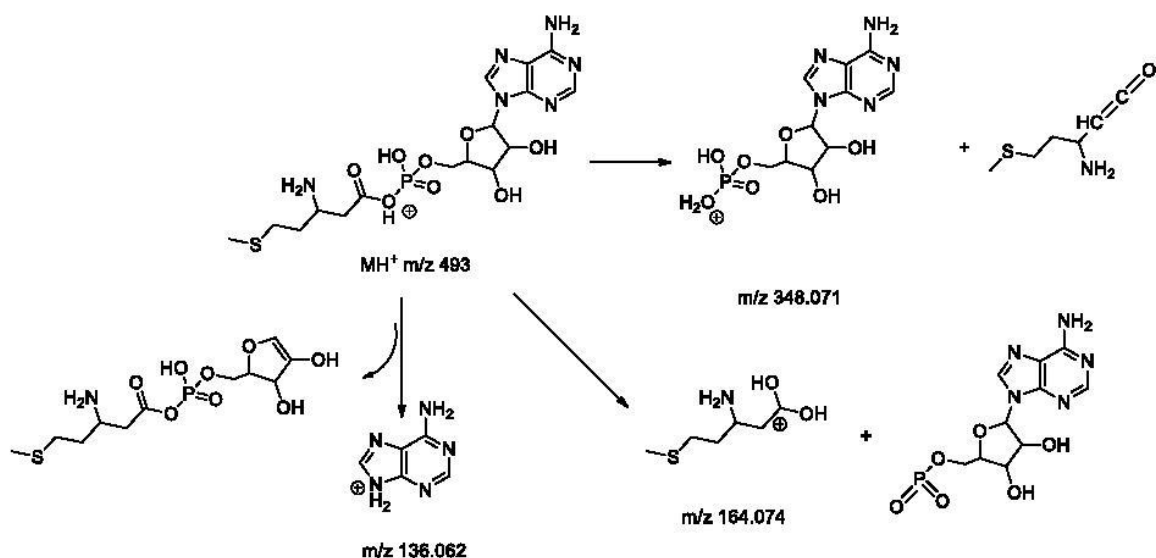
References

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Section D, Biol. Crystallogr.* 66, 213–221.
- Blanquet, S., Fayat, G., Waller, J.P., 1974. The mechanism of action of methionyl-tRNA synthetase from *Escherichia coli*. Mechanism of the amino-acid activation reaction catalyzed by the native and the trypsin-modified enzymes. *Eur. J. Biochem.* 44, 343–351.
- Blanquet, S., Fayat, G., Poiret, M., Waller, J.P., 1975. The mechanism of action of methionyl-tRNA synthetase from *Escherichia coli*. Inhibition by adenosine and 8-aminoadenosine of the amino-acid activation reaction. *Eur. J. Biochem.* 51, 567–571.
- Braman, J., Papworth, C., Greener, A., 1996. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol. Biol.* (Clifton, N.J.) 57, 31–44. <https://doi.org/10.1385/0-89603-332-5:31>.
- Brick, P., Bhat, T.N., Blow, D.M., 1989. Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.* 208, 83–98.
- Cheloha, R.W., Sullivan, J.A., Wang, T., Sand, J.M., Sidney, J., Sette, A., Cook, M.E., Suresh, M., Gellman, S.H., 2015. Consequences of periodic alpha-to-beta(3) residue replacement for immunological recognition of peptide epitopes. *ACS Chem. Biol.* 10, 844–854. <https://doi.org/10.1021/cb500888q>.
- Chin, J.W., 2017. Expanding and reprogramming the genetic code. *Nature* 550, 53–60. <https://doi.org/10.1038/nature24031>.
- Collaborative Computational Project n°4, 1994. The CCP4 suite: programs from protein crystallography. *Acta Crystallogr. D* 50, 760–763.
- Crepin, T., Schmitt, E., Mechulam, Y., Sampson, P.B., Vaughan, M.D., Honek, J.F., Blanquet, S., 2003. Use of analogues of methionine and methionyl adenylate to sample conformational changes during catalysis in *Escherichia coli* methionyl-tRNA synthetase. *J. Mol. Biol.* 332, 59–72. [https://doi.org/10.1016/S0022-2836\(03\)00917-3](https://doi.org/10.1016/S0022-2836(03)00917-3).
- Dedkova, L.M., Fahmi, N.E., Paul, R., del Rosario, M., Zhang, L., Chen, S., Feder, G., Hecht, S.M., 2011. β -Puromycin selection of modified ribosomes for in vitro incorporation of β -amino acids. *Biochemistry* 51, 401–415. <https://doi.org/10.1021/bi2016124>.
- Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. *Acta Crystallogr. D* 66, 486–501.
- Fayat, G., Fromant, M., Blanquet, S., 1977. Couplings between the sites for methionine and adenosine 5'-triphosphate in the amino acid activation reaction catalyzed by trypsin-modified methionyl-transfer RNA synthetase from *Escherichia coli*. *Biochemistry* 16, 2570–2579.
- Fersht, A.R., 1987. Dissection of the structure and activity of the tyrosyl-tRNA synthetase by site-directed mutagenesis. *Biochemistry* 26, 8031–8037.
- Fersht, A.R., Dingwall, C., 1979. An editing mechanism for the methionyl-tRNA synthetase in the selection of amino acids in protein synthesis. *Biochemistry* 18, 1250–1256.
- Fujino, T., Goto, Y., Suga, H., Murakami, H., 2016. Ribosomal synthesis of peptides with multiple beta-amino acids. *J. Am. Chem. Soc.* 138, 1962–1969. <https://doi.org/10.1021/jacs.5b12482>.
- Fukumoto, M., Kudou, D., Murano, S., Shiba, T., Sato, D., Tamura, T., Harada, S., Inagaki, K., 2014. The role of amino acid residues in the active site of L-methionine γ -lyase from *Pseudomonas putida*. *Biosci. Biotechnol. Biochem.* 76, 1275–1284. <https://doi.org/10.1271/bbb.110906>.
- Guillon, L., Schmitt, E., Blanquet, S., Mechulam, Y., 2005. Initiator tRNA binding by eIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. *Biochemistry* 44, 15594–15601. <https://doi.org/10.1021/bi051514j>.
- Hartman, M.C.T., Josephson, K., Szostak, J.W., 2006. Enzymatic aminoacylation of tRNA with unnatural amino acids. *PNAS* 103, 4356–4361. <https://doi.org/10.1073/pnas.0509219103>.
- Hartman, M.C.T., Josephson, K., Lin, C.-W., Szostak, J.W., 2007. An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS One* 2, e972–15. <https://doi.org/10.1371/journal.pone.0000972>.
- Heck, T., Limbach, M., Geueke, B., Zacharias, M., Gardiner, J., Kohler, H.-P.E., Seebach, D., 2006. Enzymatic degradation of beta- and mixed alpha, beta-oligopeptides. *Chem. Biodivers.* 3, 1325–1348. <https://doi.org/10.1002/cbdv.200690136>.
- Hyafil, F., Jacques, Y., Fayat, G., Fromant, M., Dessen, P., Blanquet, S., 1976. Methionyl-tRNA synthetase from *Escherichia coli*: active stoichiometry and stopped-flow analysis of methionyl adenylate formation. *Biochemistry* 15, 3678–3685.
- Johnson, J.A., Lu, Y.Y., Van Deventer, J.A., Tirrell, D.A., 2010. Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr. Opin. Chem. Biol.* 14, 774–780. <https://doi.org/10.1016/j.cbpa.2010.09.013>.
- Kabsch, W.J., 1988. Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J. Appl. Crystallogr.* 21, 916–924.
- Karplus, P.A., Diederichs, K., 2012. Linking crystallographic model and data quality. *Science* (New York, N.Y.) 336, 1030–1033. <https://doi.org/10.1126/science.1218231>.
- Katoh, T., Suga, H., 2018. Ribosomal incorporation of consecutive beta-amino acids. *J. Am. Chem. Soc.* 140, 12159–12167. <https://doi.org/10.1021/jacs.8b07247>.
- Larson, E.T., Kim, J.E., Zucker, F.H., Kelley, A., Mueller, N., Napuli, A.J., Verlinde, C.L.M.J., Fan, E., Buckner, F.S., Van Voorhis, W.C., Merritt, E.A., Hol, W.G.J., 2011. Structure of *Leishmania* major methionyl-tRNA synthetase in complex with intermediate products methionyladenylate and pyrophosphate. *Biochimie* 93, 570–582. <https://doi.org/10.1016/j.biochi.2010.11.015>.
- Ledoux, S., Uhlenbeck, O.C., 2008. [32 P]-labeling tRNA with nucleotidyltransferase for assaying aminoacylation and peptide bond formation. *Methods* 44, 74–80. <https://doi.org/10.1016/j.jymeth.2007.08.001>.
- Liu, C.C., Schultz, P.G., 2010. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444. <https://doi.org/10.1146/annurev.biochem.052308.105821>.
- Mahdavi, A., Hamblin, G.D., Jindal, G.A., Bagert, J.D., Dong, C., Sweredoski, M.J., Hess, S., Schuman, E.M., Tirrell, D.A., 2016. Engineered aminoacyl-tRNA synthetase for cell-selective analysis of mammalian protein synthesis. *J. Am. Chem. Soc.* 138, 4278–4281. <https://doi.org/10.1021/jacs.5b08980>.
- Maini, R., Nguyen, D.T., Chen, S., Dedkova, L.M., Chowdhury, S.R., Alcalá-Torano, R., Hecht, S.M., 2013. Incorporation of β -amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorg. Med. Chem.* 21, 1088–1096. <https://doi.org/10.1016/j.bmc.2013.01.002>.
- Mechulam, Y., Dardel, F., Le Corre, D., Blanquet, S., Fayat, G., 1991. Lysine 335, part of the KMSKS signature sequence, plays a crucial role in the amino acid activation catalyzed by the methionyl-tRNA synthetase from *Escherichia coli*. *J. Mol. Biol.* 217, 465–475.
- Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., Blanquet, S., 1999. Crystal structure of *Escherichia coli* methionyl-tRNA synthetase highlights species-specific features. *J. Mol. Biol.* 294, 1287–1297. <https://doi.org/10.1006/jmbi.1999.3339>.
- Mechulam, Y., Guillon, L., Yatime, L., Blanquet, S., Schmitt, E., 2007. Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. *Methods Enzymol.* 430, 265–281. [https://doi.org/10.1016/S0076-6879\(07\)30011-6](https://doi.org/10.1016/S0076-6879(07)30011-6).
- Mellot, P., Mechulam, Y., Le Corre, D., Blanquet, S., Fayat, G., 1989. Identification of an amino acid region supporting specific methionyl-tRNA synthetase: tRNA recognition. *J. Mol. Biol.* 208, 429–443.
- Melo Czekster, C., Robertson, W.E., Walker, A.S., Söll, D., Schepartz, A., 2016. In vivo biosynthesis of a β -amino acid-containing protein. *J. Am. Chem. Soc.* 138, 5194–5197. <https://doi.org/10.1021/jacs.6b01023>.
- Murakami, H., Kourouklis, D., Suga, H., 2003. Using a solid-phase ribozyme aminoacylation system to reprogram the genetic code. *Chem. Biol.* 10, 1077–1084.
- Murakami, H., Ohta, A., Ashigai, H., Suga, H., 2006. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* 3, 357–359. <https://doi.org/10.1038/nmeth877>.
- Ngo, J.T., Champion, J.A., Mahdavi, A., Tanrikulu, I.C., Beatty, K.E., Connor, R.E., Yoo, T.H., Dieterich, D.C., Schuman, E.M., Tirrell, D.A., 2009. Cell-selective metabolic labeling of proteins. *Nat. Chem. Biol.* 5, 715–717.
- Ohuchi, M., Murakami, H., Suga, H., 2007. The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Curr. Opin. Chem. Biol.* 11, 537–542. <https://doi.org/10.1016/j.cbpa.2007.08.011>.
- Petersson, E.J., Schepartz, A., 2008. Toward beta-amino acid proteins: design, synthesis, and characterization of a fifteen kilodalton beta-peptide tetramer. *J. Am. Chem. Soc.* 130, 821–823. <https://doi.org/10.1021/ja077245x>.
- Petersson, E.J., Craig, C.J., Daniels, D.S., Qiu, J.X., Schepartz, A., 2007. Biophysical characterization of a beta-peptide bundle: comparison to natural proteins. *J. Am.*

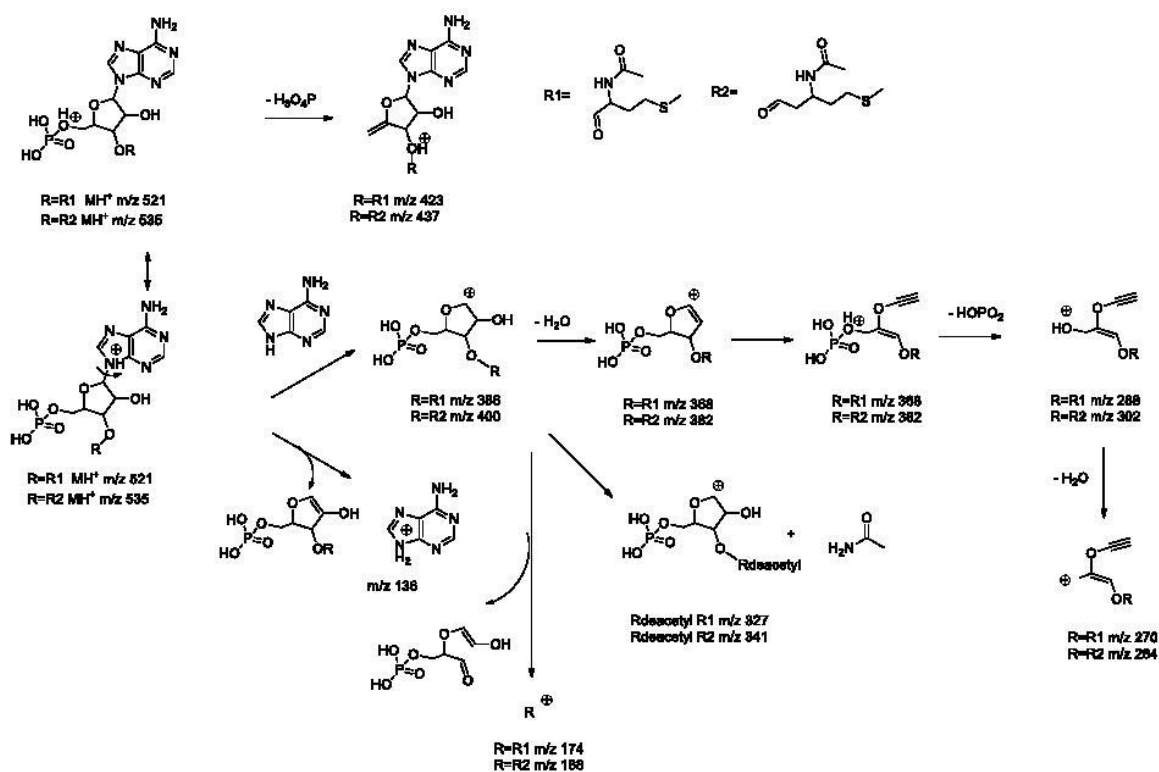
SUPPLEMENTARY MATERIAL

Use of β^3 -methionine as an amino acid substrate of *Escherichia coli* Methionyl-tRNA synthetase

Giuliano Nigro, Sophie Bourcier, Christine Lazennec-Schurdevin, Emmanuelle Schmitt, Philippe Marlière and Yves Mechulam



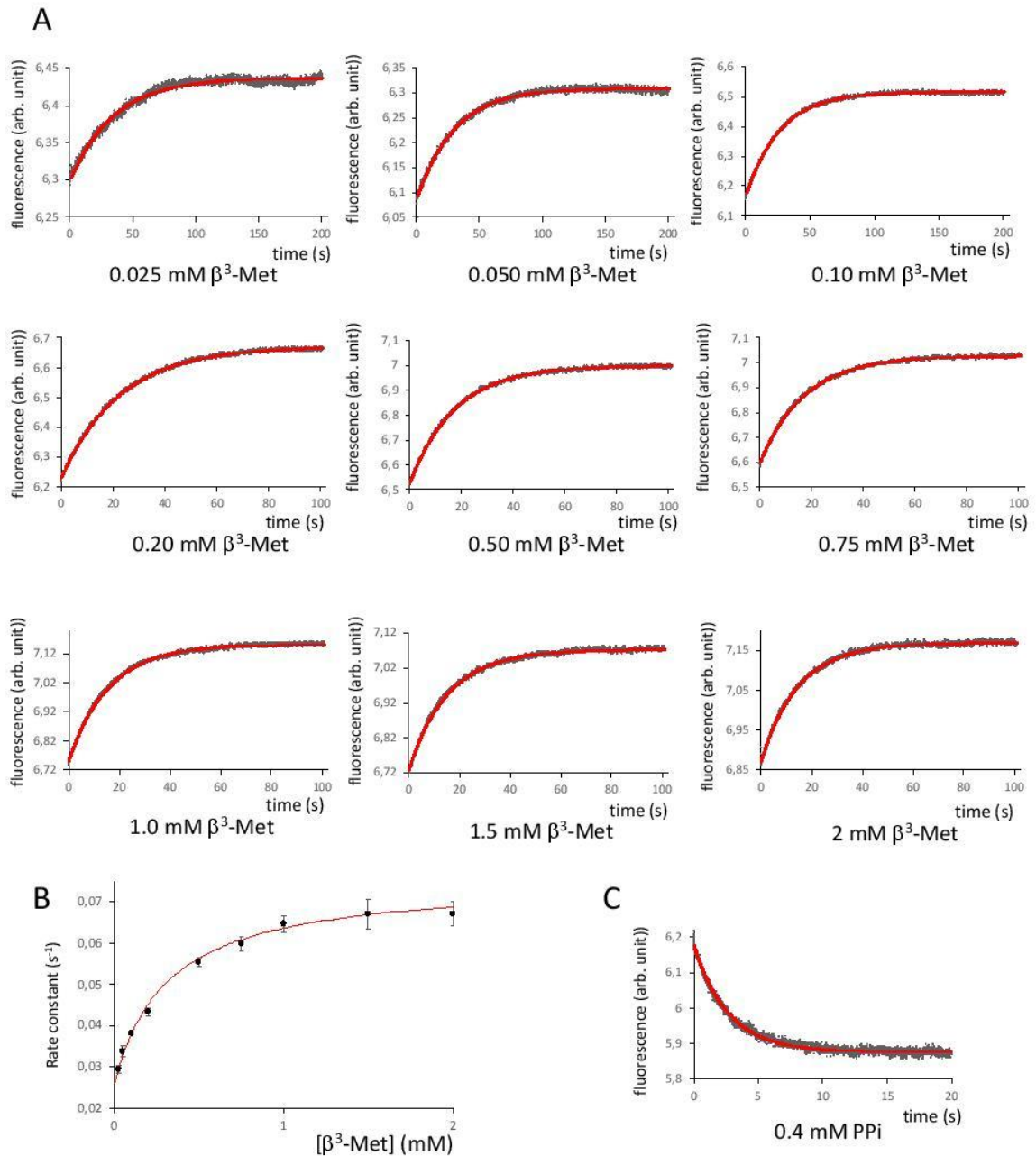
Supplementary Scheme 1. Proposed mechanism for the decomposition of m/z 493 ion [β^3 -Met adenylylateH] $^+$ observed by MS/MS (See Fig. 3B).



Supplementary Scheme 2. Proposed mechanism for the decomposition of [N-acetyl - α -methionyl-AMPH] $^+$ ($R=R1$, m/z 521) and [N-acetyl - β -methionyl-AMPH] $^+$ ($R=R2$, m/z 535) observed by LC-HR-MS/MS.

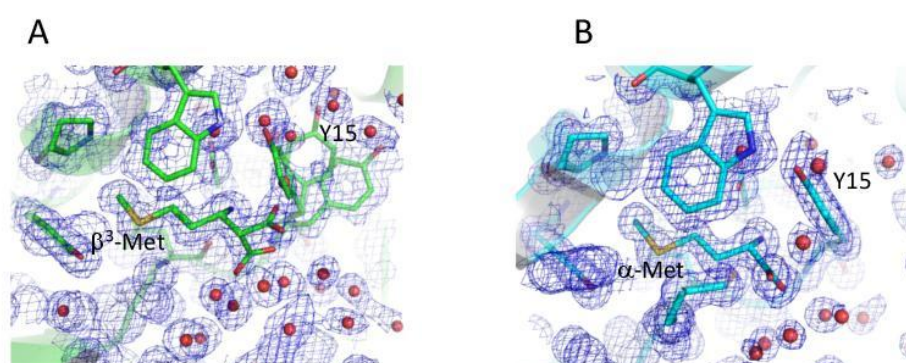
Compound	Ion (<i>m/z</i>) measured mass	Elemental composition of ion	Err (ppm)	Rdb
R _{β3} -Met	535.1378	C ₁₈ H ₂₈ N ₆ O ₉ PS	-1.4	12
R _α -Met	521.1219	C ₁₇ H ₂₆ N ₆ O ₉ PS	-0.9	12
R _{β3} -Met	400.0831	C ₁₃ H ₂₃ NO ₉ SP	-1.3	7
R _α -Met	386.0672	C ₁₂ H ₂₁ NO ₉ SP	-0.8	7
R _{β3} -Met	437.1586	C ₁₈ H ₂₅ N ₆ O ₅ S	-3.3	12
R _α -Met	423.1447	C ₁₇ H ₂₃ N ₆ O ₅ S	0.1	12
R _{β3} -Met	382.0717	C ₁₃ H ₂₁ NO ₈ SP	0.8	8
R _α -Met	368.0563	C ₁₂ H ₁₉ NO ₈ PS	0.3	8
R _{β3} -Met	341.0439	C ₁₁ H ₁₈ O ₈ PS	-1.1	11
R _α -Met	no			
R _{β3} -Met	284.0943	C ₁₃ H ₁₈ NO ₄ S	-1.0	8
R _α -Met	270.0797	C ₁₂ H ₁₆ NO ₄ S	-0.9	8
R _{β3} -Met	188.0743	C ₈ H ₁₄ NO ₂ S	-1.8	5
R _α -Met	174.0576	C ₇ H ₁₂ NO ₂ S	4.6	5
R _{β3} -Met / R _α -Met	136.0614	C ₅ H ₆ N ₅	-0.7	6
R _{β3} -Met/ R _α -Met	97.0286	C ₅ H ₅ O ₂	-1.8	4

Supplementary Table 1. HR-MS/MS of protonated N-acetyl- α -methionyl-AMP (*m/z* 521) and protonated N-acetyl- β^3 -methionyl-AMP (*m/z* 535) obtained at 2.6 min in LC-HR-MS/MS : Elemental composition of molecular and fragment ions with deviation between experimental and theoretical masses (Err) and the number of rings and double bonds in the formula (Rdb).

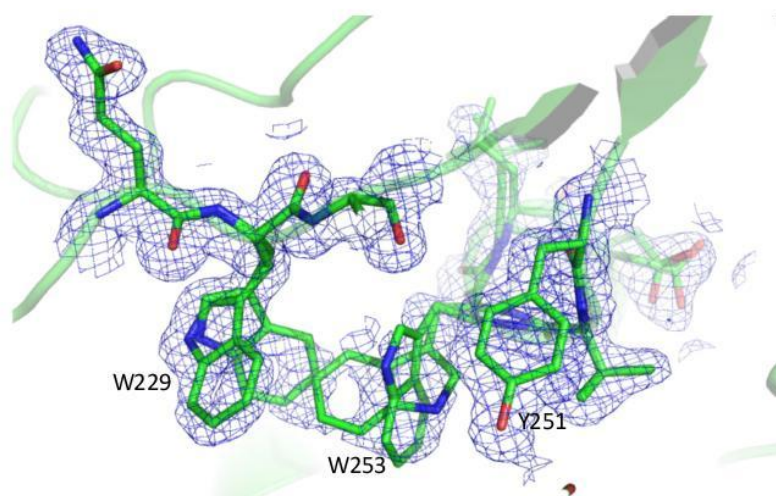


Supplementary Figure S1 : (A) Time course of β^3 -methionyl adenylate formation in the presence of 1 μ M MetRS, 2 mM ATP-Mg $^{2+}$, 10 μ M PPI and the indicated concentrations of β^3 -Met. Representative stopped-flow experiments (see Materials and Methods) are shown. The graphs show the recorded intrinsic enzyme fluorescence traces drawn in grey and the fitted exponential curves drawn with red lines. The corresponding rate constants were 0.029 s^{-1} (0.025 mM β^3 -Met), 0.034 s^{-1} (0.050 mM β^3 -Met), 0.038 s^{-1} (0.1mM β^3 -Met), 0.044 s^{-1} (0.2 mM β^3 -Met), 0.057 s^{-1} (0.5 mM β^3 -Met), 0.058 s^{-1} (0.75 mM β^3 -Met), 0.063 s^{-1} (1.0 mM β^3 -Met), 0.066 s^{-1} (1.5 mM β^3 -Met), 0.072 s^{-1} (2.0 mM β^3 -Met). Rate constants were averaged over three measurements and fitted to the theoretical equation (Hyafil et al., 1976). (B) Fit of the rate constants (black dots) to the saturation curve (red line). Deduced values for this series of experiments were k_{fp} =0.049 s^{-1} and $K_{\beta^3\text{-Met}}^{ATP}$ =0.31 mM. (C) Typical experiment showing the

pyrophosphorolysis of enzyme bound β^3 -methionyl adenylate. MetRS was preincubated with ATP-Mg²⁺ (2 mM) and β^3 -Met (25 μ M) to allow formation of adenylate, and then mixed 1:1 with a solution containing the same concentrations of ATP-Mg²⁺ and β^3 -Met, plus 0.8 mM PPI (0.4 mM final concentration after mixing). The graph shows the enzyme fluorescence trace drawn in grey and the fitted exponential curve drawn with a red line. The corresponding rate constant was 0.38 s⁻¹.



Supplementary Figure S2 : Electron density maps in the active site for MetRS:β³-Met (A) and MetRS:α-Met (B). Each panel shows the final (2mFo-DFc) maps contoured at 0.6 sigma (panel A, in order to properly visualize alternative conformations) or 1.0 sigma (panel B). The models are drawn with sticks and water molecules are shown as red spheres. Supplementary Figures S2 and S3 were drawn with Pymol. The figure shows the alternative conformation of the carboxylate of the amino acid ligand and of Y15 in the MetRS:β³-Met structure.



Supplementary Figure S3 : Alternative conformations of W229 and W253 in the structure of the V298I mutant. The model is shown in cartoon, with the relevant residues as sticks. The final (2mFo-DFc) electron density map contoured at 1.0 sigma was drawn using the carve command in Pymol.

Partie II

Afin d'améliorer l'activité de l'enzyme pour la β -méthionine, plusieurs stratégies ont été mises en place. J'ai tout d'abord testé l'affinité de la MetRS pour différents analogues afin d'évaluer l'influence de chaque groupement chimique sur l'activité. Ceci est destiné à améliorer la compréhension du mode de fonctionnement du site. Ensuite, nous avons également tiré partie de l'étude structurale de la MetRS : β -Met présentée en partie I et des différentes informations obtenues grâce aux études de mutagenèse dirigée réalisées depuis les années 90. Des mutations ont alors été effectuées dans le but d'augmenter la spécificité pour la β -méthionine. Enfin, la dernière stratégie mise en place a consisté à générer des variants *in silico*, en collaboration avec Thomas Simonson et Vaitea Opuu.

I Interaction de la MetRS avec des analogues de la méthionine

I. 1 Mesure des paramètres cinétiques de la MetRS vis-à-vis d'analogues de la méthionine

Les choix d'analogues étudiés ont été guidés par les études déjà effectuées (sélénométhionine et acide glutamique 5-méthyle ester (Fersht and Dingwall, 1979b; Kraus *et al.* , 1979; Hendrickson *et al.* , 1990; Kiick and Tirrell, 2000)) ainsi que par la disponibilité commerciale des différents composés

La misacylation d'acides aminés non canoniques par la MetRS a été étudiée dès les années 70 avec les analogues éthionine, norleucine et norvaline. Cette étude a été menée avec la MetRS de *B. stearothermophilus*. Elle a montré que les constantes cinétiques d'échange étaient sensibles aux variations dans la chaîne latérale. Cela a mis en évidence l'importance des électrons π de la chaîne latérale, de leurs positions sur celle-ci et l'influence de la longueur de la chaîne latérale ((Fersht and Dingwall, 1979b) Figure 39).

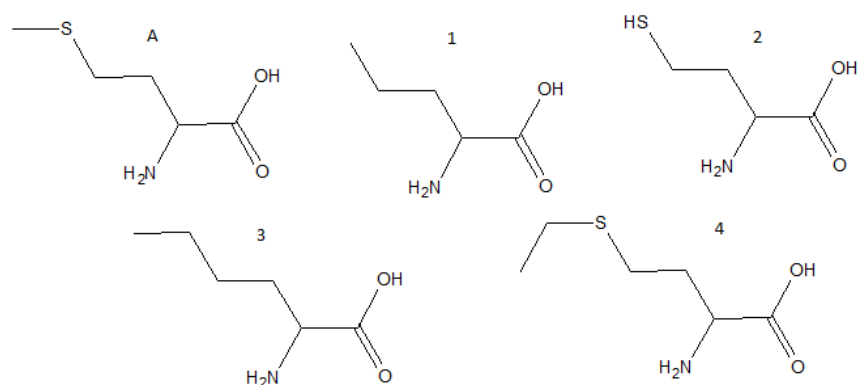


Figure 39 : Analogues de la méthionine, A : Méthionine 1 : Norvaline 2 : Homocystéine 3 : Norleucine 4 : Ethionine

Les études de l'équipe de Tirrell ont confirmé les résultats de Fersht en étudiant l'interaction de la MetRS avec une plus grande variété d'analogues présentés sur la figure 40 (van Hest *et al.* , 2000; Kiick and Tirrell, 2000; Kiick *et al.* , 2001). Ces études montrent l'importance de la longueur de la chaîne latérale et du positionnement des électrons π sur celle-ci.

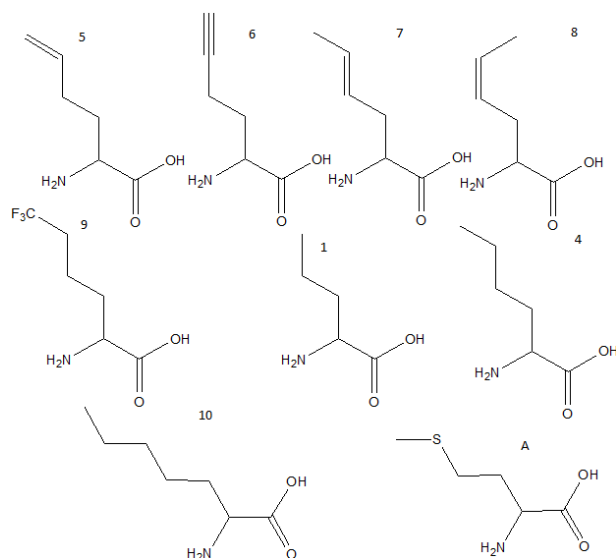


Figure 40 : Analogues de la méthionine, A : Méthionine ; 1 : Norvaline ; 4 : Norleucine ; 5 : Acide 2-amino-5-hexenoïque 6 : Acide 2-amino-5-hexynoïque ; 7 : Acide *cis*-2-amino-4-hexenoïque ; 8 : Acide *trans*-2-amino-4-hexenoïque ; 9 : Acide 6,6,6-trifluoro-2-aminoheptanoïque ; 10 : Acide 2-aminoheptanoïque

Notre choix s'est porté sur quatre composés. Trois des analogues ont une variation au niveau du groupement thio-éther et le dernier analogue a un groupement alcool à la place du groupement amine (figure 41)

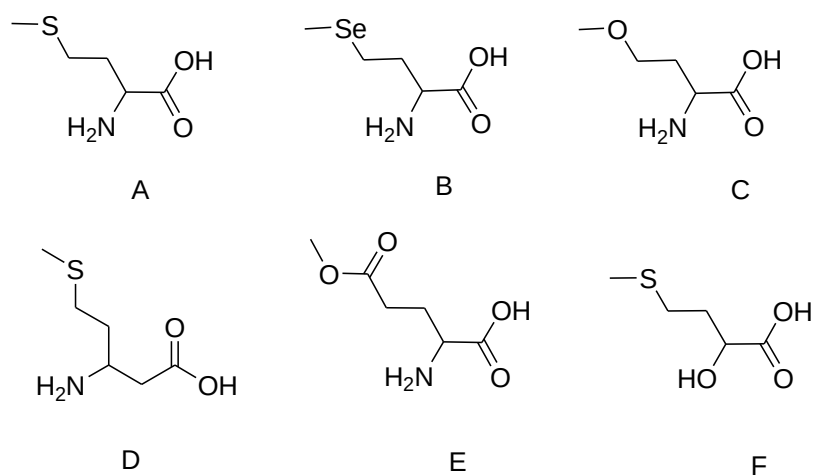


Figure 41 : Méthionine (Met, A), Selenométhionine (SeMet, B), méthoxininine (MOX, C), β-méthionine (β-Met, D), Acide glutamique 5-méthyle ester (Glu-ester, E), Desmeninol (DES, F)

Les composés ont été testés en LC-MS afin de mesurer d'éventuelles contaminations par la méthionine. Ces analyses ont montré qu'il n'y avait pas de méthionine détectable dans les solutions d'analogues (Figure 42). Ces résultats concordent avec les voies de synthèse chimique des fournisseurs qui n'utilisent pas d' α -Met parmi les réactifs.

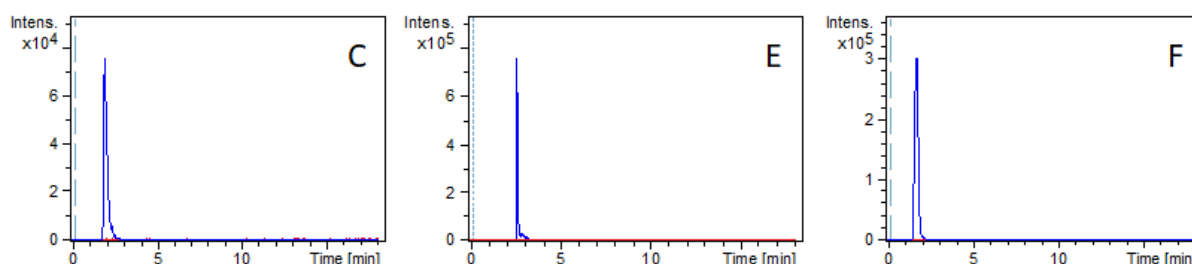


Figure 42 : Spectre LC-HRMS des composés C, E et F.

Les courants ionique extraits correspondent aux m/z de leurs ions respectifs ($[MOXH]^+ = 134.0811 \pm 0.001$; $[Glu-esterH]^+ = 162.0755 \pm 0.001$; $[DesHLi]Li^+ = 163.0557 \pm 0.001$) en bleu et celui correspondant à méthionine ($[\alpha\text{-MetH}]^+ = 150.0577 \pm 0.001$) en rouge.

Pour évaluer les affinités entre les analogues de la méthionine et la MetRS, les paramètres michaeliens de la réaction d'échange ATP-PPi ont été mesurées. L'objectif est de déterminer dans quelle mesure la structure et les groupements chimiques du substrat influencent les constantes K_m , k_{cat} ainsi que son affinité pour l'enzyme.

Les constantes K_m et k_{cat} reflètent indirectement les paramètres de la réaction d'activation de l'AA par l'ATP. En effet, à cause de la stabilité du complexe [MetRS : Met-AMP], qui a un temps de demi vie de 40 min, l'enzyme est très peu renouvelée. Par conséquent la formation de Met-AMP ne suit pas une cinétique de type Michaelis. Pour contourner cette limitation, on profite de la réversibilité de la réaction (figure 43).

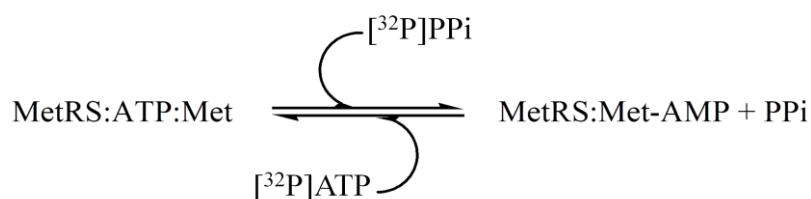


Figure 43 : Mécanisme réactionnel de la réaction d'échange isotopique [^{32}P] PPi-ATP

Si la réaction est effectuée en présence de [^{32}P]PPi, l'équilibre chimique entre les différentes espèces sera atteint rapidement après le début de la réaction. La dynamique de cet équilibre se traduit par un transfert de la radioactivité sur l'ATP jusqu'à l'équilibre isotopique. La vitesse de cette réaction d'échange isotopique entre l'ATP et le [^{32}P]PPi est étroitement liée à l'efficacité de l'activation de la méthionine (Blanquet *et al.*, 1974; Mellot *et al.*, 1989).

La constante d'affinité est mesurée grâce à la variation de fluorescence intrinsèque de la protéine lors de sa liaison avec le substrat. En effet, la MetRS possède une fluorescence intrinsèque ($\lambda_{\text{ex}}=280\text{-}295\text{nm}$; $\lambda_{\text{em}}=340\text{nm}$) liée aux résidus tryptophanes (Blanquet *et al.*, 1972; Hyafil *et al.*, 1976). La fluorescence de l'enzyme complexée à la méthionine est de 26% plus élevée que celle de l'enzyme seule. En mesurant la fluorescence à plusieurs concentrations de substrat, il est possible de déterminer la constante de dissociation.

Il faut rappeler que les liaisons de la Met et de l'ATP à la MetRS s'opèrent grâce à un couplage synergique. Ce couplage n'est pas apparent, car il est compensé par la répulsion électrostatique entre le phosphoryle α de l'ATP et le carboxylate de la méthionine. Toutefois, si cette répulsion est abolie, par exemple en supprimant le phosphoryle α de l'ATP, le couplage devient apparent. Ainsi, l'affinité de la méthionine est augmentée d'un facteur 200 lorsqu'elle est mesurée en présence de concentrations saturantes d'adénosine et de pyrophosphate (15mM Adénosine et 2mM PPi pour mesurer $K_{\text{met}}^{\text{Ado}^{\text{PP}}}$) (Blanquet *et al.*, 1975a; Blanquet *et al.*, 1975b). Ainsi, pour un analogue dont l'affinité est trop faible pour être mesurée directement, il peut devenir possible de mesurer $K_{\text{met}}^{\text{Ado}^{\text{PP}}}$. Cette mesure peut donner une idée de l'affinité de l'analogue si on fait l'hypothèse que le couplage est conservé.

Tableau 3 : Paramètres cinétiques de la réaction d'échange ATP-[³²P] PPi et constantes de dissociation

nm : non mesurable $K_D > 20\text{mM}$ ou $K_m > 20\text{mM}$, nd : non déterminé

Ligand	K_m (mM)	k_{cat} (s ⁻¹)	K_D (mM)	$K_{met}^{Ado PP}$ (μM)
Méthionine (A)	$0,02 \pm 0,002$	57 ± 10	$0,05 \pm 0,01$	$0,4 \pm 0,01$
β-Méthionine	>20	<0.002	$0,5 \pm 0,02$	120 ± 10
Acide glutamique 5-méthyle ester (E)	12 ± 2	15 ± 2	7 ± 1	ND
Méthoxinine (C)	8 ± 2	9 ± 2	>20	296 ± 30
Desmeninol (F)	>20	($v_i=0,07$ à 2mM)	>20	ND
Selenométhionine (B)	$0,02 \pm 0,002$	50 ± 10	$0,04 \pm 0,01$	ND

Les constantes obtenues avec la SeMet (B) sont les mêmes que celle obtenues avec la méthionine. Ceci est cohérent car le sélénium et le soufre sont deux éléments très proches chimiquement (environnement électronique). De la même manière, le K_D de la SeMet est similaire à celui de la méthionine, voire légèrement meilleur.

La méthoxinine (C), qui a un oxygène à la place du soufre voit son K_m augmenter d'un facteur 400 et son k_{cat} diminuer d'un facteur 5. Ceci est probablement dû à la taille de l'oxygène qui est plus petit que l'atome de soufre. Cette diminution de taille affecterait les interactions des doublets non liants avec la chaîne principale de la L13 et la Y260. La constante de dissociation de la méthoxinine est très élevée. Ceci montre que le groupement éther est trop petit et pourrait ne pas stabiliser le substrat dans la poche de reconnaissance de la méthionine.

Lorsqu'il y a une fonction ester à la place du soufre, on observe une variation du k_{cat} (diminution d'un facteur 3) et du K_m (augmentation d'un facteur 600) (E). Il est intéressant de comparer ce composé à l'éthionine car leurs chaînes latérales ont les mêmes longueurs et que le soufre se positionne au niveau de la liaison C=O de l'ester. Ce sont donc les doublets de ceux-ci qui vont prendre la place de ceux du soufre (figure 44).

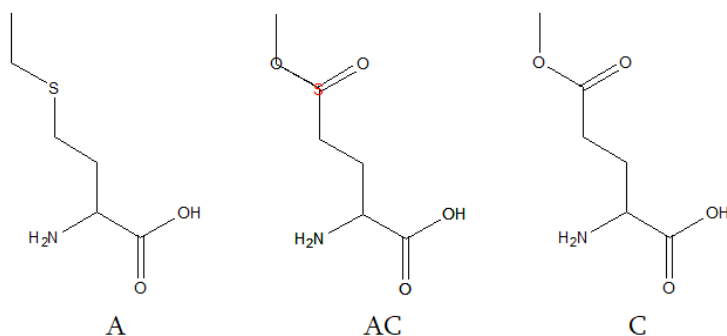


Figure 44 structure d'analogue de la méthionine A: Ethionine AC: superposition de l'Ethionine et de l'acide glutamique 5-methyl ester C: Acide glutamique 5-méthyle ester

Les valeurs des constantes michaeliennes de l'éthionine ont été mesurées pour la MetRS de *B. stearothermophilus* (Fersht and Dingwall, 1979b). Les k_{cat} sont du même ordre de grandeur et il y a un facteur 50 entre les K_m . Ces écarts ne sont pas du même ordre de grandeur que ceux obtenus entre la méthionine et l'acide glutamique 5-méthyle ester. Les variations des constantes obtenues pour l'ester ne sont donc pas uniquement dues à la longueur de la chaîne latérale mais aussi au fait que les interactions de l'atome du soufre avec l'enzyme ne sont pas les mêmes. En mesurant le K_D , on peut aussi voir une différence notable d'affinité du substrat avec une baisse d'un facteur de 136 entre le $K_{D\text{ Met-ester}}$ et le $K_{D\text{ Met}}$. Les constantes cinétiques ainsi que la constante de dissociation du desmeninol (F) n'ont pas pu être mesurées, car l'affinité pour ce composé n'est pas suffisante.

Pour expliquer les variations d'affinités de l'enzyme pour les différents substrats nous avons jugé nécessaire de résoudre les structures de l'enzyme en présence des analogues.

I. 2 Structures de la MetRS cocrystallisée avec des analogues de la méthionine

Les résultats précédents ont montré la capacité de l'enzyme à utiliser des analogues de la méthionine. Afin de mieux comprendre les bases structurales de ces propriétés, nous avons cristallisé l'enzyme en présence de ces substrats. Cela a été fait par micro ensemencement d'une solution sursaturée d'enzyme (3,6mg/ml de MetRS, 1,08M citrate d'ammonium à 6°C) contenant de 10 à 50 mM de substrat. Malgré les K_D élevés obtenus pour la méthoxinine et le desmeninol, ces composés ont pu être détectés dans la densité électronique. Dans tous les cas, des données à haute résolution ont été collectées. Ceci a permis d'affiner les valeurs d'occupation des substrats.

Les caractéristiques de ces cristaux et des données d'affinement des structures sont regroupées dans le tableau suivant :

	MetRS:SeMet	MetRS:Met-ether	MetRS:Met-ester	MetRS:Des
Données de diffraction				
Groupe d'espace	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Dimension de la maille				
a, b, c (Å)	78,39; 45,19;86,07	78,26 ; 45,77 ; 86,34	78,75 ; 45,46 ; 86,53	78,31; 45,11; 86,21
α, β, γ (°)	90;107,32;90	90;107,45;90	90;107,45;90	90;107,39;90
Résolution (Å)	1,35	1,39	1,31	1,45
R _{sym} (%) ^a	8,1 (121,1) ^b	8,4 (175)	7,4 (121,8)	11,9 (129)
I/σ(I)	7,44 (10,90)	11,4 (1,2)	10,4 (1,1)	10,2 (1,3)
Complétude (%)	96,7 (89,4)	98,7 (92,6)	98,1 (94,4)	98,9 (93,2)
Redondance	2,2 (2,6)	6,4 (5,8)	3,3 (3,2)	6,8 (6,5)
CC _{1/2} (%) ^c	99,7 (29,4)	99,8 (35,1)	99,8 (99,4)	99,7 (54,5)
Affinement				
Résolution (Å)	38,7-1,34	30,6-1,39	30,9-1,30	33,3-1,44
Nb réflexions	127763	115194	139624	104444
R _{work} /R _{free}	0,139/0,176 (0,29/0,32)	0,150/0,184 (0,275/0,322)	0,153/0,172 (0,348/0,377)	0,147/0,187 (0,303/0,351)
Nb atomes				
Protéine	4577	4649	4670	4628
Eau	712	570	652	674
Ligands	14	14	39	55
B (Å ²)				
Protéine	18,2	23,54	19,6	19,7
Eau	37	38,9	38	39,1
Ligands	33,9	34,9	41,32	47,6
r. m. s. deviations				
Longueurs (Å)	0,005	0,005	0,005	0,005
Angles (°)	0,767	0,432	0,779	0,742

Tableau 4 : Données collectées et statistiques d'affinement

- $R_{sym}(I) = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_i I_{hkl,i}}$, ou i est le nombre de réflexion d'indice hkl
- Les valeurs entre parenthèses correspondent à la tranche de résolution la plus élevée.
- CC(1/2) est le coefficient de corrélation entre deux moitiés du jeu de données choisies au hasard (Karplus and Diederichs, 2012).
- $R_{work} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$; R_{free} est calculé à partir de 5% des réflexions

I. 2. 1 Structure du complexe MetRS:SeMet

La SeMet se positionne exactement comme la méthionine dans le site actif, ce qui est cohérent avec les constantes cinétiques mesurées. La résolution de cette structure avait un deuxième intérêt. Comme cela a été vu dans la partie I, la β-Met possède un second site de liaison. Nous voulions savoir si la méthionine (ou la SeMet) pouvait se fixer à ce second site, ou bien s'il était spécifique de la β-Met. Afin de maximiser les chances de détecter une éventuelle occupation du second site, nous avons mis à profit la diffusion anormale du sélénium au seuil d'absorption, et utilisé une forte concentration du composé (20mM) lors de la cocrystallisation (Figure 45).

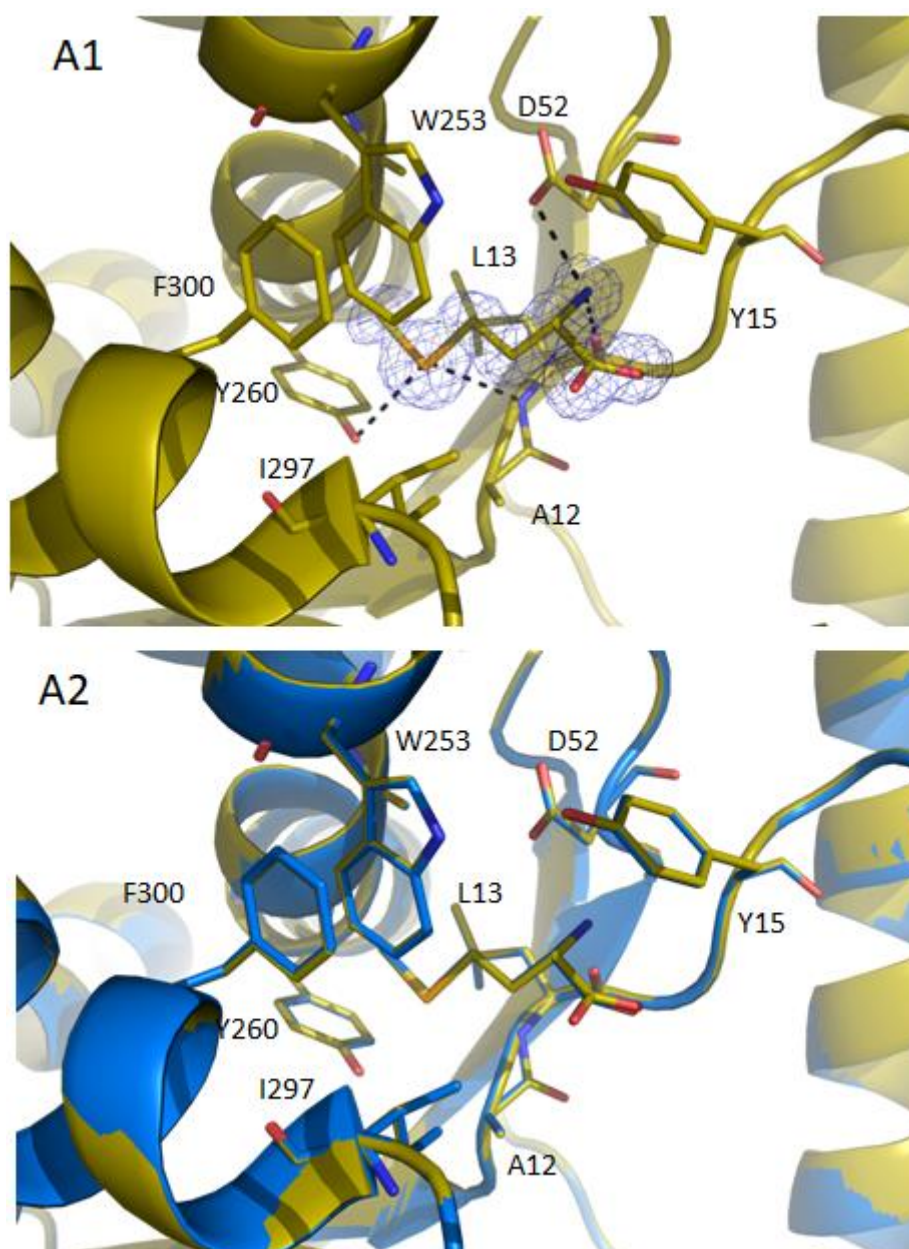


Figure 45 : Liaison de la sélénométhionine à la MetRS d'*E. coli*.

A : Structure du complexe MetRS:SeMet. B : Structure du complexe MetRS:SeMet (jaune) superposée à celle du complexe MetRS:Met (bleu). Les résidus entourant le site sont représentés en « bâtons » alors que la chaîne principale de la MetRS est représentée en « cartoon ». Sur la structure en A la densité électronique (carte 2mFo-DFc contournée à 1.0 sigma) est représentée en bleu autour du substrat en utilisant la commande carve de PyMOL. Les principales interactions sont représentées en traits pointillés. Les figures 46 à 66 sont réalisées de la même façon.

Un fort pic de densité de la carte de différence ($F_o - F_c$) a été observé dans la poche de reconnaissance du substrat mais aucun pic n'a été observé dans la seconde poche. Ainsi, cette poche est bien spécifique à la β -Met.

I. 2. 2 Structure du complexe MetRS:Méthoxinine

La MetRS a été cristallisée en présence de 20mM de méthoxinine (MOX). L'occupation du substrat a été affinée à la valeur de 75%.

La Méthoxinine a un comportement assez proche de celui la méthionine. L'oxygène de la fonction éther est toutefois déplacé de 0,2 Å vers l'amine de la L13, si on compare au soufre de la méthionine. Ceci engendre aussi un léger basculement des positions de l'amine et du carboxylate visibles sur la figure 46.

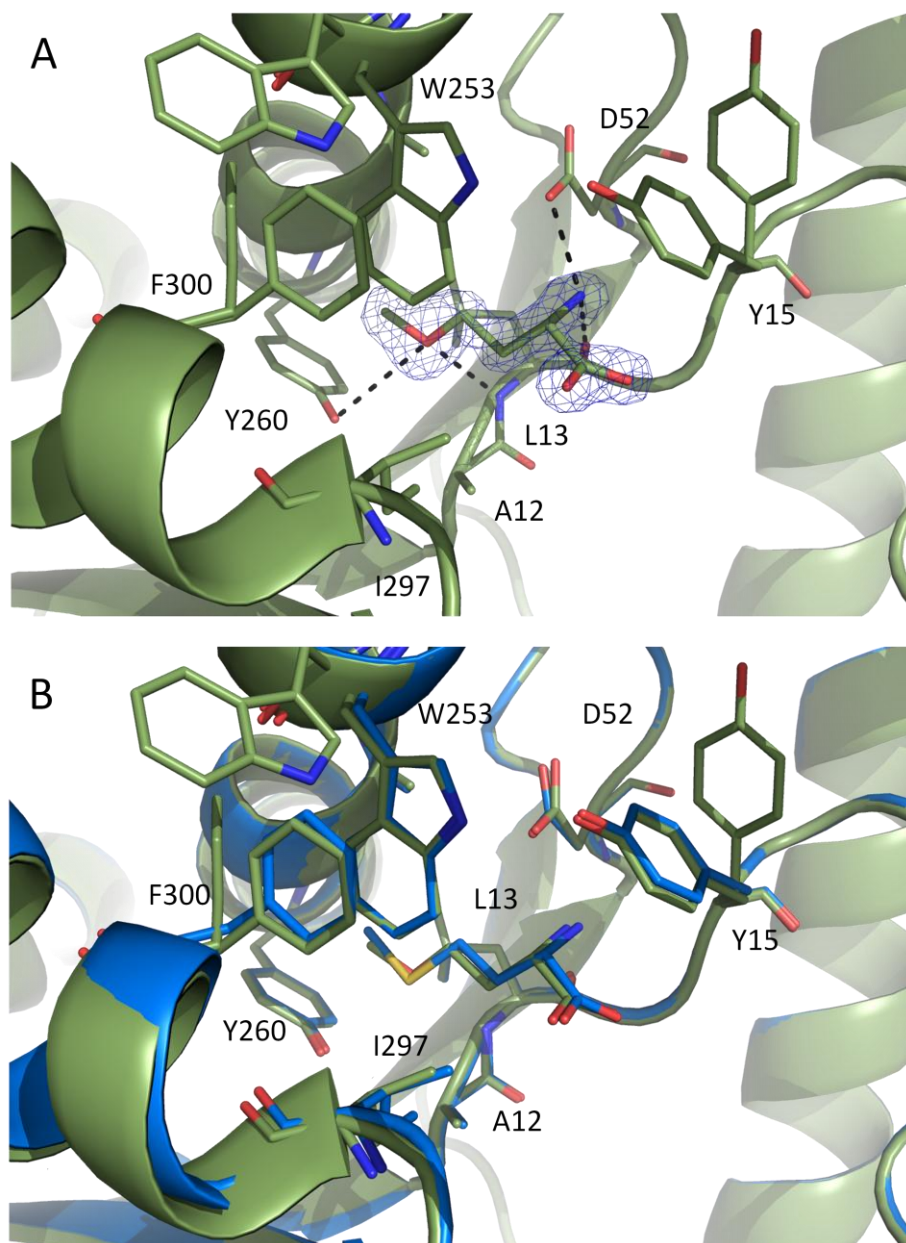


Figure 46 : Liaison de la méthoxinine à la MetRS d'*E. coli*.

A : Structure du complexe MetRS:MOX. B : Structure du complexe MetRS:MOX (vert) superposée à celle du complexe MetRS:Met (bleu).

I. 2. 3 Structure du complexe MetRS:Acide glutamique 5-méthyle ester

L'acide glutamique 5-méthyle ester (Glu-ester) a été cocrystallisé avec la MetRS à plusieurs concentrations de substrat. En fonction de celles-ci l'occupation varie. A 10mM l'occupation est affinée à 30% alors qu'à 20mM elle est affinée à 70%.

L'acide glutamique 5-méthyle ester se positionne de façon proche de la méthionine. Les doublets non liants C=O prennent la position des doublets du soufre. L'analogue Met-ester ayant une chaîne latérale plus longue d'un atome, cela induit un décalage sur la position de l'amine et du carboxylate de 0,4 Å et un léger décalage du groupement méthyle. (Figure 47). L'effet du décalage est principalement visible au niveau de la Y15 qui voit son groupement alcool décalé de 1 Å.

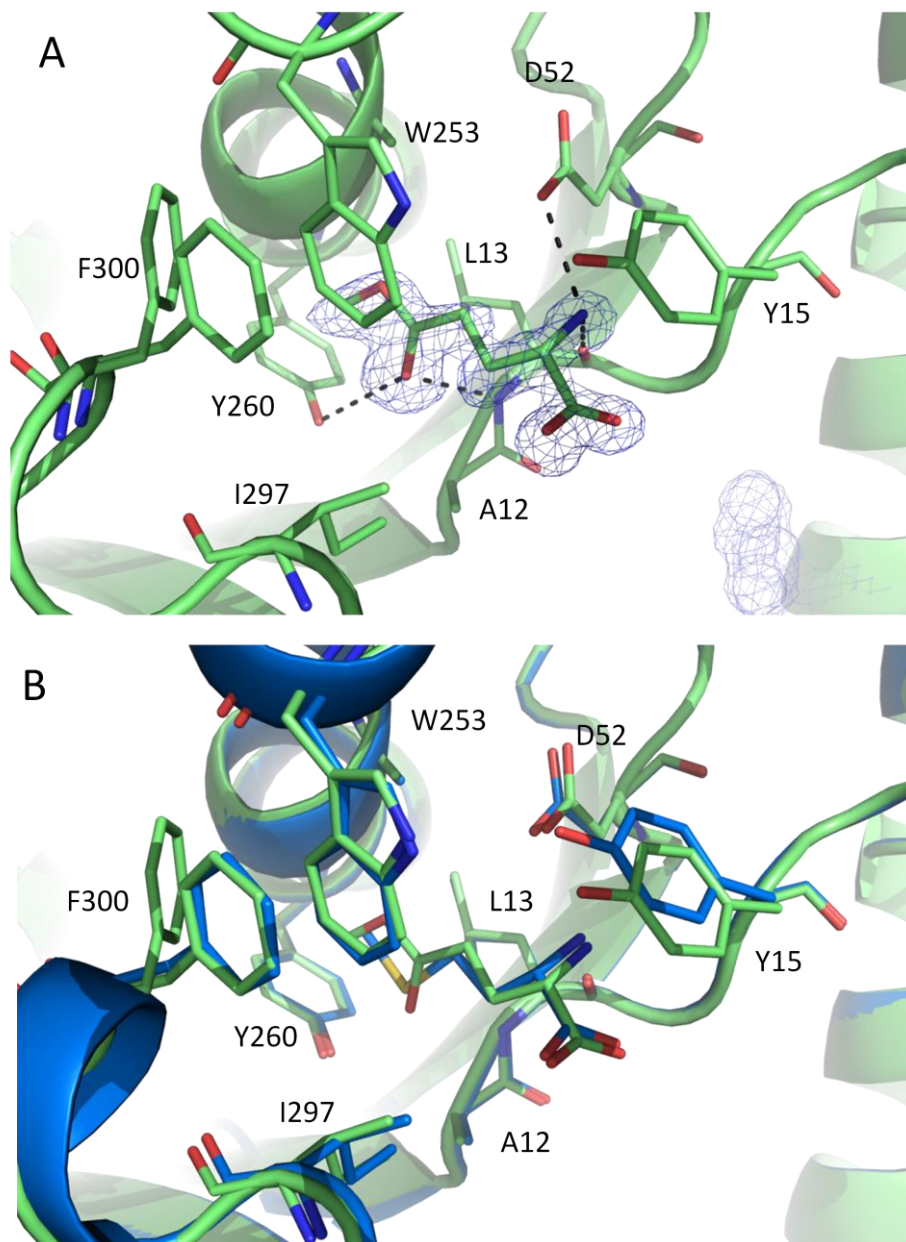


Figure 47 : Liaison de l'acide glutamique 5-méthyle ester à la MetRS d'*E. coli*.

A : Structure du complexe MetRS:Glu-ester. B : Structure du complexe MetRS:Glu-ester (vert) superposé à celle du complexe MetRS:Met (bleu).

Il y a une bonne adéquation entre les résultats structuraux et les résultats cinétiques obtenus avec la β -Met et l'acide glutamique 5-méthyle ester (Figure 48). En effet, le décalage

du substrat dans la poche provoque un décalage du carboxylate. L'ensemble des observations défavorables à la catalyse efficace explique la baisse du k_{cat} , l'augmentation du K_m ainsi que celle du K_D .

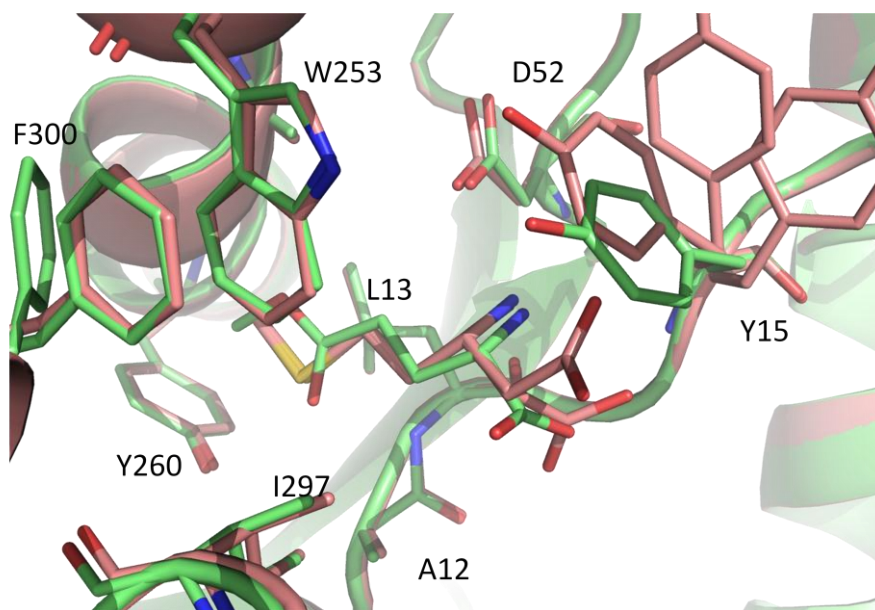


Figure 48 : Structure du complexe MetRS:β-Met (rose) superposée à la structure du complexe MetRS :Glu-ester (bleu)

On peut voir sur la figure précédente, que le décalage du groupement carboxyle de l'acide glutamique 5-méthyle ester n'est pas aussi important que observé avec la β méthionine. Ceci explique pourquoi la β-Met est un substrat encore moins bon que l'acide glutamique 5-méthyle ester (Figure 48).

I. 2. 4 Structure du complexe MetRS:Desmeninol

La structure de la MetRS cocrystallisée avec le desmeninol (DES) montre quelques différences avec la structure du complexe MetRS : Met. Dans ce cas, l'amine est remplacée par un alcool. L'occupation du site est affinée à 83% pour les cristaux obtenus en présence de 43mM de substrat. On observe que le résidu Y15 ne bascule plus (figure 49). Cela peut s'expliquer par la différence de liaison entre les groupements -NH₂ et -OH avec le D52 et la L13. En effet, le groupement -NH₂ forme une liaison probablement légèrement plus longue due à sa charge. Ce changement d'orientation fait que le groupement -OH ne peut former de liaison π avec la Y15. Ceci peut expliquer le décalage par rapport au positionnement de la méthionine visible notamment au niveau du groupement carboxylate et du groupement alcool.

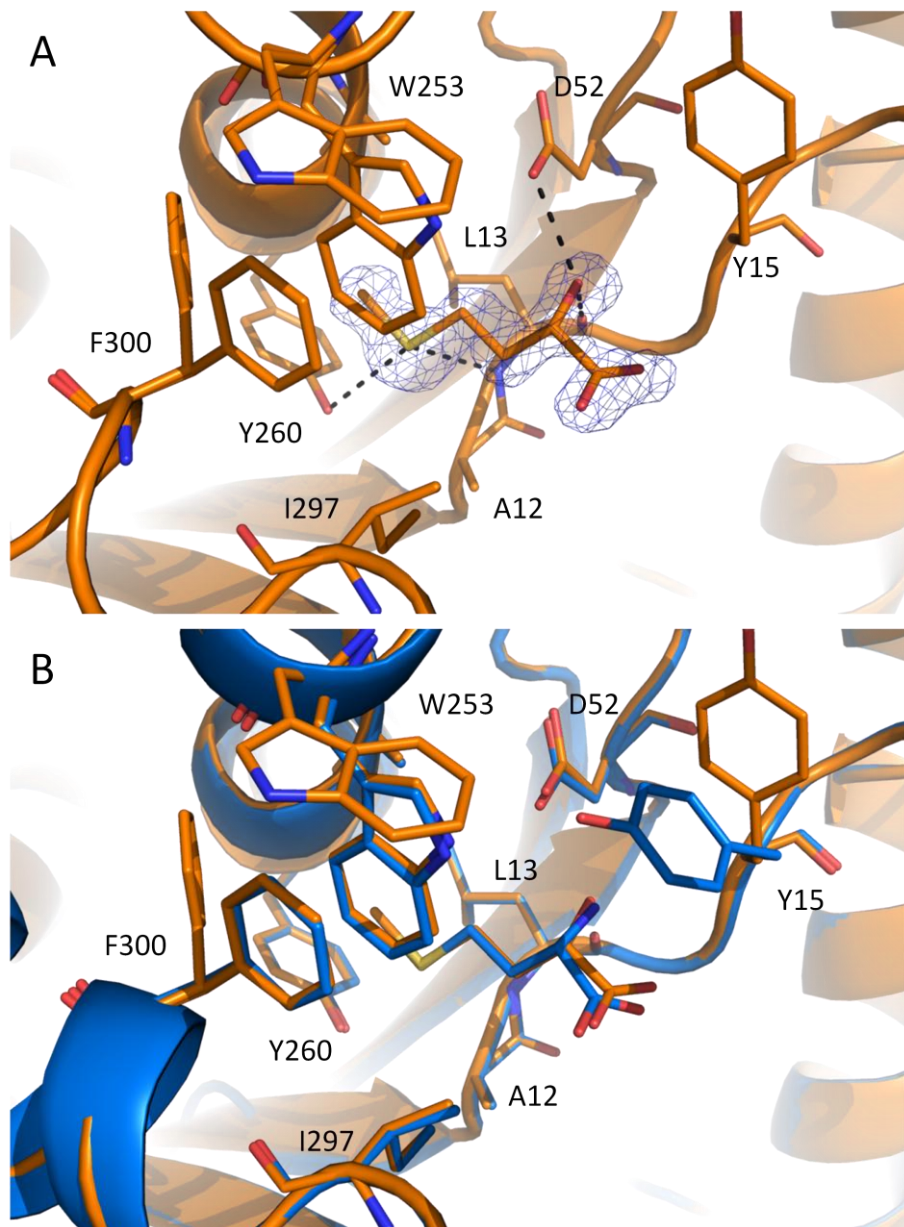


Figure 49 : Liaison du Desmeninol à la MetRS d'*E. coli*.

A : Structure du complexe MetRS:DES. B : Structure du complexe MetRS:DES (orange) superposé au complexe MetRS:Met (bleu).

L'observation de la Y15 qui ne bascule pas et des constantes michaeliennes qui ne sont pas mesurables montrent que l'interaction entre le substrat et la Y15 ainsi que le basculement de celle-ci sont importants pour la catalyse. Ceci se confirme en comparant ces résultats à ceux de D. Fourmy (Fourmy, 1993). Le variant YF15 montre un k_{cat} diminué d'un facteur 10 et un K_m augmenté d'un facteur 5, alors que le variant Y15A à une vitesse initiale de $0,62 \text{ s}^{-1}$ (à 2 mM Met), soit une efficacité catalytique diminuée de trois ordres de grandeur

I. 3 Conclusion

Bien que le rôle global des résidus qui entourent la poche de reconnaissance de la méthionine soit connu, il est intéressant de connaître l'importance des groupements chimiques du substrat. Il avait été montré que la position du soufre était importante (Fersht and Dingwall, 1979b; van Hest *et al.* , 2000; Kiick *et al.* , 2001). L'étude de l'interaction entre les analogues de la méthionine avec une différence au niveau du soufre, nous informe que la taille de l'atome et le positionnement des doublets non-liants de celui-ci sont importants. De plus, nous avons pu observer que la modification de l'amine en alcool provoquait une chute drastique de l'affinité du substrat pour l'enzyme, probablement due au positionnement des doublets non-liants de l'alcool qui remplace l'amine.

Ainsi, certaines modifications de la chaîne latérale sont assez bien tolérées alors que, comme on l'a vu dans la partie I, le positionnement du carboxylate joue un rôle important.

II Mutagenèse ponctuelle en vue d'améliorer l'activité de la MetRS pour la β -Met

Après avoir défini le rôle et l'importance de chaque groupement chimique, grâce à l'étude cinétique et structurale de la MetRS avec différents analogues de la méthionine, nous avons comparé ces résultats avec les données structurales du complexe MetRS : β -Met. Ce recoupement d'informations nous a permis, de définir une série d'acides aminés à modifier et de mesurer les activités de ces variants vis-à-vis de la Met et de la β -Met.

II. 1 Sélection des positions

Comme montré dans la partie I, une fois la β -Met positionnée dans la poche de reconnaissance de la méthionine, celle-ci peut prendre deux conformations distinctes. L'une d'entre elles semble plus favorable, car elle permet d'une part à la Y15 de basculer et d'autre part au groupement carboxyle de la β -Met de se trouver dans une position plus favorable pour réagir avec le phosphate α de l'ATP (figure 50).

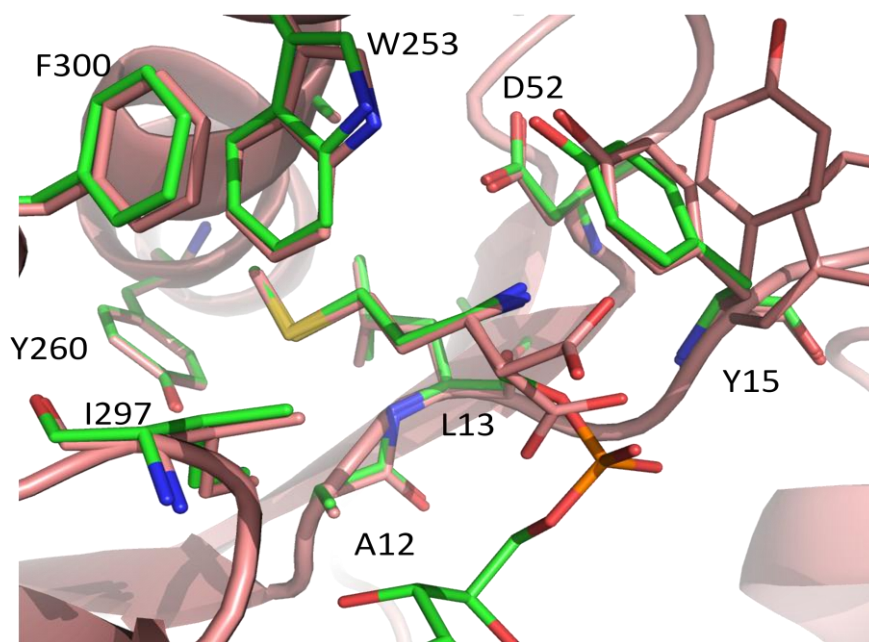


Figure 50 : Structure du complexe MetRS: β -Met (rose) superposée à la structure de complexe MetRS : méthioninyl-adénylate (bleu).

Pour favoriser la seconde conformation de la β -Met, il faut forcer le basculement du groupement carboxyle vers sa position catalytique.

Dans cette optique, nous avons décidé de modifier le résidu A12 par un acide aminé susceptible de créer une liaison hydrogène avec le carboxyle de la β -met permettant de stabiliser le groupement. Quatre modifications ont été réalisées A12S, A12T, A12N et A12D. L'autre option envisagée a été de modifier le D52 en Acide Glutamique ce qui pourrait permettre au substrat de tourner dans la poche et ainsi de se décaler pour que la Y15 ait une plus grande liberté de se positionner de façon adéquate.

II. 2 Etude cinétique des variants de la MetRS

Après avoir produit ces variants, leurs vitesses d'activation ont été mesurées en utilisant comme substrat la méthionine ou la β -méthionine (Tableau 5).

Les vitesses des variants ont été mesurées dans les mêmes conditions que lors de l'étude de l'enzyme sauvage avec les différents substrats. Il est important de rappeler que la β -Met est contaminée par de la méthionine à hauteur de 0.09%. Nous avons vu dans le chapitre 1 que la contamination pouvait être fortement diminuée par traitement à la γ -lyase. Toutefois, les préparations d'enzymes sont également contaminées en Met. Nous ne sommes pas parvenus à trouver de traitement permettant une mesure fiable de la vitesse d'échange isotopique dépendante de la β -Met. Les vitesses mesurées seront donc en réalité des vitesses initiales d'activation de la méthionine inhibées. En faisant le ratio V_{β}/V_{Met} , on pourra mesurer l'influence de la β -Met sur la vitesse initiale d'activation de la Met et donc la capacité de la β -Met à agir comme un inhibiteur compétitif (Tableau 5).

Tableau 5 : Vitesses initiales d'activation de la Méthionine et de la β -méthionine par différents variants (2mM Met ou 2mM β -Met, 2mM ATP et 2mM PPI)

Mutation	V_i (s^{-1})	$V_{i\beta}$ (s^{-1})	ratio V_{β}/V_{Met}
WT	50 ± 10	$1,2 \pm 0,2$	0,024
D52E	$0,29 \pm 0,6$	$0,004 \pm 0,001$	0,012
A50V	$0,41 \pm 0,06$	$0,006 \pm 0,002$	0,015
A12S	$0,97 \pm 0,2$	$0,024 \pm 0,008$	0,025
A12T	$0,025 \pm 0,005$	$< 0,002$	ND
A12D	$0,15 \pm 0,04$	$< 0,002$	ND
A12D + D52E	$< 0,006$	$< 0,002$	ND
A12N	$< 0,006$	$< 0,002$	ND

La plupart des mutations a provoqué une perte d'activité majeure. Toutefois la mutation A12S semble avoir un effet moins délétère que les autres. De plus, le ratio V_{β}/V_{Met} pour le variant A12S est proche de celui de l'enzyme WT.

Nous avons donc mesuré les constantes de Michaelis de ce variant (Tableau 6).

Tableau 6 : Constantes d'aminocyclisation de la MetRS WT et de son variant A12S pour la Met et la β -Met (contaminée)

	$K_{cat\ Met}$ (s ⁻¹)	$K_m\ Met$ (μ M)	$k_{cat\ \beta-met\ app}$ (s ⁻¹)	$K_m\ \beta-Met\ app$ (mM)	K_i (μ M)	$\frac{k_{cat}}{K_m}$	$\frac{K_i}{K_m}$	$K_D\ Met$ (μ M)	$K_D\ \beta-Met$ (mM)
WT	57 $\pm$ 0,9	20 $\pm$ 0,2	1,2 $\pm$ 2	0,57 $\pm$ 0,1	370 $\pm$ 14	3100	20	50 $\pm$ 5	1.2 $\pm$ 0,2
A12S	1,5 $\pm$ 0,03	50 $\pm$ 10	0,020 $\pm$ 0,002	2,6 $\pm$ 0,5	3300 $\pm$ 200	28	62	70 $\pm$ 7	1,4 $\pm$ 0,2

On voit clairement une baisse du k_{cat} et une augmentation du K_m (Met ou β -Met) entre le A12S et l'enzyme sauvage (Tableau 6). Le modèle d'inhibition décrit dans la partie I nous permet de déterminer la constante d'inhibition K_i de la β -Met. Celle-ci est plus élevée pour le variant portant la mutation A12S que celle de l'enzyme sauvage. L'affinité de la β -Met est donc spécifiquement diminuée par rapport à celle de la Met. Il aurait été très informatif de mesurer la vitesse de formation de l'adénylate (k_f) de la β -Met avec le variant A12S mais cela n'a pas été possible car nous n'avons pas observé de variation de fluorescence après ajout d'ATP au complexe A12S : β -Met. Nous ne pouvons donc pas conclure quant à une éventuelle amélioration de la catalyse vis-à-vis de la β -Met.

Afin de vérifier les conséquences de la mutation A12S au niveau structural, ce variant a été cristallisé seul, en présence de Met et en présence de β -Met.

II. 3 Etude structurale des variants A12S, A12T, A12N et D52E

Les cristaux ont été obtenus dans les mêmes conditions que celles décrites dans la partie précédente avec les analogues. Les caractéristiques des structures et des cristaux obtenus sont indiquées dans le tableau suivant.

	A12S:β ³ -Met	A12S:Met	A12S	A12N	A12T	D52E
Données de diffraction						
Groupe d'espace	P2 ₁	P2 ₁	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Dimension de la maille						
a, b, c (Å)	78,16 ; 45,14 ; 86,21	78,69 ; 45,34 ; 85,41	79,11 ; 45,65 ; 86,20	78,84 ; 45,77 ; 86,58	78,73 ; 45,75 ; 86,15	78,38 ; 45,78 ; 86,94
α, β, γ (°)	90 ; 107,35 ; 90	90 ; 107,36 ; 90	90 ; 107,45 ; 90	90 ; 107,33 ; 90	90 ; 107,48 ; 90	90 ; 109,03 ; 90
Résolution (Å)	1,28	1,39	1,64	1,55	1,80	1,25
R_{sym} (%)^a	3,8 (135) ^b	9,1 (120)	8,8 (90,0)	13,3 (118,6)	14,2 (123,9)	6,5 (83,9)
I/σ(I)	14 (1,2)	11,4 (1,2)	13,7(1,6)	9,8 (2,0)	9,55 (1,3)	14,0 (2,0)
Complétude (%)	98,5 (93,3)	98,9 (94,5)	99,4 (96,9)	96,8 (87,6)	98,5 (93,9)	99,5 (97,4)
Redondance	6,4 (6,1)	4,9 (4,6)	6,7 (6,3)	6,4 (6,1)	6,4 (6,0)	6,4 (5,5)
CC_{1/2} (%)^c	99,9 (60,7)	99,8 (60,3)	99,8 (74,6)	99,5 (56,1)	99,6 (66,8)	99,9 (72)
Affinement						
Résolution (Å)	30,7-1,27	48,8-1,39	32,00-1,59	33,41-1,53	33,37-1,79	47,8-1,25
Nb réflexions	147999	115456	79044	84809	55103	161686
R_{work}/R_{free}	0,132/0,164 (0,283/0,311)	0,145/0,176 (0,295/0,296)	0,176/0,202 (0,340/0,403)	0,176/0,196 (0,521/0,492)	0,174/0,24 (0,314/0,344)	0,138/0,167 (0,330/0,348)
Nb atomes						
Protéine	468	4664	4392	4495	4392	4617
Eau	770	641	701	655	486	796
Ligands	58	75	27	26	20	25
B (Å²)						
Protéine	18,3	22,1	23,4	16,6	26,6	17,1
Eau	38,1	36,6	34,3	29,8	35,5	34,4
Ligands	31,4	55,7	42,7	28,7	44,4	34,1
r.m.s. deviations						
Longueurs (Å)	0,005	0,005	0,005	0,006	0,006	0,014
angles (°)	0,795	0,801	1,144	0,947	0,808	1,349

Tableau 7 : Données collectées et statistiques d'affinement

- $R_{sym}(I) = \frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_i I_{hkl,i}}$, ou i est le nombre de réflexion d'indice hkl
- Les valeurs entre parenthèse correspondent à la tranche la plus haute.
- CC(1/2) est le coefficient de corrélation entre deux moitiés du jeu de données choisies au hasard (Karplus and Diederichs, 2012).
- $R_{work} = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|}$; Rfree est calculé à partir de 5% des réflexions

II. 3 Structures des variants A12T, A12N et D52E

Sur la figure suivante, on peut voir que T12 ou N12 interagissent avec I297 induisant un changement de conformation (Figure 51).

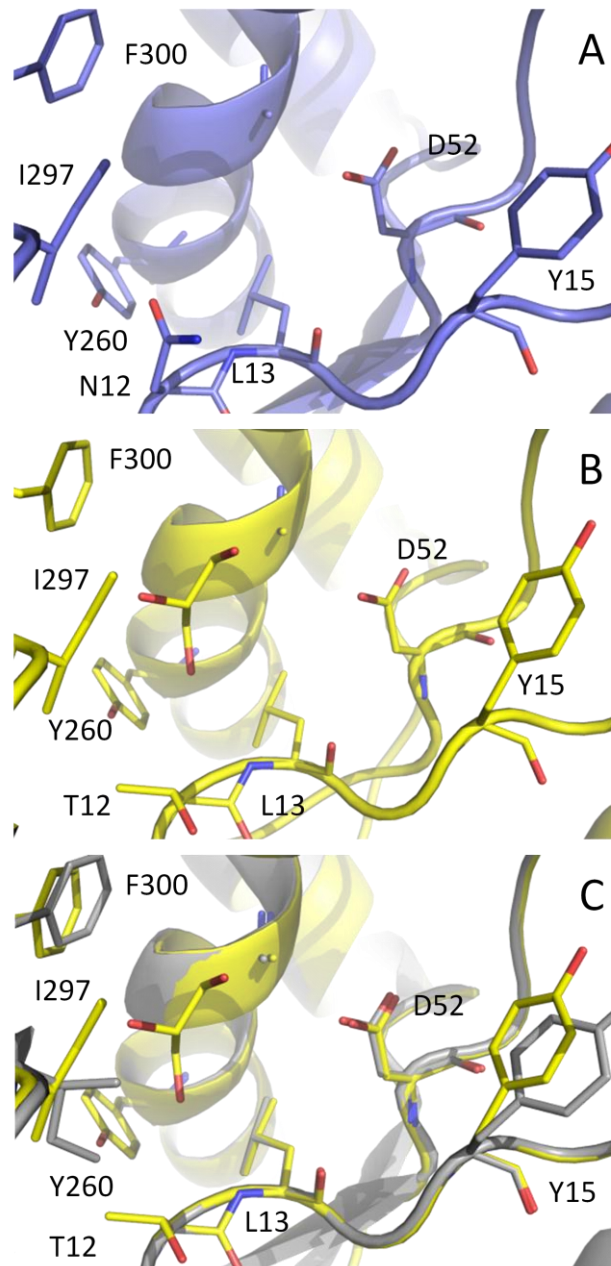


Figure 51 : Structure de la poche de reconnaissance de la méthionine de variants de la MetRS d'*E. coli*

A : Variant A12N B : Variant A12T C : Variant A12T (jaune) superposé à WT apo (gris).

I297 est impliquée dans la formation de la poche de reconnaissance de la Met. L'hélice 296-302 se décale de presque 1 Å lorsque la Met est liée. Ce décalage permet à la chaîne latérale de l'I297 de stabiliser la méthionine. De manière cohérente, le variant I297A montre une grande baisse de l'affinité pour le substrat (Fourmy, 1993).

Ainsi le changement de conformation de l'I297 dans les variants A12N et A12T ne lui permet probablement pas de stabiliser la méthionine. De plus, la position d'I297 dans les variants est susceptible de gêner le basculement de F300 nécessaire à l'ajustement induit (Figure 52).

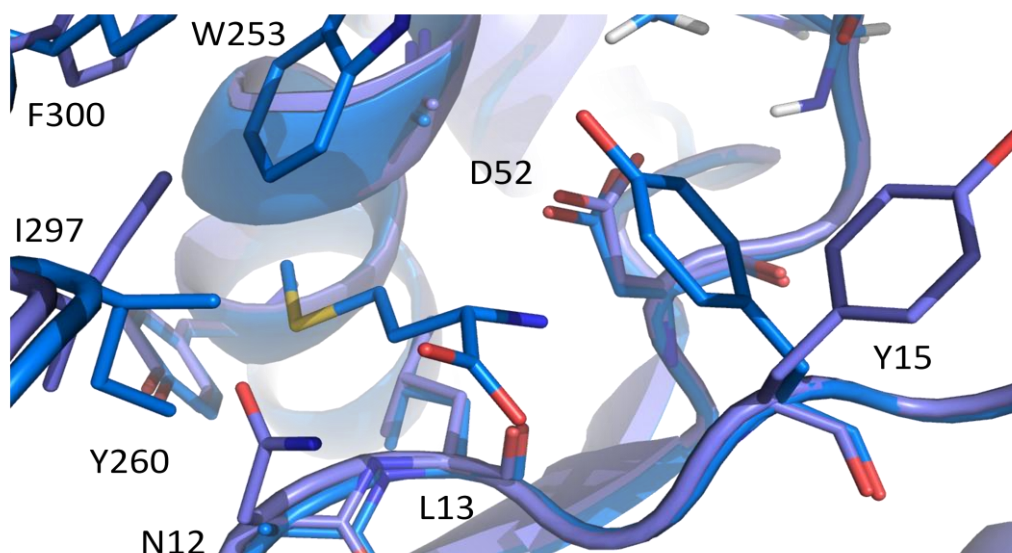


Figure 52 : Structure du variant A12N de la MetRS d'*E. coli* (violet) superposée à celle du complexe WT:Met (bleu).

Sur la structure du variant D52E (figure 53), il est possible de voir que le groupement carboxylate est légèrement décalé et a deux conformations possibles. Une seule de ces conformations est relativement favorable à la reconnaissance du substrat. De plus, l'écart entre le carboxylate du E52 et l'amine de la méthionine est trop grand pour une interaction efficace.

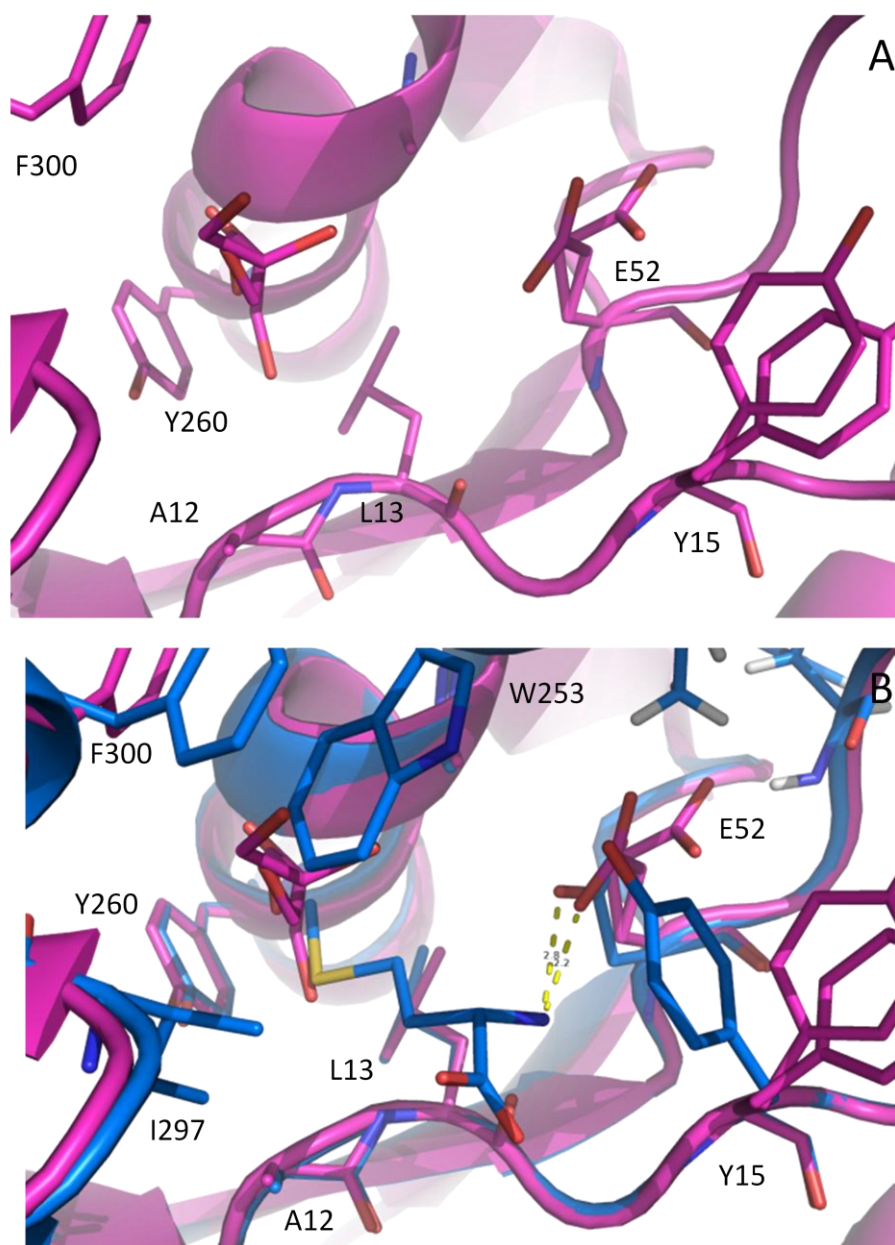


Figure 53 : Influence de la mutation D52E sur la structure de la MetRS d'*E. coli*

A : Structure du variant D52E de la MetRS de *E. coli* B : Structure du variant D52E (rose) de la MetRS de *E. coli* superposé au complexe MetRS:Met (bleu). Les distances entre les résidus en 52 et la méthionine sont représentées en traits pointillés (2,8 Å pour l'Asp et 2.2 Å pour le Glu).

II. 3. 2 Structure du variant A12S en présence de méthionine et de β -méthionine

Pour approfondir la compréhension de l'interaction entre le variant et le substrat acide aminé, nous avons souhaité déterminer les structures des complexes A12S :Met et A12S : β -Met (Figure 54).

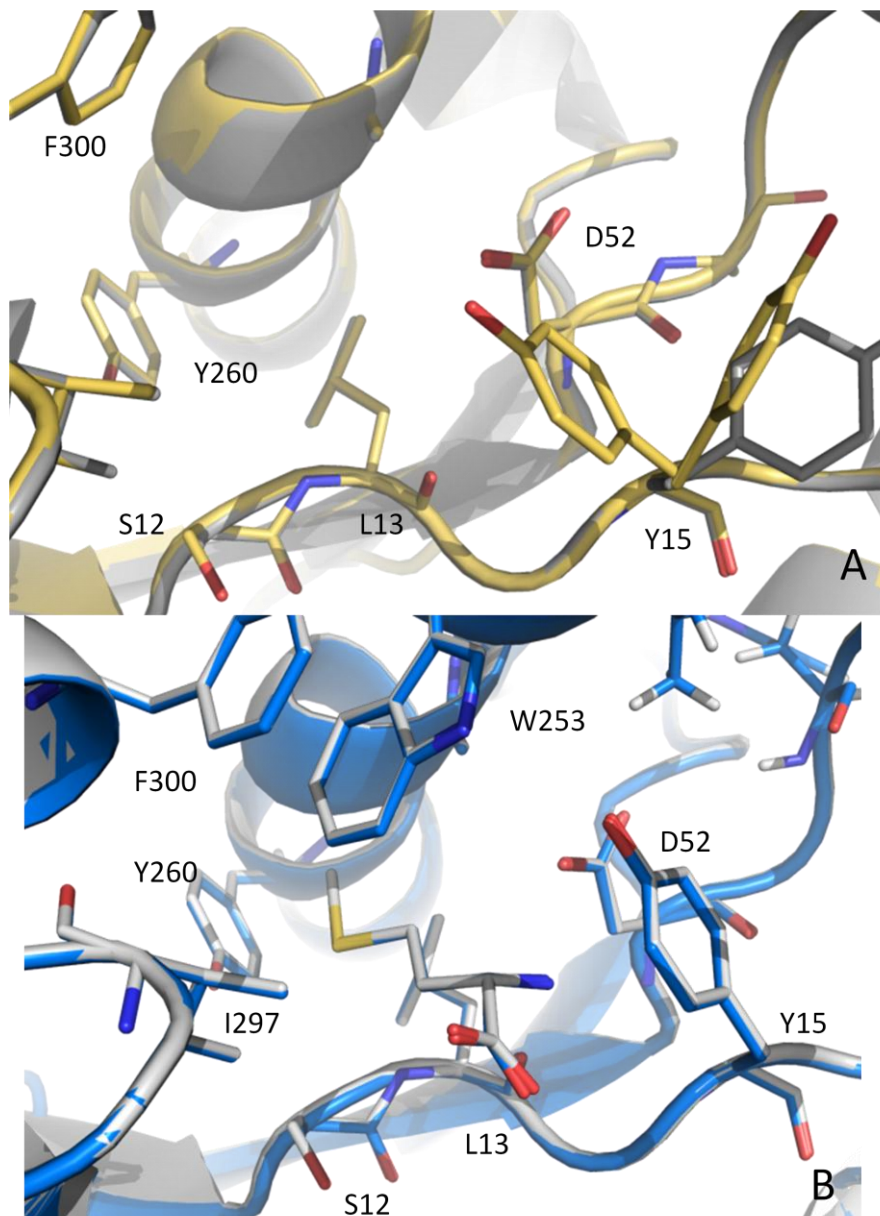


Figure 54 : Influence de la Mutation A12S sur la structure de la MetRS apo et celle du complexe MetRS :Met

A : Structure de la MetRS apo (gris) superposée du variant A12S apo (jaune). B : Structure du complexe MetRS:Met (bleu) superposée à celle du complexe A12S:Met (gris).

Les structures obtenues sont très semblables à celles obtenues avec l'enzyme WT (figure 54). Toutefois, l'affinement de l'occupation montre que le carboxylate de la β -Met est majoritairement dans la position supposée favorable à la réaction dans le variant A12S (Figure 55).

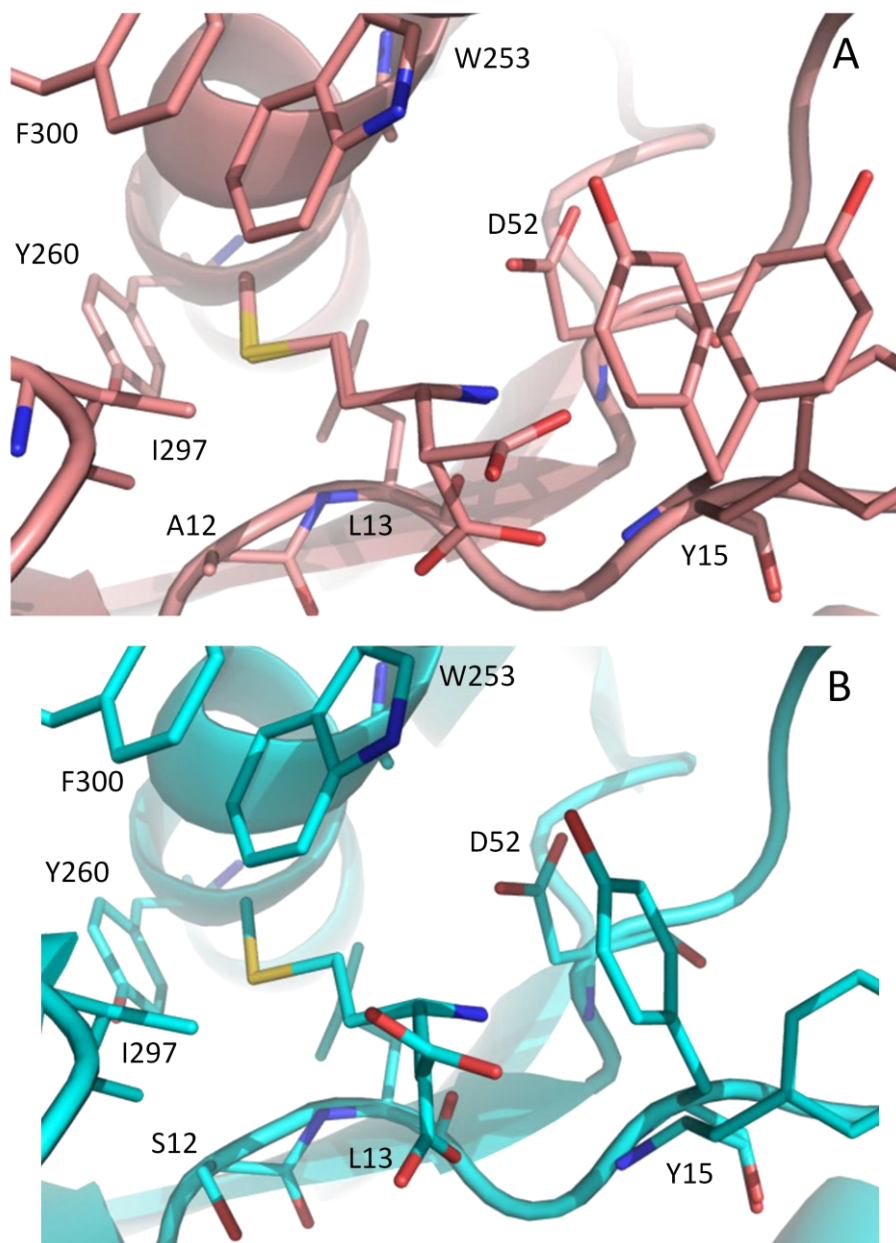


Figure 55 : Influence de la Mutation A12S sur la structure du complexe MetRS : β -Met

A : Structure du complexe MetRS: β -Met. B : Structure du complexe A12S: β -Met.

En superposant la structure du variant A12S avec celle d'un complexe de la MetRS avec un analogue du méthionyl adénylate (le méthionyl sulfamoyl adénosine, 1PFY) on peut constater une gêne stérique entre la sérine 12 et l'ATP. Ceci pourrait expliquer que la catalyse avec ce variant reste très modeste (figure 56) même si le carboxylate est mieux placé.

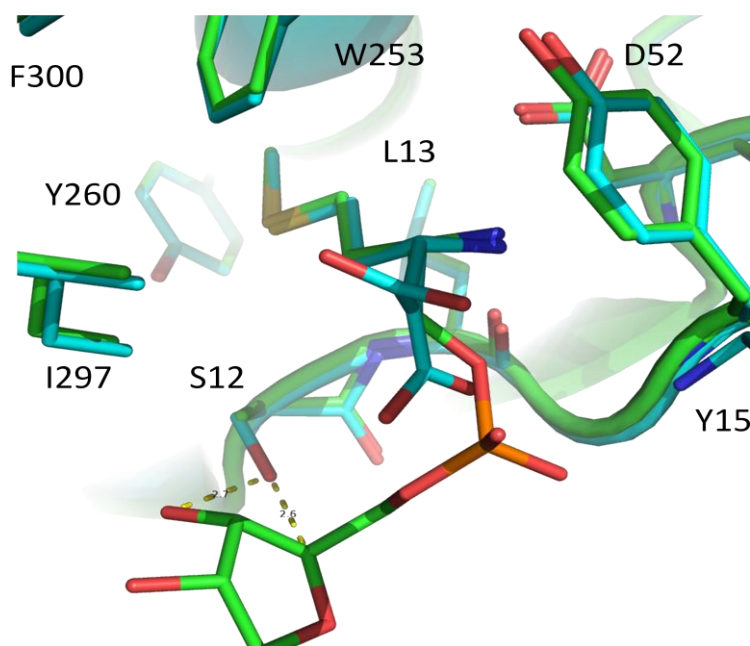


Figure 56 : Structure complexe MetRS:β-Met superposée à celle du complexe méthioninyl-adénylate.

La distance entre le méthioninyl adénylate et la serine est représentée en tiret jaune (2,7 Å entre la Ser et le OH et 2,6 Å entre la Ser et le CH)

II. 4 Conclusion

Malgré la baisse d'activité observée avec le variant A12S et le fait que ce variant ne montre pas de spécificité améliorée pour la β-Met, on observe sur la structure un positionnement plus favorable de la β-Met. La stratégie basée sur une définition intuitive des mutations à réaliser n'ayant pas donné les résultats escomptés, nous nous sommes tournés vers une approche par modélisation, basée sur des simulations *in silico*.

III Méthode de sélection *in Silico* en collaboration avec l'équipe de Thomas Simonson

III. 1 Sélection des variants

III. 1. 1 Principe

Les simulations numériques ont été effectuées par Vaitea Opuu et Thomas Simonsons. Dans un premier temps, un modèle apo, a été créé à partir de la structure du complexe avec un analogue du méthionyl-adénylate, le méthioninyl-adénylate (1PG0). A partir de celui-ci, un premier modèle a été réalisé en positionnant une méthionine et puis en effectuant quelques étapes de minimisation pour modéliser un méthionyl-adénylate. Ensuite, en utilisant la structure du complexe MetRS : β -Met décrite dans la partie I et le modèle apo, un modèle d'un complexe β -Met-adénylate a été réalisé. Les parties carboxylate et ammonium de la β -Met ont été contraintes pour que leurs configurations respectives soient identiques à celles observées dans le complexe MetRS : α -Met.

III. 1. 2 Sélection des positions à muter

L'idée générale est de générer des variants en testant les acides aminés possibles à quelques positions, et de calculer l'affinité du variant pour le Met-adénylate et le β -Met-adénylate. Les variants les plus spécifiques du β -Met-adénylate sont ceux qui nous intéressent et seront étudiés expérimentalement.

Pour le choix des positions à muter, nous avons décidé de nous intéresser aux acides aminés les plus proches du substrat. (A12, L13, A256, I297 et H301).

Rappelons que le triple variant L13S-Y260L-H301L, capable d'activer l'azidonorleucine (Tanrikulu *et al.* , 2009), a une conformation proche de celle de l'enzyme liée à la Met, même en l'absence de ligand (Schmitt *et al.* , 2009). L'hypothèse faite lors de l'étude structurale était que la mutation H301L provoquait un changement de conformation des résidus aromatiques (Schmitt *et al.* , 2009). J'ai produit ce variant et résolu sa structure. Celle-ci confirme l'hypothèse faite en 2009. Ainsi, H301L est liée à des changements de conformation majeurs et n'a donc pas été utilisée dans cette étude.

Nous avons vu dans la partie précédente que muter l'A12 pouvait facilement provoquer des gênes stériques et donc interagir avec l'adénylate, la Met ou pouvait influencer la conformation de l'I297. Ce n'était donc pas un bon candidat. Ainsi, nous avons finalement choisi les positions L13, A256 et I297 comme positions à muter (Figure 57).

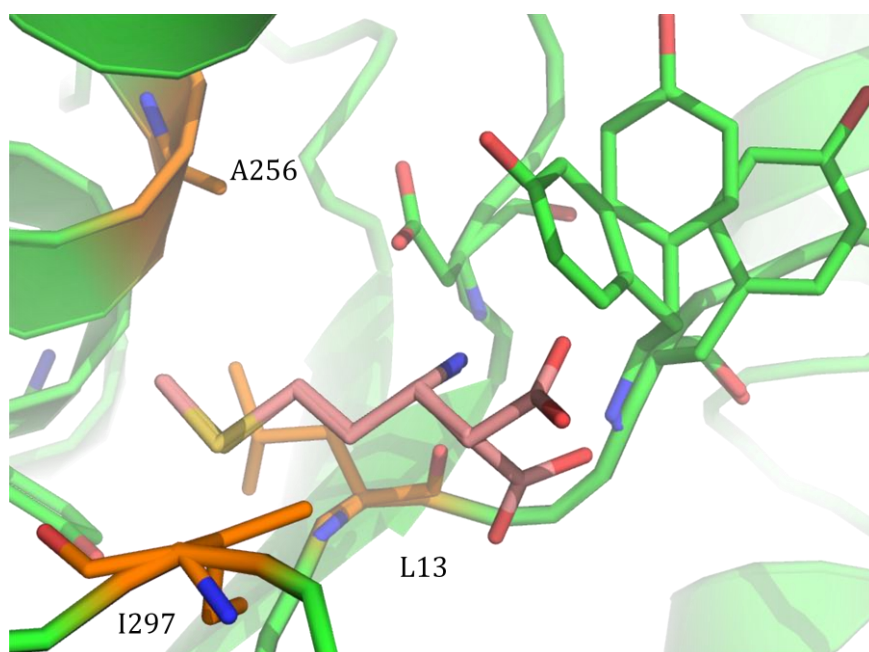


Figure 57 : Site de reconnaissance de la méthionine complexé à la β -Met. Les chaînes latérales des résidus sélectionnés pour être modifiés sont colorées en orange.

III. 1. 3 Echantillonnage

L'échantillonnage s'effectue de la façon suivante. Dans un premier temps, les chaînes latérales sont retirées de la chaîne principale. Ensuite, des chaînes latérales sont aléatoirement positionnées en utilisant une bibliothèque de positions (Tuffery *et al.*, 1991). Les énergies des interactions des rotamères sont ensuite calculées par rapport au modèle initial. On fait ensuite varier à chaque position l'acide aminé et la conformation de la chaîne latérale. A partir du modèle Apo un premier échantillonnage des mutations pour les 3 positions est effectué pour estimer l'énergie libre des séquences à l'état Apo par la méthode de Monte-Carlo (MC) adaptative (Bhattacharjee and Wallin, 2013; Simonson *et al.*, 2013; Villa *et al.*, 2018). Pour l'échantillonnage des séquences dans le complexe MetRS : α -Met-adénylate, l'énergie des variants dans l'état Apo est soustraite à celle du complexe. De ce fait, les variants sont échantillonnés en fonction de la différence APO/HOLO. Ce procédé est ensuite

appliqué au complexe MetRS : β -Met-adénylate pour obtenir les variants les plus affins pour la β -Met.

Dans le contexte de l'échantillonnage pour l'affinité, nous ne tenons pas compte de la stabilité du site actif. En effet, ces variant affins peuvent entrainer des déformations de la chaîne principale dans le site actif. Pour estimer la stabilité des variants, un modèle déplié théorique a été défini (Wernisch *et al.* , 2000; Dahiyat and Mayo, 2008). Nous calculons par la suite le changement d'énergie libre entre l'état déplié et l'état APO pour en déduire une estimation de la stabilité.

III. 2 Liste des variants proposés par l'étude *in silico*

III. 2. 1 Variants spécifiques du méthionyl adénylate

Dans un premier temps, nous souhaitons éprouver le modèle avec des variants sélectionnés pour leur affinité pour le Met-adénylate. Les variants sont comparés entre eux pour leur affinité pour le substrat avec comme référence l'affinité de l'enzyme sauvage. Les onze variants avec la meilleure affinité calculée pour le méthionyl-adénylate sont regroupés dans le Tableau suivant. La stabilité de chaque variant est également calculée.

Tableau 8 : Douze meilleurs variants proposés par l'étude *in silico* classés en fonction de leur affinité pour le méthionyl adénylate

Séquence (13/256/297)	Affinité-met-adénylate (kcal/mol)	Stabilité (kcal/mol)
CDV	-1,36	4,52
CDI	-1,32	5,92
CDC	-0,97	5,5
CDA	-0,88	5,24
ADV	-0,66	5,7
ADC	-0,4	6,73
ADA	-0,26	6,45
MAV	-0,23	1,28
MAI	-0,20	2,53
LAV	0,16	-1,33
MAC	-0,09	2,28
LAI (WT)	0	0

On voit dans cette liste que certains variants proposés ont la mutation A256D qui semble suspecte. En effet, on peut voir dans le tableau 8 que les variants avec la mutation A256D confèrent à l'enzyme une mauvaise stabilité. Effectivement, ce variant a été produit et sa

vitesse en échange isotopique ATP-PPi dépendant de la méthionine est à la limite du seuil de détection ($v_i=0,002s^{-1}$ à 2mM met). Les mutants comportant la mutation A256D seront donc écartés.

On a finalement choisi de sélectionner des variants qui présentent des valeurs de stabilité raisonnables. La première série de variants a principalement des mutations en position 13 et 297. A cette liste s'ajoutent des variants avec une proline en position 13 car, bien que cette mutation soit interdite par le processus de criblage, elle pourrait être favorable. En effet, dans l'étude de l'équipe de D. Tirrell, les variants les plus spécifiques de l'azidonorleucine avaient une proline en 13 (Tanrikulu *et al.* , 2009).

III. 2. 2 Variants spécifiques du β -méthionyl adénylate

Dans le tableau suivant, on retrouve les 20 meilleurs variants issus de l'étude *in silico*, à priori plus affins pour le β -Met-adénylate. Ceux-ci sont classés en fonction de leur affinité pour le β -Met-adénylate (Tableau 9).

Tableau 9 : Paramètres thermodynamique calculés pour les vingt meilleurs variants vis-à-vis du β -Met-adénylate.

Séquence (13/256/297)	Affinité- β -Met-adénylate (kcal/mol)	Stabilité (kcal/mol)
MSA	-4,89	3,76
MSC	-4,85	4,06
MST	-3,97	6,21
MSV	-3,95	3,06
MSS	-3,91	5,77
MAA	-3,50	2,01
MAC	-3,41	2,28
MAT	-2,70	4,56
NSA	-2,45	5,52
MAV	-2,38	1,28
MSI	-2,36	4,36
CSA	-2,18	-6,34
MAS	-2,17	3,92
NSC	-2,17	5,77
HSA	-2,12	2,65
CSC	-2,11	-6,18
LSC	-2,06	1,25
LSA	-1,97	1,05
SSA	-1,79	-2,46
ASA	-1,78	-5,92
MAD	-1,77	-9,86

Parmi les 30 meilleurs variants, nous en avons sélectionné 20 qui étaient assez proches des variants obtenus pour la méthionine pour pouvoir comparer plus directement les résultats. Nous avons volontairement laissé de côté ceux qui comportaient un acide aminé chargé. En effet, la poche de reconnaissance de la méthionine est apolaire et un résidu chargé risquerait de créer des interactions électrostatiques défavorables à la catalyse.

III. 3 Etude cinétique

III. 3. 1 Caractérisation en présence de méthionine

Les constantes michaeliennes mesurées avec certains des variants de la première série ont été regroupées dans le tableau suivant. Les variants y sont dans le même ordre que dans la classification issue du modèle *in silico*.

Tableau 10 : Paramètres cinétiques de la réaction d'échange [32P] PPI-ATP avec la méthionine classés. L'ordre des variants est l'ordre décroissant des affinités calculées.

Variant	$k_{cat\ Met} (s^{-1})$	$K_m\ Met (mM)$	$k_{cat} / K_m (s^{-1}/\mu M)$
MAV	$12 \pm 0,3$	$0,42 \pm 0,06$	28
LAV	$27 \pm 0,4$	$0,41 \pm 0,04$	66
MAC	$13 \pm 0,1$	$0,91 \pm 0,04$	14
WT (LAI)	$57 \pm 0,9$	$0,02 \pm 0,002$	3100
LAC	$43 \pm 1,4$	$0,40 \pm 0,07$	108
MAT	$5,2 \pm 0,4$	$1,1 \pm 0,4$	4,9
CAV	$51 \pm 0,9$	$2,2 \pm 0,13$	22
CAI	$57 \pm 0,6$	$0,14 \pm 0,01$	410
LAT	$22 \pm 0,5$	$0,85 \pm 0,05$	26
CAC	$31 \pm 0,3$	$3,2 \pm 0,1$	9,8
SAI	$46 \pm 1,2$	$0,14 \pm 0,02$	330
SAC ¹	$37 \pm 2,4$	$3,0 \pm 0,3$	12
PAI ¹	$21 \pm 0,02$	$7,4 \pm 2$	2,9

On ne trouve pas de cohérence claire avec entre la classification théorique et la classification expérimentale. Certains variants ont été prédits meilleurs que la MetRS sauvage. Or, selon les résultats obtenus, aucun variant testé n'a de meilleures constantes que celle de l'enzyme WT.

On peut supposer que ces différences entre le modèle théorique et les résultats expérimentaux sont dues aux limites intrinsèques du modèle qui n'optimise pas strictement l'affinité pour l'état de transition.

III. 3. 2 Evaluation de l'activité catalytique des variants en présence de β -Met

Pour mesurer l'activité des variants avec la β -Met il aurait fallu idéalement mesurer leur vitesse de réaction à l'état pré-stationnaire. Or nous avons observé avec le variant A12S qu'il n'était pas forcément possible de mesurer les constantes k_f et K_{met}^{ATP} . En effet, la variation de fluorescence n'est pas observable si l'affinité du substrat est trop faible.

Nous avons donc utilisé la β -Met contaminée pour mesurer les k_{cat} et K_m apparents ($k_{cat \beta-Met \text{ app}}$ et $K_{m \beta-Met \text{ app}}$) de la réaction d'échange isotopique ATP-PPi. Ces valeurs reflètent l'inhibition exercée par la β -Met sur l'activation de la Met contaminante. Les valeurs des constantes mesurées se trouvent dans le tableau suivant.

Tableau 11 : Paramètres de la réaction d'échange isotopique ATP :PPi mesurés en présence de méthionine ou de β -méthionine. Dans ces mesures le stock de β -Met utilisé est contaminé à 0.09% par de la méthionine.

- Valeurs calculés à partir de la pente à l'origine
- NM : Non mesurable

	k_{cat} (s ⁻¹)	K_m (mM)	k_{cat} / K_m (s ⁻¹ /mM)	$k_{cat \beta-Met \text{ app}}$ (s ⁻¹)	$K_{m \beta-Met \text{ app}}$ (mM)	$k_{cat \beta-Met \text{ app}} / K_{m \beta-Met \text{ app}}$ (s ⁻¹ /mM)
WT	57 ± 0,9	0,020 ± 0,002	3073	1,2 ± 0,2	0,57 ± 0,1	1,8
CAI	57 ± 0,6	0,14 ± 0,01	412	1,0 ± 0,003	3,3 ± 0,3	0,29
SAI	46 ± 1,2	0,14 ± 0,02	326	1,1 ± 0,01	4,9 ± 2,2	0,21
LAC	43 ± 1,4	0,40 ± 0,07	108	0,3 ± 0,003	3,4 ± 0,2	0,080
LAV	27 ± 0,4	0,41 ± 0,04	66	0,5 ± 0,005	5,9 ± 0,2	0,080
MAV	12 ± 0,3	0,42 ± 0,06	28	0,2 ± 0,006	11 ± 0,8	0,019
MSI	20 ± 0,5	0,73 ± 0,07	27	0,2 ± 0,006	11 ± 1,0	0,020
SSI	52 ± 1,2	0,75 ± 0,07	70	0,5 ± 0,015	8,4 ± 0,5	0,059
LAT	22 ± 0,5	0,85 ± 0,05	26	0,2 ± 0,003	7,2 ± 0,4	0,021
MAC	13 ± 0,1	0,91 ± 0,04	14	0,08 ± 0,002	6,2 ± 0,6	0,013
MAT	5,2 ± 0,4	1,1 ± 0,4	4,9	0,03 ± 0,002	6,3 ± 0,9	0,0054
CAV	51 ± 0,9	2,2 ± 0,13	23	0,5 ± 0,02	22 ± 6,5	0,020
SAC	37 ± 2,4	3,0 ± 0,3	12	0,2 ± 0,003	12 ± 0,5	0,013
CAC	31 ± 0,3	3,2 ± 0,5	9,8	0,07 ± 0,004	11 ± 1,2	0,0062
LAS ¹	1,7 ± 0,2	5,6 ± 0,78	0,4 ± 0,01	0,003 ± 3*10 ⁻⁴	6,8 ± 1	0,0004±3*10 ⁻⁵
MSV ¹	8,3 ± 0,9	5,4 ± 0,76	1,5 ± 0,08	0,025 ± 5*10 ⁻³	14 ± 3	0,0017±7,2*10 ⁻⁵
PAI ¹	20 ± 2,2	6,3 ± 0,89	3,1 ± 0,15	0,05 ± 0,01	15 ± 4	0,003 ± 1,6*10 ⁻⁴
SSC	38 ± 2,3	8,3 ± 0,9	4,55	0,09 ± 0,003	12 ± 1,0	0,0071
MST ¹	3,4 ± 0,4	7,4 ± 1,1	0,5 ± 0,01	0,004±10 ⁻³	7,5 ± 1,6	5*10 ⁻⁴ ±3*10 ⁻⁵
LAA ¹	5,7 ± 0,8	12 ± 2,1	0,5 ± 0,01	0,007±10 ⁻³	6,7 ± 1,2	0,001 ± 6*10 ⁻⁵
MSS ¹	0,1 ± 0,01	5,8 ± 0,95	0,02 ± 10 ⁻³	NM	NM	NM
MSA ¹	0,3 ± 0,06	18 ± 4	0,02 ± 10 ⁻³	NM	NM	NM
SSA ¹	0,9 ± 0,1	6,5 ± 1,3	0,01 ± 001	NM	NM	NM

De plus, l'analyse des courbes de vitesse d'échange en fonction de la concentration de la β -Met nous permet de déterminer K_i , la constante d'inhibition, et ce en utilisant l'équation décrite dans la partie I. Nous avons fixé le taux de contamination λ à 0,09%, conformément à l'analyse de l'enzyme WT. D'autre part, pour minimiser la quantité de travail et donc cribler le plus grand nombre de mutants, nous avons mesuré la vitesse de formation du β -méthionyl-adénylate à l'état pré-stationnaire à une concentration de β -Met fixée à 10 mM pour chaque variant choisi. Pour calculer l'efficacité catalytique d'activation, nous avons supposé que K_i représentait une bonne estimation de $K_{\beta\text{-Met}}^{ATP}$. A partir de K_i , nous avons déduit $k_{f\beta\text{-Met}}$ par l'équation :

$$k_{f\beta\text{-Met}} = k_{10\text{mM } \beta\text{-Met}} \frac{[S] + K_i}{[S]}$$

Avec $[S]=10\text{mM}$

Les valeurs ainsi estimées de $\frac{k_{f\beta\text{-Met}}}{K_i}$ sont rassemblées tableau 12. Afin d'évaluer la spécificité de chaque variant, nous avons fait le rapport de l'efficacité catalytique d'échange ($\frac{k_{f\beta\text{-Met}}}{K_i}$) sur l'efficacité d'activation de la β -Met ($\frac{k_{cat}}{K_{m\text{met}}}$).

Les variants recherchés sont ceux pour lesquels le rapport ($\frac{(k_f/K_i)_{\beta\text{-Met}}}{(k_{cat}/K_m)_{Met}}$) est plus grand que pour l'enzyme sauvage, et qui pourraient donc avoir une spécificité vis-à-vis de la β -Met améliorée.

Nous sommes bien conscients que cette approche n'est pas idéale puisqu'elle compare des paramètres d'échange isotopique ATP-PPi dépendant de la méthionine à des paramètres d'activation de la β -Met à l'état pré-stationnaire. Néanmoins, cette approche permet de réaliser les expériences plus simples pour aboutir à une évaluation rapide des meilleurs candidats parmi les variants. Pour les variants retenus, cette première analyse devra être poursuivie par des études plus complètes d'activation de la Met et de la β -Met à l'état pré-stationnaire.

Tableau 12 : Efficacités catalytiques d'activation de la méthionine ou de la β -méthionine (ND : non déterminé)

1 : Trop d'approximation pour calculer une erreur pour k_f excepté pour WT et VI298

2 : Calculé à partir des valeurs de k_f (Partie I)

	$\frac{k_{cat\ Met}}{K_m\ Met}$	K_I (mM)	$\frac{K_I}{K_m}$	$k_{10mM\ \beta-Met}$ ($\times 10^{-3} s^{-1}$)	$k_{f\ \beta-Met}^1$ ($\times 10^{-3} s^{-1}$)	$k_{f\ \beta-Met}/K_I$ (s^{-1}/M)	$\frac{(k_f/K_I)_{\beta-Met}}{(k_{cat}/K_m)_{Met}} (\times 10^{-6})$
CAC	9,8	$6,3 \pm 1$	1,9	$13 \pm 0,3$	21	3,4	345
MAC	14	$7,7 \pm 0,2$	8,5	$9,0 \pm 0,2$	16	2,1	145
SAC	12	20 ± 1	6,6	$7,0 \pm 0,1$	22	1,1	91
VI298	5200	$0,21 \pm 0,006$	18	56 ²	57 $\pm$ 5	270	52
LAC	107	$3,5 \pm 0,1$	8,9	$13 \pm 0,3$	18	5,0	47
WT	3073	$0,37 \pm 0,01$	20	49 ²	51 $\pm$ 3	138	45
LAT	26	$6,2 \pm 0,1$	7,3	$3,0 \pm 0,06$	5,2	0,83	32
MAV	28	$7,1 \pm 0,3$	17	$3,0 \pm 0,09$	5,4	0,75	27
CAI	412	$2,6 \pm 0,06$	19	$16 \pm 0,5$	21	8,9	19
SAI	325	$4,2 \pm 0,08$	30	$16 \pm 0,3$	23	5,4	16
CAV	23	$38 \pm 2,9$	17	$3,0 \pm 0,1$	13	0,35	15
LAV	66	$11 \pm 0,8$	27	$4,0 \pm 0,5$	8,0	0,74	11
SSI	70	$7,5 \pm 0,1$	10	<2	<3,5	ND	ND
MAT	4,9	$8,2 \pm 0,3$	7,7	<2	<3,6	ND	ND
MSI	27	$8,4 \pm 0,2$	11	<2	<3,7	ND	ND
MST ¹	$0,46 \pm 5 \times 10^{-4}$	19 ± 2	2,1	<2	<5,7	ND	ND
LAS ¹	$0,31 \pm 0,01$	21 ± 2	3,4	<2	<6,1	ND	ND
MSV ¹	$1,5 \pm 0,08$	39 ± 4	5,7	<2	<10	ND	ND
PAI ¹	$3,1 \pm 0,1$	47 ± 2	6,4	<2	<11	ND	ND
SSC	4,5	60 ± 20	7,4	<2	<14	ND	ND
LAA ¹	$0,48 \pm 0,02$	180 ± 140	17	<2	<38	ND	ND

Parmi les variants testés, la moitié ne montre pas de variation de fluorescence suffisante pour que la vitesse de formation du β -Met-adénylate soit mesurée. Pour mieux visualiser les rapports $\frac{(k_f/K_I)_{\beta-Met}}{(k_{cat}/K_m)_{Met}}$, nous les avons représentés dans le graphique suivant.

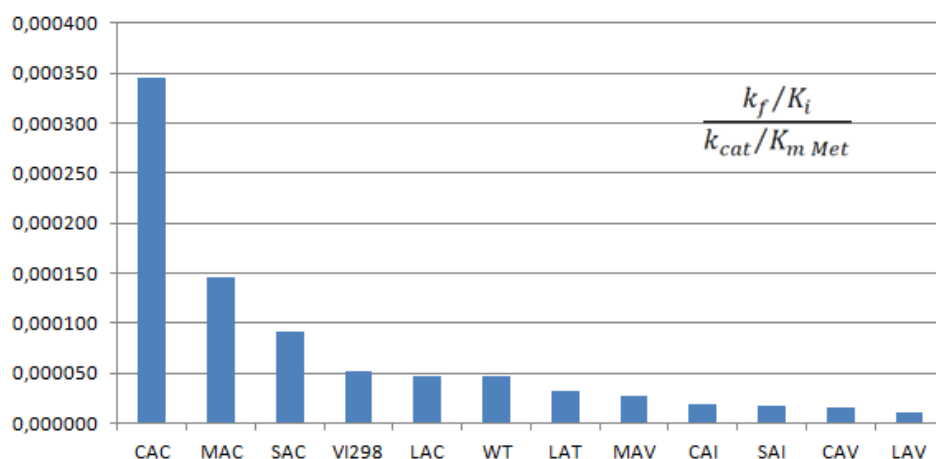


Figure 58 : Graphique comparant les rapports des $\frac{(k_f/K_i)_{\beta\text{-Met}}}{(k_{cat}/K_m)_{Met}}$ des différent variants produits et testé à partir de l'étude *in silico*

On peut voir que le variant VI298 est très légèrement plus spécifique de la β -Met que l'enzyme sauvage comme cela a été montré dans la partie I. Les résultats montrent que les meilleurs candidats pour améliorer la spécificité vis-à-vis de la β -Met sont les variants MAC, SAC et CAC. L'étude de ces variants devra être approfondie.

III. 4 Conclusion

Nous avons évalué l'importance des groupements chimiques du substrat et leurs contributions à la sélectivité de la poche de reconnaissance de la méthionine. En particulier la taille du soufre et la répartition des charges sur l'amine de la méthionine sont très importants. Nous avons ensuite trouvé un variant, A12S, qui favorisait la conformation « active » de la β -méthionine dans la poche de reconnaissance de la méthionine. Toutefois cette mutation diminue grandement l'activité de l'enzyme car elle empêche le positionnement adéquat du phosphate α de l'ATP. En effet, on ne peut d'ailleurs pas mettre en évidence l'activation de la β -Met à l'état pré-stationnaire. Nous avons pu mesurer expérimentalement l'activité de certains des variants proposés par l'étude *in silico*. Il a toutefois été compliqué de corréler les résultats. Ceci pourrait être lié aux limites du modèle et au fait que l'enzyme change de conformation lors de la réaction.

Malgré tout, les résultats préliminaires ont permis de trouver des mutants qui pourraient être plus spécifiques pour la β -Met. Il reste maintenant à mesurer leurs véritables k_f et $K_{\beta Met}^{ATP}$.

Partie III

Afin d'améliorer le système de traduction pour optimiser l'incorporation des AA β dans des peptides *in vivo*, nous avons décidé de mettre en place une méthode d'évolution dirigée. Nous avons vu dans l'introduction qu'il existait déjà plusieurs méthodes pour améliorer les AARS et ARNt orthogonaux. Bien que les méthodes pour générer des variants puissent être adaptées, les méthodes de criblage sont difficilement applicables pour la β -méthionine. En effet, ce n'est pas la fonctionnalité chimique de la chaîne latérale qui est modifiée mais la chaîne principale. Nous avons donc décidé de développer une méthode d'évolution dirigée spécifique aux acides aminés β en collaboration avec Valérie Pezo et Philippe Marlière. En pratique, une fois le système de criblage au point, nous comptons utiliser un chemostat décrit par Philippe Marlière pour faire évoluer une souche bactérienne dans une direction donnée, en appliquant une pression sélective (Mutzel *et al.* , 2007).

I Méthode d'évolution dirigée

I. 1 Principes

Nous cherchons à créer une pression de sélection qui va forcer à mieux incorporer les acides aminés β . Nous devons ainsi trouver une enzyme dont l'activité sera rendue dépendante de l'incorporation de β -Met. En alternant des conditions de croissance non permissives et permissives, nous sélectionnerons des bactéries ayant une prolifération de plus en plus rapide et stable en conditions non permissives.

Cette méthode a été éprouvée, notamment pour faire évoluer une souche auxotrophe pour la thymine (Δ deoCABD thyA::Erm p::deoA+) et la rendre résistante à l'absence de thymine (de Crécy-Lagard *et al.* , 2001). Dans ce cas, la souche a subi des cycles de croissance en présence d'une concentration minimale de thymine afin qu'elle évolue pour ne plus en avoir besoin.

I. 2 Choix d'une enzyme de sélection.

Une première contrainte majeure était de trouver une activité enzymatique essentielle pour la croissance cellulaire, dont l'activité n'était *a priori* pas contournable. Pour ce faire nous avons décidé de travailler sur la voie métabolique du désoxythymidylate monophosphate (dTMP) (Figure 59). Celui-ci est primordial à la survie bactérienne.

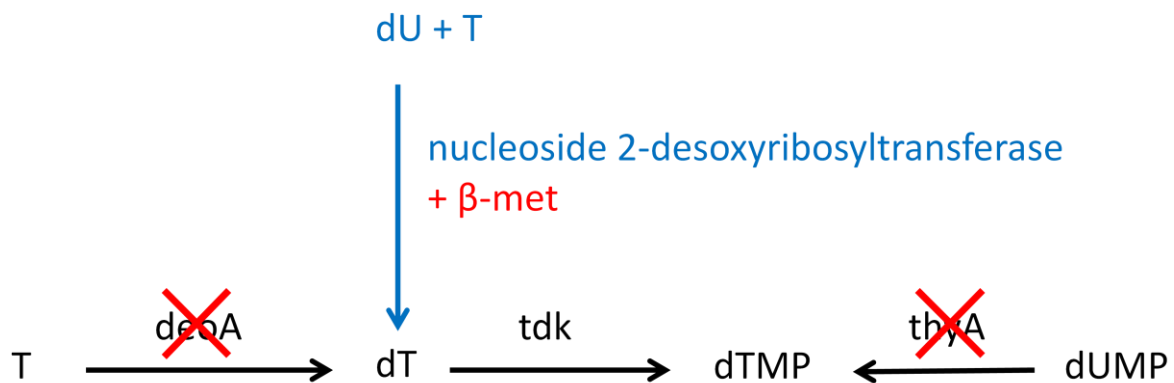


Figure 59 : Schéma de la voie métabolique du dTMP en noir chez *E. coli* en bleu la voie alternative présente chez *Lactobacillus leichmannii* et *Lactobacillus helveticus*

Chez *E. coli*, deux voies peuvent fabriquer du dTMP. La première voie, directe, utilise la thymidylate synthase (*thyA*) pour produire le dTMP par déméthylation du dUMP. La voie indirecte passe par la synthèse de désoxythymidine (dT) et phosphate à partir de thymine (T) et de ribose monophosphate, catalysée par la thymidylate phosphorylase (*deoA*). La désoxythymidine est ensuite phosphorylée en dTMP par la thymidylate kinase (*tdk*). Si ces deux voies sont bloquées simultanément, en inactivant *thyA* et *deoA*, le dTMP ne peut pas être produit et la souche n'est pas viable. On peut alors ajouter le gène de la nucléoside 2-desoxyribotransférase (NDT) trouvé chez certaines bactéries comme *Lactobacillus leichmannii*. Cette enzyme est capable de fabriquer du dT en transférant le désoxyribose de la desoxyuridine (dU) à la thymine. Ainsi, en présence de dU et de T, la croissance de la souche devient possible si une NDT est active. Nous avons retenu ce système car il a été montré très robuste dans sa dépendance à la NDT (de Crécy-Lagard *et al.*, 2001). De plus, la NDT est un hexamère constitué par un polypeptide de relativement petite taille (157 acides aminés), dont la structure 3D et le mécanisme catalytique sont bien connus.

II Construction d'une NDT dépendante de l'incorporation d'une β -méthionine

Afin de créer un crible efficace dépendant de l'incorporation de la β -méthionine, nous avons défini la NDT comme l'enzyme dans laquelle la β -méthionine doit s'incorporer pour que la souche soit viable. Globalement, le système que l'on doit obtenir est schématisé sur la figure suivante.

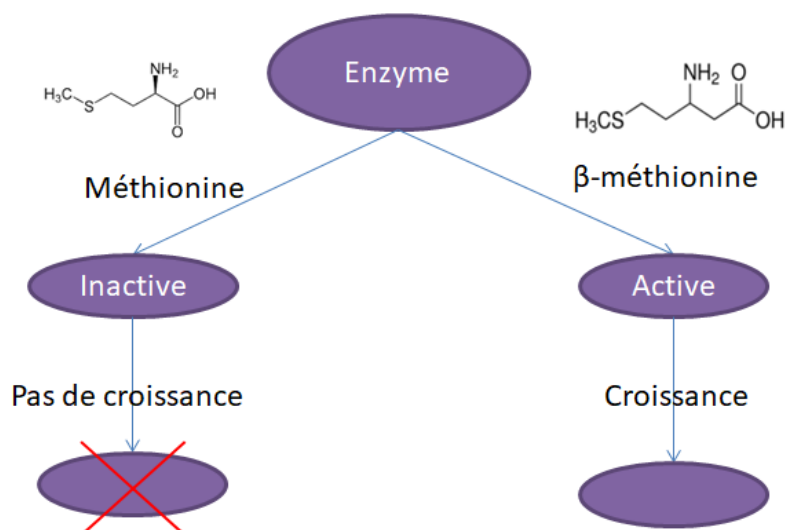


Figure 60 : Représentation schématique de la méthode d'évolution dirigée

Notre stratégie pour créer la dépendance de l'enzyme à l'incorporation de β -méthionine est basée sur les propriétés de l'acide aminé, c'est à dire l'allongement de la chaîne principale, le changement d'alternance des chaînes latérales dans la chaîne peptidique et de propension des AA- β à faire des hélices. L'idée directrice a été de remplacer deux acides aminés par un seul. Ainsi en remplaçant deux acides aminés α par un AA- β , l'alternance des chaînes latérales est conservée et la longueur bien que diminuée est augmentée par rapport à celle obtenue si l'on incorporait un AA- α (Figure 61). On peut donc espérer trouver des sites auxquels l'incorporation d'un AA- α donnera une enzyme inactive alors que celle d'un AA- β restaurera l'activité (Figure 61).

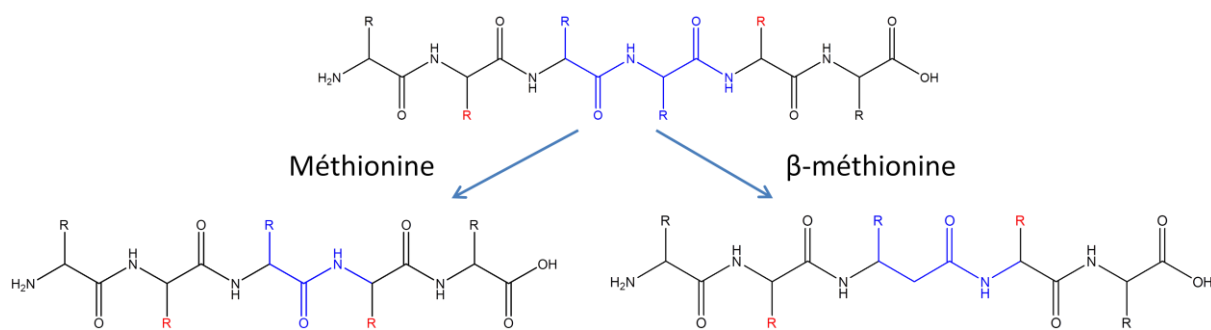


Figure 61 : Correspondance du remplacement de deux acides aminés par un AA- α (gauche) ou un AA- β (droite)

II. 1 Choix des positions à muter

II. 1. 1 La N-Desoxyribosyltransferase

Les N-desoxyribosyltransferases catalysent le transfert du groupement 2'-desoxyribosyl d'un desoxynucleoside donneur à une base nucléique acceptrice. Il en existe deux classes. La première, appelée purine desoxyribosyltransferase (PDT), transfère le groupement désoxyribose d'une purine à une autre. La seconde classe catalyse le transfert du groupement 2'-desoxyribosyl entre des purines ou des pyrimidines, mais ayant une préférence pour les desoxypyrimidines comme donneurs (Porter *et al.* , 1995). Ces dernières sont aussi appelées nucléoside desoxyribosyltransferases (NDT).

Ces enzymes ont été étudiées chez plusieurs organismes, en particulier la NDT de *L. leichmannii* (Armstrong *et al.* , 1996a) et la PDT de *L. helveticus* (Anand *et al.* , 2004). Leurs structures sont connues avec et sans substrat. Ces protéines sont des hexamères constitués de trois homodimères (Figure 62).

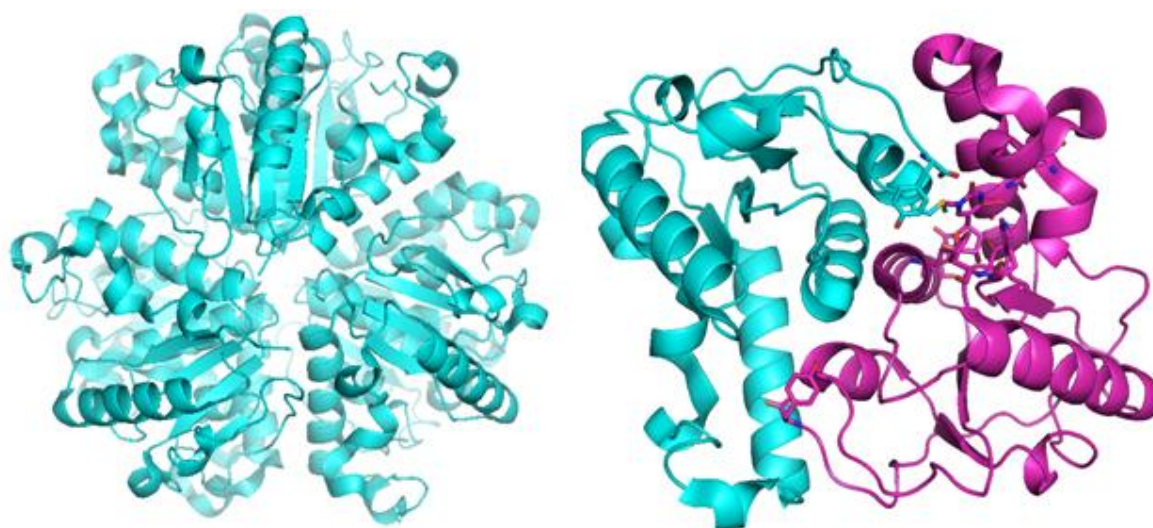


Figure 62 : Structure de la NDT hexamérique de *Lactobacillus leichmannii* cocrystallisée avec de la 5-méthyl-2'-deoxypseudouridine (analogue de la thymidine) à gauche. Structure d'un dimère de NDT à droite (1F8Y)

II. 1. 2 Choix des positions à muter

Afin d'identifier des positions essentielles à l'activité de la nucléoside 2-Desoxyribosyltransferase susceptibles d'être remplacées par de la β -méthionine, nous avons soigneusement analysé la structure 3D de l'enzyme et les résultats biochimiques disponibles (Armstrong *et al.* , 1996a). De cette analyse, huit mutations ont été réalisées, où nous avons choisi de remplacer deux codons par un codon méthionine. Par souci de conservation des chaînes latérales, la série de mutagenèses a principalement ciblé des méthionines ou des résidus comportant des chaînes latérales apolaires.

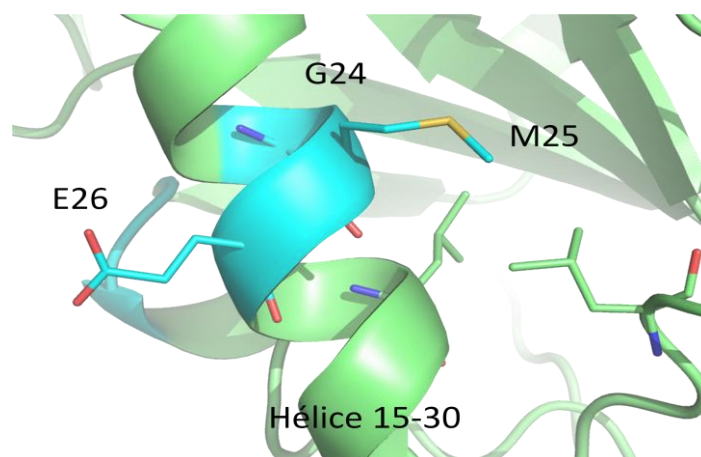


Figure 63 : Structure la NDT de *L. leichmannii* à proximité des résidus 24 à 25.

La première région mutée, 24-25 et 25-26, a été choisie pour deux raisons. D'une part nous souhaitons conserver la chaîne latérale de la méthionine qui semble importante pour la structure de la protéine. D'autre part, cette méthionine est dans une hélice et la β -Met permettrait de conserver cette conformation (figure 63). En faisant une mutation à cette position, on joue sur la formation de l'hélice et le positionnement de la méthionine, en espérant que seule une β -Met puisse permettre de placer la chaîne latérale à une position adéquate.

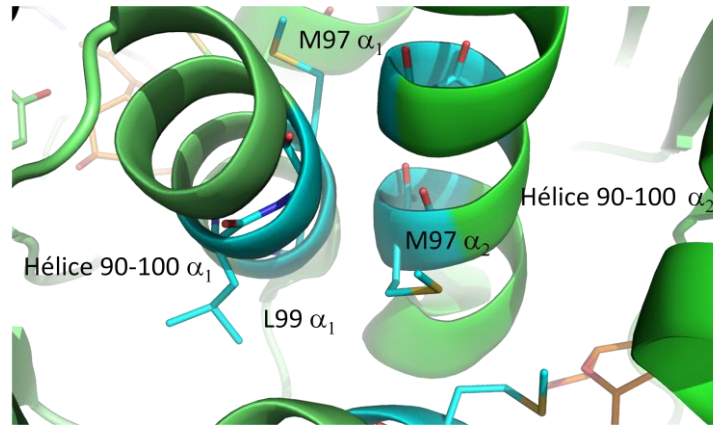


Figure 64 : Structure la NDT de *L. leichmannii* à proximité des résidus 96-99.

La seconde région mutée se trouve sur l'hélice 90-100, où l'on trouve le E98 qui est primordial pour la catalyse (Armstrong *et al.*, 1996a). De plus cette hélice est un élément intervenant dans la dimérisation par des interactions de type Van der Waals. Il est donc important de garder la chaîne latérale de la méthionine à ces positions (Figure 64). Les mutations GM96-97M et LG99-100M traduisent ces attentes.

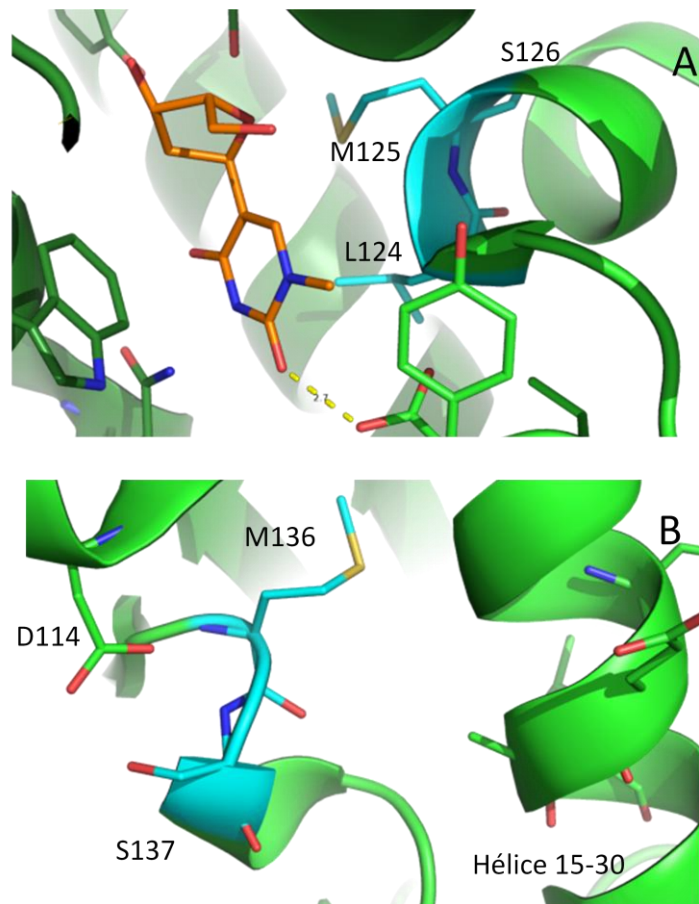


Figure 65 : Structure la NDT de *L. leichmannii* à proximité des résidus 124-126 (A) et des résidus 136 et 137 (B).

La M125 fait partie des acides aminés qui entourent la poche de reconnaissance du substrat. De plus, celle-ci se trouve dans une hélice (figure 65).

La M136, tout comme la M25, a un rôle principalement structural. Ici, on se retrouve au début d'une hélice qui serait décalée avec un AA- α mais dont la position pourrait être mieux conservée avec un AA- β (figure 65).

Les études cinétiques effectuées sur la NDT de *L. leichmannii*, ont montré l'importance de plusieurs acides aminés comme l'E98 cité précédemment, mais aussi la Y157, le résidu C-terminal. Il a été montré que son excision diminuait grandement la vitesse de la réaction (Armstrong *et al.* , 1996b). De plus, l'étude structurale de l'enzyme cocrystallisée avec le substrat montre que le groupement carboxyle C-terminal interagit avec le substrat dans le site actif (figure 66). Cette position est particulièrement porteuse d'espoir. En effet, on peut s'attendre à ce que le remplacement des deux derniers résidus par un seul ne soit pas suffisant pour positionner le carboxylate en position favorable, alors que le remplacement des deux résidus par une β -Met, pourrait permettre d'atteindre une position plus adéquate pour la catalyse.

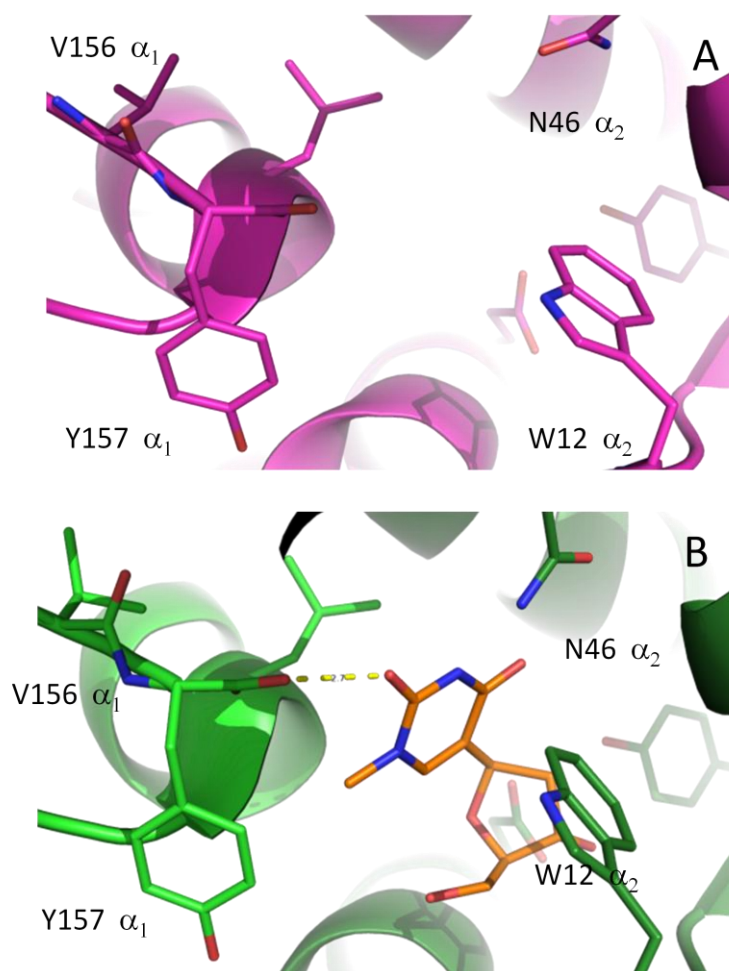


Figure 66 : Liaison du 5-méthyl-2'-deoxypseudouridine à la NDT

A : Structure de la NDT de *Lactobacillus leichmannii* (1F8X). B : structure du complexe NDT: 5-méthyl-2'-deoxypseudouridine (1F8Y). L'interaction entre le carboxylate en C-terminal et le substrat est représentée par des traits jaunes.

Le résidu M82 a été écarté car, cet acide aminé se situant dans un brin β , la présence d'un AA- β au sein de celui-ci ne permettrait pas a priori la formation du feuillet à cause de l'orientation de la chaîne principale.

Sur chacune des positions sélectionnées, on espère une désactivation de l'enzyme ou une grande baisse de son activité si la méthionine est incorporée, et une activation de l'enzyme ou une amélioration de son activité si la β -Met est incorporée. Finalement, les mutations retenues sont répertoriées dans le tableau suivant. Les variants ont été construits dans le contexte du gène de la NDT de *L. leichmannii* inséré dans un plasmide sous contrôle d'un promoteur constitutif (un promoteur modifié venant du bactériophage T5).

Tableau 13 : Liste des mutations réalisées dans le gène codant pour la NDT (Figure 67)

Mutation	AM24-25M	ME25-26M	GM96-97M	LG99-100M	LM124-125M	MS125-126M	MS136-137M	VY156-157M
----------	----------	----------	----------	-----------	------------	------------	------------	------------

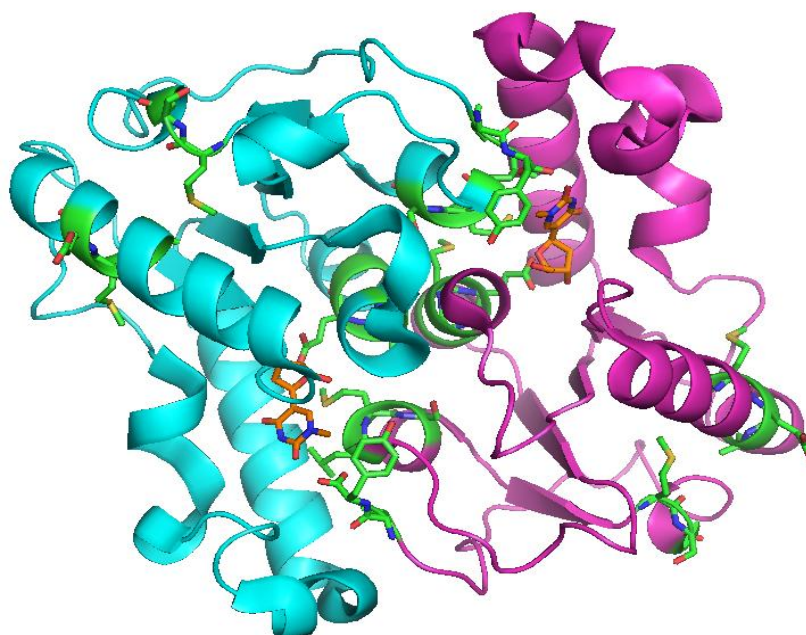


Figure 67 : Structure d'un dimère de NDT.

Le substrat est représenté en orange et les positions mutés en vert (1F8Y). Les deux monomères sont colorés en rose et en bleu.

II. 2 Test des variants en présence de méthionine

Pour tester les mutations de l'enzyme il a été nécessaire d'introduire les plasmides dans une souche qui ne produisait pas de dT et qui avait donc besoin soit d'un apport extérieur en dT, soit de NDT active pour produire du dT à partir de dU. Cette souche, THY0, nous a été fournie par nos collaborateurs, Philippe Marlière et Valérie Pezo (Marlière *et al.* , 2011). Dans un premier temps, nous avons vérifié que, comme espéré, chacune des mutations conduisait bien à une NDT inactive. Chaque plasmide muté a été introduit dans la souche THY0. La croissance a été testée en présence de dT ou en présence de dU+T sur une boîte Müller-Hinton agar, qui est un milieu riche sans dT. Ces cultures ont été faites à 30°C pour mieux observer les différences de vitesses de croissance (Figure 68).

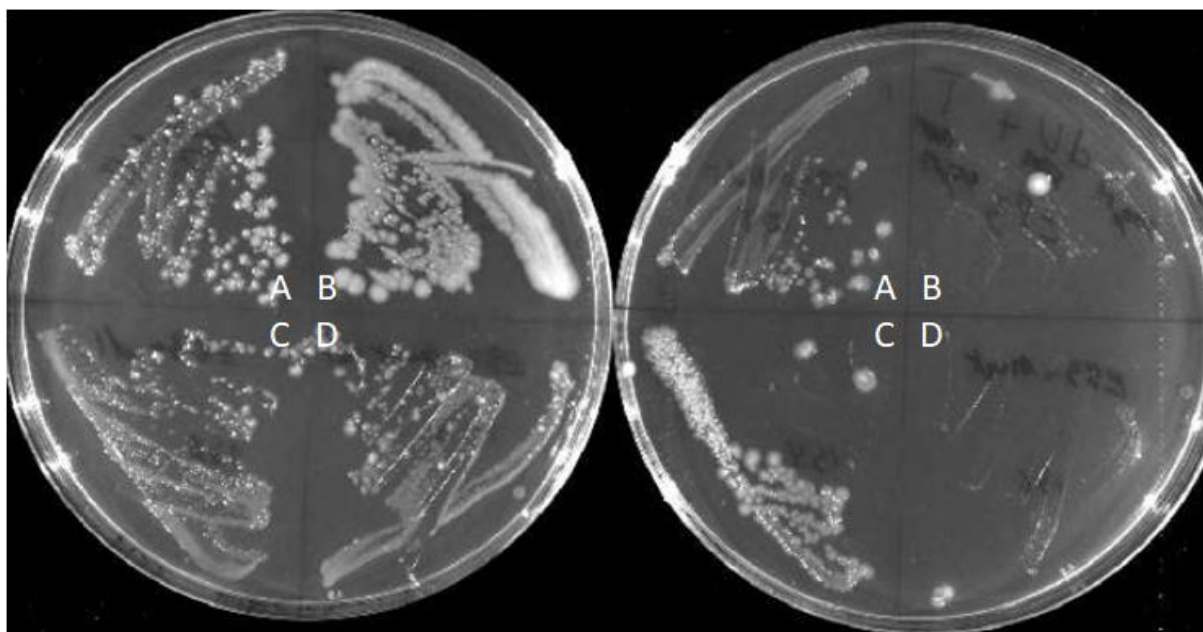


Figure 68 : Culture sur boîte MH dT (300µM) à gauche et DU+T (300µM) à droite après 48h

A : Souche THY0 + plasmide NDT WT

B : Souche THY0

C : Souche THY0 + plasmide NDT VY156-157M

D : Souche THY0 + plasmide NDT MS136-137M

Ces tests ont été faits sans antibiotique dans le milieu pour pouvoir contrôler en parallèle la souche dépourvue de plasmide. Cependant, toutes les souches avec un plasmide ont été stabilisées au préalable en présence d'ampicilline, pour s'assurer de la présence du plasmide. Aucun des plasmides exprimant les gènes mutés n'a permis la croissance dans la condition non permissive, excepté le variant C-ter VY156-157M qui a donné une croissance lente (Tableau 14). Nous avons vérifié en parallèle que la souche dépourvue de plasmide ne se développe pas alors que celle avec le plasmide portant le gène de l'enzyme non modifié se développe. L'inhibition de l'activité de l'enzyme par les mutations est donc un succès.

Tableau 14 : Résultats des tests de criblage négatif

Mutation	dT	dU+T
Souche vide	ok	non
NDT-WT	ok	ok
AM24-25M	ok	non
ME25-26M	ok	non
GM96-97M	ok	non
LG99-100M	ok	non
LM124-125M	ok	non
MS125-126M	ok	non
MS136-137M	ok	non
VY156-157M	ok	Croissance lente

Il reste alors à tester si l'activité peut être restaurée en présence de β -méthionine. Ces tests ont été effectués. Toutefois aucun des variants de la NDT n'a donné de résultat positif. Nous avons à ce stade voulu déterminer si la β -méthionine pouvait effectivement être importée dans le cytoplasme de la bactérie.

III Perméabilité d'*E. coli* à la β -Met

III. 1 Test de toxicité

Dans un premier temps, nous avons fait l'hypothèse que si la β -Met entre dans le cytoplasme d'*E. coli*, sa liaison à la MetRS (voir partie I) et/ou son incorporation dans les protéines pourrait provoquer une toxicité. Des tests de toxicité avec une souche auxotrophe pour la méthionine (Δ MetA) ont été effectués. Ces tests ont été faits sur boîte en milieu minimum dans plusieurs conditions différentes. Dans un premier temps nous avons utilisé une concentration de β -Met constante sur la boîte. Dans un second temps, nous avons déposé la β -méthionine dans un puits pour créer un gradient de concentration (Lemeignan *et al.*, 1993).

Lorsque l'acide aminé est en concentration homogène sur la boîte, on observe une très légère croissance. Si la β -Met est déposée dans un puits, on observe de la croissance uniquement à haute concentration (>100mM dans le puits). De plus, à haute concentration d'AA dans le puits, on observe un halo de toxicité en plus du petit halo ténu de pousse (très peu visible). On suppose donc que c'est la méthionine contaminante qui permet à la souche auxotrophe de se développer et que la β -Met provoque la toxicité. Cette conclusion est confirmée lorsque ce test est réalisé avec de la β -Met traitée à la MGL. On n'observe alors plus de croissance. En revanche, cela ne permet pas de conclure si la β -Met entre dans la cellule ou si elle inhibe le transport actif de la méthionine.

Nous avons donc réalisé des tests de toxicité sur une souche d'*E. coli* prototrophe pour la méthionine (MG1655). Ces expériences n'ont montré aucune toxicité de la β -Met. On ne peut donc finalement pas conclure avec certitude si la β -Met entre ou non dans la cellule. Ces expériences ont été faites en collaboration avec Valérie Pezo

III. 2 Perméabilité à la β -Met via un dipeptide

Afin d'assurer la perméabilité à la β -Met, nous avons mis en place une stratégie basée sur l'utilisation d'un dipeptide. En effet, chez *E. coli*, il existe un système de transport actif qui permet de faire rentrer certains dipeptides (Perryt and Gilvarg, 1984). De plus, il a été montré qu'il était possible d'utiliser ce procédé pour faire entrer d'autres molécules structurellement proches d'un acide aminé (Kingsbury *et al.*, 1984).

III. 2. 1 Synthèse du dipeptide β -Met-His

L'idée générale est d'utiliser un dipeptide contenant la β -Met et un acide aminé α , puis de tester son apport dans une souche auxotrophe pour l'acide aminé α . Une fois dans le cytoplasme, le dipeptide sera clivé par la N-peptidase. La β -méthionine et l'acide aminé α seront alors disponibles.

Il a été décidé de synthétiser du dipeptide L- β^3 -Met-His. Le peptide synthétisé est L- β^3 -Met-His est non His-L- β^3 -Met pour deux raisons. La première est que l'extrémité N-ter de la L- β^3 -Met est la même que celle de la méthionine ce qui augmente l'analogie avec le dipeptide Met-His et la seconde est que la synthèse est plus simple dans ce sens. J'ai réalisé cette synthèse dans des conditions classiques à partir d'une résine His-Wang, où l'histidine est attachée en C-ter à la résine. Le principe d'une synthèse peptidique sur support solide est simple. Un acide aminé protégé est fixé à une résine puis celui-ci est déprotégé avant d'y fixer l'acide aminé suivant qui est lui aussi protégé. Ce cycle est répété jusqu'à ce que le peptide soit terminé. Dans ce cas, il suffit simplement d'effectuer une dernière déprotection puis de cliver le peptide (Figure 69).

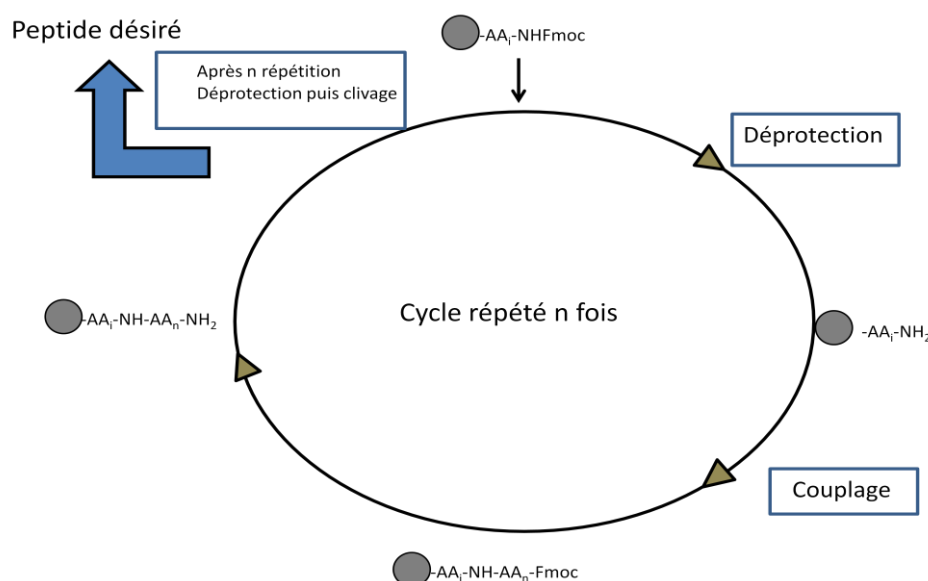


Figure 69 : Schéma représentant le concept de la synthèse peptidique en phase solide

Dans le cas de la synthèse du dipeptide, il n'y a donc qu'un seul AA à ajouter. A la fin de la synthèse, des impuretés sont précipitées à l'éthanol. Nous avons vérifié qu'il n'y avait que des traces (<0.1%) d'histidine par LC-MS.

III. 2. 2 Test de l'incorporation du dipeptide

Dans un premier temps, nous avons vérifié que le dipeptide entrainait dans *E. coli* en testant la levée d'une auxotrophie à l'histidine (EV1, Δ His). Les bactéries ont été cultivées en milieu riche avant d'être lavées puis resuspendues pour faire un dépôt homogène sur la boîte afin d'observer des clones isolés (5ml, 0,04 DO₆₀₀). Sur la figure suivante, on peut observer un halo de croissance en fonction de la concentration du dipeptide dans le puits. Cette observation confirme l'entrée du dipeptide dans la souche et son clivage.

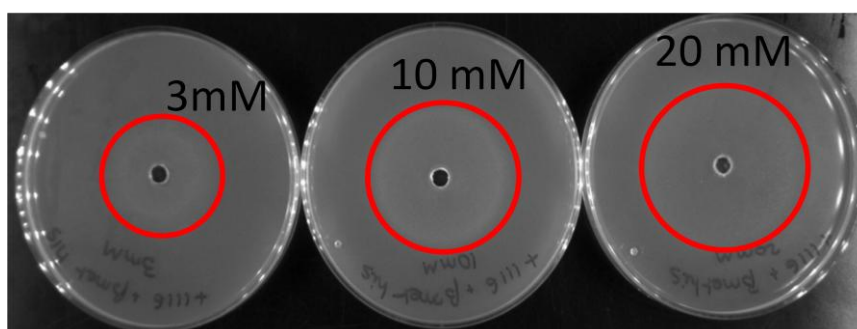


Figure 70 : Test d'incorporation du dipeptide L- β^3 -Met-His dans une souche d'*E. coli* Δ His

Le test a été fait à trois concentrations différentes de dipeptides ; 3mM, 10mM et 20mM.

Nous avons ensuite testé l'effet du peptide sur une souche prototrophe (MG1655). On observe cette fois un léger halo de toxicité (30mm). Ainsi, nous disposons maintenant d'un moyen de faire pénétrer la β -Met dans le cytoplasme d'*E. coli* et de démontrer sa toxicité.

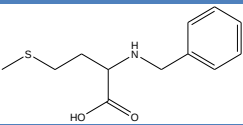
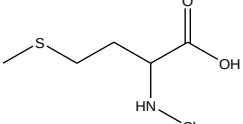
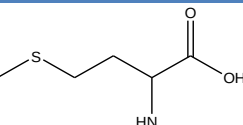
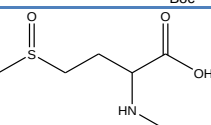
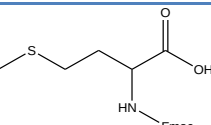
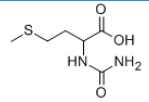
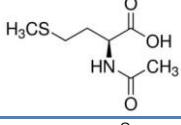
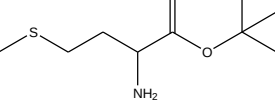
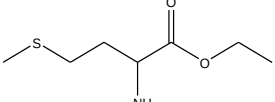
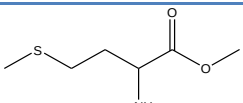
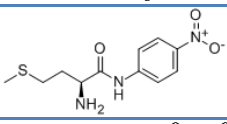
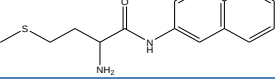
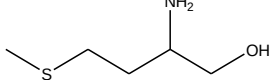
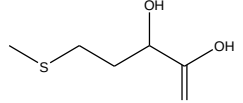
IV Amélioration de l'incorporation d'acides aminés analogues de la méthionine *in vivo*

Compte tenu de ces résultats, nous avons supposé que pour mettre en place un système susceptible d'aboutir à l'incorporation de β -Met *in vivo*, l'équilibre des concentrations d' α -Met et β -Met devait être ajusté avec soin. En effet, une concentration d' α -Met trop importante aboutira à une compétition trop forte en inhibant l'aminacylation, déjà peu efficace, de la β -Met. A l'inverse, s'il n'y a pas suffisamment d' α -Met, l'incorporation de β -Met dans les peptides risque d'être trop toxique. Nous avons donc décidé de trouver un mauvais précurseur de l' α -Met afin de limiter la concentration intracellulaire de l'acide aminé. Un mauvais précurseur se définit comme un composé qui peut être métabolisé mais dont la vitesse de métabolisation limite la croissance de la souche.

IV. 1 Recherche d'une mauvaise source de méthionine

De nombreux composés susceptibles d'être métabolisés par *E. coli* et transformés en méthionine ont été testés. Nous avons choisi deux catégories de précurseurs, ceux qui ont un groupement protecteur sur un des groupements caractéristiques de la méthionine et ceux qui possèdent un groupement chimique légèrement différent. La liste des composés testés ainsi que leur structure est présentée dans le tableau 15. Ces précurseurs ont été testés sur une souche auxotrophe pour la méthionine. Pour cela, on dépose sur une boîte une couche homogène non saturée de bactéries puis on dépose dans un puits un volume (50 μ L) de solution à concentration connue de précurseur. On obtient ainsi un gradient de concentration qui permet d'évaluer la capacité de la souche à utiliser le composé. Dans certains cas, nous avons dû utiliser un tampon contenant du diméthyle formamide (DMF, 10 ou 30 %) pour parvenir à une solubilité suffisante. Nous avons vérifié que 10 % ou 30 % de DMF n'inhibait pas la croissance.

Tableau 15 : Liste des précurseurs potentiels de la méthionine testés.

Précurseur	Structure
N-Benzyl-D-méthionine méthyl ester HCl	
N-Benzyl-L-méthionine méthyl ester HCl	
N-Benzyloxycarbonyl-D-méthionine	
N-Benzyloxycarbonyl-L-méthionine	
Boc-D-Met-OH	
Boc-L-Met-OH	
fmoc-met(o)-OH	
Fmoc-D-Met-OH	
Fmoc-Met-OH	
4-méthylsulfanyl-2-ureido-butyrac acid	
N-Acétyl-D-méthionine	
N-Acétyl-L-méthionine	
L-Méthionine tert-butyl ester HCl	
L-Méthionine éthyl ester HCl	
D-Méthionine éthyl ester HCl	
L-Méthionine méthyl ester HCl	
D-Méthionine méthyl ester HCl	
L-Méthionine p-nitroanilide	
L-Méthionine b-naphthylamide	
L-méthioninol	
D-méthioninol	
Desmeninol	

Les composés N-Benzyl-L-méthionine méthyl ester HCl, N-Benzylloxycarbonyl-L/D-méthionine, fmoc-met(o)-OH, Fmoc-L/D-Met-OH et le L/D-méthioninol n'ont pas permis d'observer la croissance d'une souche lorsqu'ils ont été utilisés à la concentration de 100mM ou même à saturation. Les composés N-Acétyl-L-méthionine, L/D Méthionine éthyl ester HCl, L/D-Méthionine méthyl ester HCl et Désmeninol ont permis d'observer la croissance d'une souche lorsqu'ils ont été utilisés à la concentration de 10mM. Cette croissance était équivalente à celle obtenue avec la méthionine. Ces précurseurs sont donc trop facilement convertis en méthionine.

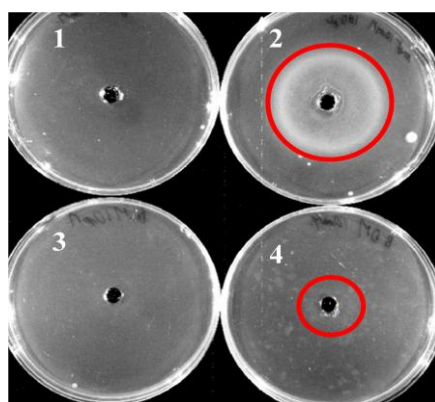


Figure 71 : Boite de culture avec 100 μ l de précurseur de la méthionine

1 : $MgCl_2$ 10mM: 2 : Met 10mM 30% DMF 3 : Boc-L- méthionine 10mM 30% DMF 4 : Boc-D- méthionine 10mM 30% DMF. A 10mM Boc-L/D-Met un halo ténu de croissance est présent même s'il est difficilement visible sur la photographie.

Les composés qui sont de mauvaises sources de méthionine conduisent à des petits halos de croissances ou des disques ténu de croissances (Figure 71). Les meilleurs candidats sont répertoriés dans le tableau 16.

Tableau 16 : Liste des précurseurs qui ne sont pas de bonnes sources de méthionine

L'ordre indique le classement de la meilleure mauvaise source de méthionine à la moins bonne.

Composé	Type de croissance observé et condition	ordre
N-Acétyl-D-méthionine	Disque ténu de croissance à 100mM	1
4-méthylsulfanyl-2-ureido-butyric acid	Disque ténu de croissance à 10mM, 10%DMF	2
Boc-L-Met-OH	Disque ténu de croissance à 10mM, 10%DMF	3
Boc-D-Met-OH		
N-Benzyl-D-méthionine méthyl ester HCl	Disque ténu de croissance à 10mM et halo de croissance avec un halo qui pousse moins à 100 mM	4
L-Méthionine p-nitroanilide	Petit halo intense (10mM 30%DMF)	5
L-Méthionine b-naphthylamide		6
L-Méthionine tert-butyl ester HCl	Halo moins intense et 2 fois plus petit que celui obtenu avec la met à la même concentration (10mM)	7

Au final, nous avons choisi les composés L et D Boc car ils sont un bon compromis entre efficacité et disponibilité commerciale.

IV. 2 Test du composés Boc-D-Met comme mauvaise source de méthionine

Valérie Pezo a réalisé des suivis de croissance en milieu liquide à plusieurs concentrations de précurseur. Comme on peut le voir sur la figure suivante, une concentration optimale de Boc-D-Met 3mM a été déterminés.

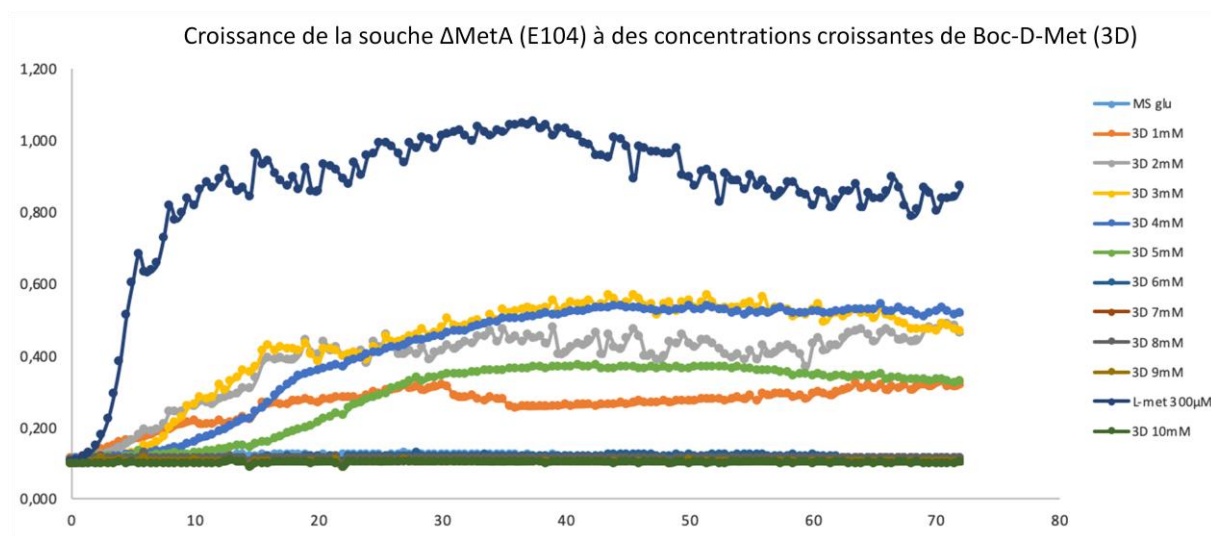


Figure 72 : Test de croissance d'une souche Δ met A en présence de Boc-D-Met (3D)

MS glu : MS glucose ajouté à la place de la méthionine ou du précurseur.

On remarquera que la DO maximale atteinte augmente en fonction de la concentration de Boc-D-Met jusqu'à atteindre un palier (figure 72).

Une fois cette concentration optimale déterminée, il a été nécessaire de vérifier si cette mauvaise source de méthionine pouvait en effet améliorer l'incorporation d'un AANC analogue de la méthionine. Une expérience dans des boîtes de milieu minimum contenant une concentration optimale de Boc-D-Met de 3mM a été réalisée en plaçant de la norleucine dans le puits. La norleucine est connue pour pouvoir s'incorporer à la place de la méthionine et être toxique. En présence de Boc-Met le halo de toxicité observé est plus grand que celui observé en présence de méthionine. Le Boc-met permet donc d'améliorer l'incorporation d'AANC analogue de la méthionine.

Une expérience identique, mais en mettant cette fois le dipeptide β -Met-His dans le puits est en cours de réalisation. Elle nous permettra, nous l'espérons, de trouver des conditions

dans lesquelles la β -Met peut être incorporée dans les protéines sans toxicité trop importante.

V Conclusion et perspectives

Nous ne sommes pas encore parvenus à mettre en place un système d'évolution dirigée fonctionnel, mais avons progressé dans cette direction. Nous avons en effet défini un système susceptible de permettre la sélection d'une incorporation accrue de β -Met. Il faut évidemment vérifier que, comme espéré, l'une au moins des mutations de la NDT pourra être compensée par l'introduction d'une β -Met. Une démarche claire et progressive pour y parvenir nécessite dans un premier temps la démonstration que la β -Met peut être incorporée, même très marginalement, dans les polypeptides d'*E. coli*. En préalable, il fallait mettre en place une méthode pour faire rentrer la β -Met dans *E. coli* et trouver une source médiocre d' α -Met. Nous sommes maintenant en mesure de rechercher des conditions de croissance d'une souche auxotrophe pour la méthionine, en présence de β -Met et de Boc-Met (L- ou D-), permettant l'incorporation partielle de β -Met. La MetRS sera éventuellement surproduite. Cette incorporation sera suivie par spectrométrie de masse à haute résolution. Ensuite, nous pourrions vérifier si l'une des mutations de la NTD est compensée dans ces conditions.

Discussion et perspectives

Un premier résultat de notre travail est la caractérisation de la reconnaissance de la β -Met par la MetRS d'*E. coli*. Les résultats issus de l'Université de Yale (Melo Czekster *et al.* , 2016) laissaient penser que l'enzyme pouvait utiliser la β^3 -Met avec une grande efficacité. Nous montrons en utilisant différentes méthodes que l'efficacité de la MetRS vis-à-vis de la β^3 -Met est en fait très modérée. Nous pensons que ces résultats sont au moins en partie liés à une contamination des β -Met commerciales par de l' α -Met. Les effets de telles contaminations, même infimes (moins que 1/1000 dans notre cas), peuvent être très importants, que l'on utilise l'échange isotopique ATP-PPI, la fluorescence ou les tests d'aminacylation. Cet effet est connu depuis longtemps et a conduit à de nombreuses interprétations erronées. Il a été soigneusement considéré par Alan Fersht dans le cas de l'activation d'analogues de la méthionine, et avait été résolu en oxydant la méthionine contaminante au Nickel de Raney (Fersht and Dingwall, 1979b). Toutefois, cette méthode ne peut pas s'appliquer à notre situation, puisque les chaînes latérales des α - et β -Met sont identiques. Par contre, nous avons pu prouver que la méthionine- γ -lyase est capable d'agir sélectivement sur la méthionine α , ce qui nous a permis de fortement diminuer le taux de contamination. Ceci s'est avéré suffisant pour observer la formation d'adénylate en fluorescence ainsi que l'aminacylation de l'ARNt^{Met}. La spectrométrie de masse à haute résolution nous a en particulier permis de prouver que les β -Met-adénylate et β -Met-ARNt étaient bien formés. Par contre, il ne nous a pas été possible de réduire de manière contrôlable la contamination en α -Met au point de pouvoir mesurer de manière certaine l'activité d'échange isotopique ATP-PPI. Au final, les efficacités de catalyse de l'activation et de l'aminacylation de la β -Met sont respectivement réduites de 4 et 3 ordres de grandeurs par rapport à l' α -Met, essentiellement à cause de l'étape catalytique. La structure à haute résolution du complexe MetRS: β -Met montre effectivement que le carboxylate du substrat n'est pas correctement positionné. Il est donc crucial de considérer avec le plus grand soin les questions liées à la contamination des réactifs.

Ce résultat rend d'autant plus pertinente la démarche qui consiste à tenter d'améliorer la spécificité de la MetRS vis-à-vis de la β -Met. L'intérêt de cette étude a par ailleurs été encore renforcé par des résultats parus au cours de ce travail de thèse. Il a en particulier été montré qu'il était possible, avec un système de traduction *in vitro* modifié, d'incorporer plusieurs acides aminés β à la suite (Katoh and Suga, 2018). De plus, cette étude montre que la méthionine fait partie des acides aminés β les mieux tolérés par la synthèse ribosomale. Nous avons tout d'abord étudié les interactions de l'enzyme avec divers analogues de l' α -Met afin de mieux comprendre le rôle de chaque groupement du substrat dans la liaison et dans la catalyse. A partir de ces résultats et de la structure du complexe MetRS: β -Met, nous avons réalisé une première série de mutations intuitives pour améliorer la catalyse. Aucun des variants de cette série n'a malheureusement permis d'améliorer l'activité, même si structuralement le variant A12S favorise une conformation *a priori* favorable de la β -Met.

Un modèle *in silico* a ensuite été utilisé pour générer des variants plus actifs envers la β -Met. Plusieurs dizaines de variants ainsi suggérés ont été produits et purifiés. Leur caractérisation fonctionnelle reste assez longue et complexe, mais une première série de mesures liées aux activités envers l' α -Met et la β -Met suggère que trois triples variants aux positions 13, 256 et 297 (CAC, MAC et LAC) montrent une spécificité accrue pour la β -Met. Nous avons démarré une étude plus approfondie de ces trois enzymes afin de déterminer si nous avons effectivement réussi à améliorer la spécificité. Si cela est bien le cas, une étude structurale de ces variants pourra être entreprise.

Un autre axe important de ce travail est la construction d'un système de criblage permettant d'améliorer l'incorporation de β -Met en utilisant l'évolution dirigée. Nous nous sommes heurtés à plusieurs difficultés. La première était de trouver une activité indicatrice adéquate, capable de répondre positivement à l'incorporation d'une β -Met. Nous pensons que le système de la NTD décrit ici reste prometteur, mais le test du système a pour l'instant buté sur le problème de la mise en évidence d'une incorporation, même faible, de la β -Met dans les protéines *d'E. coli*. Un obstacle a été la perméabilité de la bactérie à l'AANC, mais nous avons pu le résoudre en le fournissant sous la forme d'un dipep-

tide. Une autre difficulté était d'identifier une source de méthionine médiocre, permettant d'optimiser le rapport β -Met/ α -Met dans l'incorporation tout en évitant la toxicité. Nous pensons que cette difficulté est presque résolue, et comptons mettre en évidence l'incorporation de β -Met dans les protéines *d'E. coli* en utilisant la spectrométrie de masse. Il sera alors possible de déterminer si au moins un des variants de la NTD peut devenir actif en présence de β -Met. A ce stade, la souche destinée au criblage pourra être construite. Dans le gène de NTD muté adéquat, nous introduirons un codon ambre à la place du codon Met. La souche produira également un $\text{ARNt}_{\text{m}}^{\text{Met}}$ à anticodon ambre. Le système sera amorcé avec un plasmide portant le gène d'une MetRS mutée capable de reconnaître un $\text{ARNt}^{\text{ambre}}$ (Schmitt *et al.*, 1993), ainsi que des mutations améliorant la spécificité envers la β -Met, comme celles identifiées dans la partie 2. Cette méthode pourrait permettre de sélectionner des mutations dans la MetRS qui augmentent sa spécificité vis-à-vis de l'AANC. L'avantage de l'évolution dirigée *in vivo* est que la pression évolutive s'exercera sur tout le système de traduction et pas seulement sur la MetRS. On pourrait donc obtenir des mutations dans d'autres composants tels que le ribosome, EF-Tu, EF-P ou encore l'ARNt, voire dans les voies métaboliques des ARNt soufrés.

Matériels et Méthodes

I Matériels

I. 1 Liste des souches

Nom de la souche (<i>E. coli</i>)	Génotype	Source et phenotype
XL1 – blue	<i>endA1 gyrA96(nal^R) thi-1 recA1 relA1 Δlac glnV44 F'[::Tn10 proAB⁺ lac^R Δ(lacZ)M15] hsdR17(rk⁻ mk⁺)</i>	
Rosetta	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm</i> λDE3 pRARE (Cam ^R)	http://www.uni-frankfurt.de/fb15/mikro/euroscarf/data/
XA102	D(<i>lac-pro</i>) <i>gyrA rpoB metB argE</i> Δ <i>supE ara</i>	(Coulondre and Miller, 1977) Met - Pro - Lac - Rif ^R
E103	Δ <i>metA</i> :: <i>kan</i>	Auxotrophe méthionine dérivé de MG1655 (V. Pézo)
THY0	Δ <i>thyA</i> :: <i>erm</i> Δ <i>deo</i> Δ <i>udp</i> :: <i>kan</i>	Auxotrophe thymine dérivé de MG1655, (Marlière <i>et al.</i> , 2011)
MG1655	<i>rph-1</i>	
BLR(DE3)/FB810	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal lac ile dcm</i> Δ(<i>srl-recA</i>)::Tn10(tet ^R) λDE3	
EV1	ΔHis :: <i>kan</i>	Auxotrophe histidine dérivé de MG1655 (V. Pézo)

I. 2 Liste des plasmides

Gènes	Nom	Plasmide/source	Description du plasmide
M547 +	pBSM547 +	pBSM547+ (Mellot et al. , 1989; Schmitt et al. , 1993)	MetG 547 sous contrôle de son promoteur et du promoteur lac ColE1
M547+ histag	pBSM547 + histag	Ce travail	MetG 547 avec une étiquette histidine sous contrôle de son promoteur et du promoteur lac ColE1
-	pET15blpa	pET15blpa (Guillon et al. , 2005)	
M547 +	pET15 blpa M547	Ce travail	MetG 547 avec une étiquette histidine sous contrôle du promoteur lac ColE1
cat+ ntd+	PIC15 – ISSB	V. Pézo	NDT sous contrôle d'un promoteur dérivé du bacteriophage T5 ,Cm-res
Amp+ ntd+	PXEN28 – ISSB	V. Pézo	NDT sous contrôle d'un promoteur dérivé du bacteriophage T5 ,Amp-res

I. 3 Liste des milieux de culture et tampons

Milieu	Composition
2TY	Tryptone (bacto) 16g/L, extrait de levure 10g/L, NaCl 5g/L
Luria-Bertani (LB)	Tryptone (bacto) 10g/L, extrait de levure 5g/L, NaCl 5g/L
Terrific Broth Auto Inductible (TBAI)	Tryptone (bacto) 12g/L, extrait de levure 24g/L, MgSO ₄ 0,15g/L, (NH ₄) ₂ SO ₄ 3,3g/L, KH ₂ PO ₄ 6,5g/L, Na ₂ HPO ₄ 7,1g/L, glucose 0,5g/L, α-Lactose 2,0g/L
Muller Hinton Agar	Infusion de bœuf solide 2g/l, Amidon 1,5g/l, hydrolysate de caséine 17,5g/l, Agar 17g/l
M9	KH ₂ PO ₄ 6,5g/L, Na ₂ HPO ₄ 7,1g/L, NaCl, Glucose, Thiamine

Etapas utilisé	Nom du tampon	Composition
Dosage	Tampon de dosage	20mM Tris, HCl (pH 7,6), 2mM MgCl ₂ , 10mM 2-mercaptoethanol, 0,1mM EDTA
Mesure de la variation de fluorescence	Tampon de fluorescence	20mM Tris, HCl (pH 7,6), 2mM MgCl ₂ , 10mM 2-mercaptoethanol, 0,1mM EDTA
Cassage de bactéries	Tampon de cassage	10 mM HEPES, 3 mM 2-mercaptoethanol et 500mM NaCl
Tampon d'elution pour la colonne d'affinité Talon	TP A	10 mM HEPES-OH pH 7, 3mM 2-mercaptoethanol et 500mM NaCl
	TP B	10 mM HEPES-OH pH 7, 3mM 2-mercaptoethanol, 500mM NaCl et 250mM imidazole
Tampon pour la colonne Echangeuse d'ion Q-sepharose Hiload	TP A	10 mM HEPES-OH pH 7, 10mM 2-mercaptoethanol
	TP B	10 mM HEPES-OH pH 7, 10mM 2-mercaptoethanol et 1M NaCl
Stockage des protéines à -20°C	Tampon glycerol	10 mM HEPES-OH pH 7, et 55% glycérol

II Méthodes

II. 1 Production et purification de protéines

II. 1. 1 Clonage

Le gène codant pour la MetRS monomérique M547 de pBSM547+ a été cloné dans le plasmide pET15bla pour ajouter une étiquette histidine à la MetRS.

L'amplification du gène est effectuée dans une solution de 25 µL contenant 10 ng de plasmide, 50 pmole de chaque amorce nucléotidique, 200 µM de chaque désoxyribonucléotides, 2 U d'ADN polymérase thermostable Taq DNA polymérase (Roche) ainsi que 0,4 U d'ADN polymérase Cloned Pfu (Stratagen). La PCR consiste en 30 cycles (15 s à 95°C, 30 s à 50°C et 1 min à 72 °C) précédés d'une phase de dénaturation à 95°C (2min). La PCR s'achève avec une phase de 5 min à 72°C. L'ADN est ensuite digéré avec des enzymes de

restriction adéquats (Not 1 et FspB1) avant précipitation à l'éthanol et redissolution dans 25 μ L d'eau.

La ligation a été réalisée dans une solution de 25 μ L contenant 10 ng de vecteur, 0,5 mM d'ATP, 2,5 U d'ADN Ligase. Ce volume inclut 2 μ L de la solution de fragment ainsi que 2,5 μ L de tampon 10x de la ligase. La réaction est faite pendant 1 h à température ambiante. L'enzyme est ensuite inactivée en incubant 10 min à 70°C. Finalement, les vecteurs reconstitués non désirés sont linéarisés par digestion avec SmaI.

1 μ L du mélange réactionnel est utilisé pour transformer des bactéries compétentes de la souche XL1 Blue. Après l'électroporation, les bactéries sont diluées dans 3 ml de HTOP avant d'être étalées sur boîte de LB Agar Amp. Après avoir incubé la boîte durant une nuit à 37°C, huit colonies sont repiquées et incubées à 37°C, pour une nuit, à deux reprises pour s'assurer de la clonalité. Les plasmides sont vérifiés par analyse de restriction puis par séquençage.

II. 1. 2 Mutagénèse

Les variants de la MetRS ont été préparés grâce à la méthode QuickChange™ Sited directed Mutagenesis (Stratagene).

Le protocole utilisé consiste à préparer 100 μ L contenant 50 ng de plasmide, 15 pmole de chaque amorce, 200 μ M de chaque désoxynucléotide, 2,5 U de polymérase PfuTurbo (Polymerase *Pyrococcus furiosus*) dans un tampon fourni. La réaction commence par 5 min à 95°C avant 12 cycles comportant 30 s à 98°C puis 30s à 50°C et 2min à 72°C. L'ADN est ensuite digéré par l'enzyme de restriction Dpn1 avant transformation dans la souche XL1 Blue. Des clones sont alors isolés puis l'ADN plasmidique isolé et séquencé.

La séquence complète du gène muté a été systématiquement vérifiée.

II. 2. 1 Production

La MetRS et ses variants ont été produits et purifiés après expression des plasmides recombinants dans la souche BLR(DE3) en milieu auto induit TBAI (50 μ g/mL ampicillin) à 37°C pendant 20 h sous agitation. La culture est ensuite centrifugée à 7000 g. Pour isoler la MetRS le culot est ensuite repris dans le tampon de cassage.

II. 2. 2 Purification

Les bactéries sont désintégrées par sonication pendant 5 min. Les protéines sont purifiées sur des systèmes chromatographiques. Le lysat clarifié est injecté sur une colonne Talon (Clontech) puis la résine est lavée à 10 mM imidazole avant élution à 125 mM imidazole. La seconde partie de la purification consiste en une purification sur une colonne échangeuse d'ions Q-Hiload (GE-Healthcare). La colonne est équilibrée à 5mM NaCl avant d'injecter la protéine puis de l'éluer par un gradient allant de 5mM NaCl à 500mM NaCl. Pour 250 ml de cultures des colonnes de 5 à 10 ml sont préparées pour chaque variant afin d'éviter la contamination d'un variant à l'autre. La quantité de protéine obtenue varie entre 20 et 25 mg. Les protéines purifiées sont stockées à -20°C dans le tampon glycérolé.

II. 2 Enzymologie

Pour toutes les expériences suivantes les substrats ont été dissouts dans les tampons adéquats et si nécessaire ajustés à pH 7,5

II. 2. 1 Digestion de la β -méthionine à la γ -lyase (MGL)

La β -méthionine (achetée chez Fluorochem Ltd. (UK)) est dissoute dans le Tampon standard puis le pH est ajusté à 7,5.

Pour digérer la méthionine contenue dans la β -méthionine, celle-ci est traitée à la MGL. Le traitement se fait dans 100 μ L à 160 mM β -méthionine (pH 7,5), 150 μ M de pyridoxal phosphate (PLP), 20 mM de Tris-HCl pH 7,6, 0,1 mM EDTA, 10mM 2-mercaptoethanol et 1 μ L de γ -lyase de *Pseudomonas putida* à 1 U/ μ L. La digestion se fait à 37°C pour 14 h avant incubation à 95 °C pendant 2 min pour arrêter la réaction.

II. 2 Caractérisation de la MetRS

La concentration d'enzyme est calculée en utilisant le coefficient d'absorption massique à 280 nm de la MetRS (Cassio and Waller, 1971). Une unité d'absorbance à 280 nm donne une concentration de 0,58 mg/ml

Quel que soit le tampon, si l'enzyme est diluée en dessous de 500 nM, le tampon contient en plus 200 μ g/ml d'albumine de sérum bovin.

II. 2. 1 Echange isotopique ATP-PPI

Les paramètres cinétiques des enzymes purifiées ont été déterminés pour la réaction d'échange isotopique [^{32}P]PPI-ATP et pour la réaction d'aminacylation de l'ARN_{tr}^{Met}

La vitesse de la réaction d'échange [^{32}P]PPI-ATP est mesurée à 25°C dans un volume final de 100 μL de tampon standard (20mM Tris-HCl (pH 7,6), 7 mM Mg^{2+} , 10 mM 2-mercaptoethanol, 0,1 mM EDTA), dans lequel est ajouté 2 mM [^{32}P]PPI, 2 mM ATP et 2 mM méthionine (Blanquet *et al.* , 1974; Mellot *et al.* , 1989). Pour la mesure des K_m , les concentrations de méthionine et de β -Met varient respectivement de 0,07 mM à 2 mM et de 0,125 à 16 mM. La réaction commence par l'ajout de l'enzyme. Sa concentration varie en fonction de son activité.

II. 2. 1. 1 Test d'aminacylation

La vitesse d'aminacylation a été mesurée de deux façons. La première permet de mesurer les vitesses d'aminacylation grâce à des acides aminés marqués au ^{14}C . La seconde permet de mesurer la vitesse quand les AA ^{14}C ne sont pas disponibles commercialement. Cette méthode nécessite de l'ARNt marqué en 3' au ^{32}P

II. 2. 1. 2 Mesure de la vitesse de la réaction avec de la L-[^{14}C] méthionine

Les vitesses de catalyse de la réaction d'aminacylation sont mesurées dans le tampon standard contenant en plus 150 mM KCl, 2 mM ATP, 25 μM de L-[^{14}C] méthionine et 8 μM d'ARN_{tr}^{met}.

III. 2. 2 Aminacylation de l'ARNt marqué au ^{32}P

Le marquage de l'ARNt par du ^{32}P se fait en échangeant l'adénosine en 3' de l'ARNt grâce à de l'ARNt nucleotidyltransferase d'*E. coli* purifiée au laboratoire grâce à un plasmide fourni par K.Tomita (Université de Tokyo). La méthode utilisée est celle décrite par Ledoux et al 2008 (Ledoux and Uhlenbeck, 2008). La réaction se fait dans 50 μL de solution à 37°C. Cette solution contient 1 μM d'ARNt, 50 μM de pyrophosphate de sodium, 10mM de MgCl_2 , 5 mM de glycine pH9, 0,3 μM d'ATP [α - ^{32}P] et 0,02 μM de nucleotidyltransferase. La réaction démarre après l'ajout de l'enzyme. Après 5 minutes de réaction, 5 μL d'une solution contenant 1 μM de CTP et 0,1 U de pyrophosphatase sont ajoutés. La solution est

incubée 2 min supplémentaires à 37°C. La solution est ensuite extraite avec du phénol puis de l'éther. La phase aqueuse est ensuite traitée sur une colonne de tamis moléculaire Micro bio-Spin P-6 pour retirer les restes de phénol et d'ATP. La fraction récupérée est ensuite précipitée à l'éthanol puis lavée à l'éthanol 100%. Le précipité est ensuite repris dans le tampon utilisé pour la réaction d'aminocyclisation.

L'aminocyclisation est faite à 37°C dans 20 µL le tampon standard contenant 7 mM de MgCl₂, 150

mM de KCl, 2 mM d'ATP et 2 mM d'ARNt [³²P]. Pour les mesures de vitesse d'aminocyclisation de la Met et de la β-Met, les acides aminés sont à une concentration de 2 mM et la MetRS à des concentrations de respectivement 1 nM et 20 nM.

Pour chaque mesure, 3 µL du mélange réactionnel sont prélevés et mélangés à 3 µL d'une solution de Nucléase S1 à 0,1 U dans 0,3 M d'acétate de sodium. La digestion est effectuée 10 min à 37°C. 1µL de cette solution est ensuite déposé sur une plaque CCM de PEI cellulase (Bio-Rad). La migration s'effectue avec une phase mobile contenant 85/10/5 (H₂O/1 M NH₄Cl/Acide acétique). La plaque est ensuite séchée puis analysé au phosphoimager Typhoon.

II. 3. Etude par fluorescence

II. 3. 1 Fluorescence à l'équilibre

Les mesures de la fluorescence intrinsèque de l'enzyme ont été faites sur un spectrofluorimètre JASCO. L'excitation s'effectue à 295 nm et l'émission est enregistrée à 340 nm.

Pour mesurer la constante de dissociation, une dilution enzymatique de 800 µL est préparée dans le tampon standard auquel on a ajouté 2 mM de MgCl₂. Cette solution est placée dans une cuve de quartz, puis les courbes de saturation sont obtenues par l'ajout d'aliqots de 5 µL contenant des concentrations croissantes de substrat. Les concentrations final varient ainsi de 0,008 à 1,4 mM pour la Met et de 0,06 mM à 8 mM pour la β-Met.

Dans le cas des mesures pour le couplage synergique, le tampon contient 15 mM d'adénosine et 2 mM de PPi et les gammes de concentrations du substrat sont décalées. Il en va de même pour la mesure des K_D des analogues de la méthionine.

Pour un complexe enzyme : substrat à l'équilibre, l'expression de la constante de dissociation est $\frac{[E][S]}{[ES]}$.

Dans le cas où le K_D est grand devant la concentration d'enzyme ($0,8\mu\text{M}$), l'expression de la fluorescence peut être décrite de la manière suivante : $F = F_0 + \frac{(F_\infty - F_0)[S_t]}{K_D + [S_t]}$

F_0 et F_∞ sont les fluorescences à concentration nulle et à concentration infinie de substrat $[S_t]$ et $[E_t]$ correspondent aux concentrations totales de substrat et d'enzyme.

Dans les autres cas l'équation de la fluorescence peut être écrite de la façon suivante :

$$F = F_0 - \frac{(F_\infty - F_0)}{[E_t]} \left([E_t] + K_D + [S_t] - \sqrt{([E_t] + K_D - [S_t])^2 + 4K_D[S_t]} \right)$$

II. 3. 2. Fluorescence à l'état pré-stationnaire

La formation ou la pyrophosphorolyse du méthionyl-adénylate s'accompagnent d'une variation de la fluorescence. L'enregistrement des variations de fluorescence permet, grâce à un appareil de cinétique rapide, de mesurer l'ensemble des paramètres cinétiques et d'équilibre de la réaction d'interconversion (k_f , k_b , K_{ATP}^{met} , $K_{PP}^{met-AMP}$) (Hyafil *et al.*, 1976).

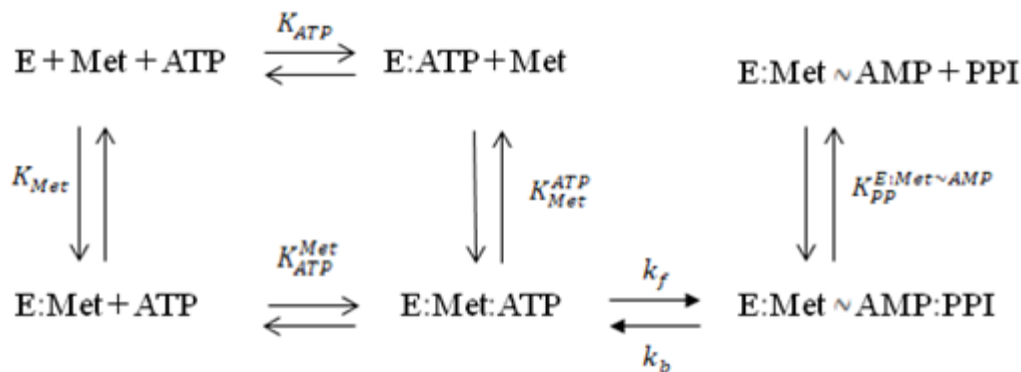


Figure 73 : Schéma des états pré-stationnaires de la réaction d'activation de la méthionine

La constante de vitesse k_f est mesurée en suivant la formation du méthionyl- adénylate en présence de différentes concentrations de méthionine, d'une concentration saturante d'ATP et d'une concentration faible de PPI, mais grande devant la concentration d'enzyme. Ces expériences sont réalisées dans le tampon standard (20 mM Tris-HCl (pH 7,6), 10 mM 2-mercaptoethanol, 0,1 mM EDTA), avec une concentration en magnésium libre de 2 mM. Le mélange réactionnel est thermostaté à 25°C. Les constantes de temps des relaxations mono-exponentielles sont calculées avec Origin®. Les constantes sont ensuite

déterminées par régression non-linéaire. L'expression de la constante de vitesse peut être calculée à partir du schéma réactionnel. On obtient alors l'équation suivante :

$$k = k_f G^{-1} + k_b H^{-1}$$

Avec

$$G = 1 + \left(\frac{K_{ATP}^{met}}{[ATP]} \right) \left(1 + \left(\frac{K_{Met}}{[Met]} \right) \left(1 + \frac{[ATP]}{K_{ATP}} \right) \right)$$

et

$$H = 1 + \frac{K_{PP}^{Met-AMP}}{[PP]}$$

II. 4 Mise en évidence de la catalyse par LC-MS et LC-MS/MS

II. 4. 1 Activation de l'acide aminé

La réaction est effectuée dans 0,8 ml d'une solution contenant 20 mM de Tris-HCl, 0,1 mM EDTA, 10m M 2-mercaptoethanol, 2 mM ATP, 2 mM MgCl₂, 1 mM de β-met et 0,5 μM MetRS. La réaction est suivie en fluorescence. Après 10 min de réaction, 100 μL sont analysés en LC-MS et LC-MS/MS après ajout de 10% d'acide formique.

II. 4. 2 Aminoacylation de l'ARNt^{Met}

L'ARNt est aminoacylé lors d'une incubation de 30 min à 25°C d'une solution (1 mL) contenant 20 μM d'ARNt, 2 mM d'ATP, 7 mM de MgCl₂, 150 mM de KCl₂, 20 mM de Tris-HCl, 0,1 mM EDTA, 10 mM 2-mercaptoethanol, 1 μM MetRS et 0,4 mM méthionine ou 1 mM β-méthionine. 111 μL d'une solution de NaAc 3M sont ajoutés au mélange réactionnel avant de précipiter l'ARNt à l'éthanol. La suspension ainsi obtenue est séparée dans 4 tubes, centrifugée 10 min à 15300 rpm. Le surnageant est ensuite retiré. Les culots sont lavés 2 fois avec 500 μL d'éthanol 100%.

Un des culots est repris dans 100 μL une solution aqueuse à 5mM NaAc pH 5,5. Cette solution est centrifugée 10 min à 15300 rpm. Le surnageant est ensuite refroidi à 0°C avant d'ajouter 100 μL de DMSO puis 20 μL d'acide acétique et 20 μL d'anhydride acétique. Le mélange réactionnel est gardé à 0°C pendant 30 min. La réaction est ensuite arrêtée en ajoutant 24 μL NaAc pH 4,8 et 600 μL d'éthanol 100%. La suspension est centrifugée 10min à 15300rpm, le surnageant est ensuite retiré puis le culot est lavé à l'éthanol 100%.

Le N-acetylaminoacyl-ARNt est ensuite repris dans 90 μ L d'une solution contenant 80 μ L d'eau et 10 μ L d'une solution 30 mM de NaAc et 10 mM de ZnAc. La digestion est démarrée par l'ajout de 10 μ L de nucléase S1 à 1 U/ μ L. La réaction est arrêtée après 20 minutes à 37°C en plaçant le tube dans la glace. La digestion à la nucléase S1 est également effectuée sur un échantillon non acétylé. Ces deux solutions sont acidifiées en ajoutant 10% d'acide formique puis analysées en LC-MS et LC-MS/MS. Une dilution d'un facteur mille est généralement nécessaire avant analyse.

II. 5 Cristallisation de la MetRS et acquisition des données cristallographiques

II. 5. 1 Préparation des cristaux de MetRS

La cristallisation de la MetRS et de ses variants a été faite dans une seule condition (1,08 M citrate d'ammonium 10 mM phosphate de potassium et 10 mM 2-mercaptoethanol). La concentration de substrat dans les gouttes est variable en fonction du substrat. (1mM Met, 10, 20, 50 mM β -Met et 20 et 50 mM analogues). Les cristaux ont été cultivés en gouttes assises ou en gouttes suspendues. Le micro ensemencement a été fait à partir de cristaux de MetRS WT. Le temps de croissance après ensemencement est variable, entre deux semaines et deux mois en fonction du substrat ou de la mutation. Les cristaux ont été plongés dans la solution de cristallisation supplémentée de 25% de glycérol avant d'être refroidis dans l'azote liquide.

II. 5. 2 Principe de la cristallographie par diffraction des rayons X

Lorsqu'un rayonnement électromagnétique, et en particulier les rayons X, traverse la matière, il interagit avec les particules rencontrées. Deux phénomènes résultent de cette interaction :

- l'absorption
- la diffusion

Pour un rayon X et plus généralement pour les ondes électromagnétiques, la diffusion est essentiellement due aux interactions avec les électrons. Lorsque la matière traversée présente une organisation spatiale périodique, comme les cristaux, un phénomène de diffrac-

tion se produit. Les ondes ainsi diffusées interfèrent avec le réseau cristallin, ce qui entraîne que les rayons X diffractent uniquement dans certaines directions privilégiées. L'onde diffractée sera non nulle uniquement pour la famille de plans réticulaires (hkl, ou h, k et l sont les indices de Miller de la famille de plans) satisfaisant la loi de Bragg :

$$2d \sin \theta = \lambda n$$

Avec d : la distance entre deux plans (hkl), λ : la longueur d'onde, θ : l'angle entre le rayonnement incident et le plan réticulaire (hkl) et n : un entier positif

A chaque plan réticulaire correspond donc un point dans l'espace de diffraction, ce qui forme un réseau réciproque. Le facteur F_{hkl} caractérise l'onde diffractée en amplitude et en phase par les plans réticulaires (hkl).

$$F_{hkl} = \iiint_v \rho(xyz) e^{2i\pi(hx+ky+lz)} dx dy dz$$

Avec v : le volume de la maille

Et ρ : la densité électronique

Dans le cas de cristaux moléculaires, ce facteur peut s'écrire de la façon suivante :

$$F_{hkl} = \sum_j f_j e^{2i\pi(hx_j+ky_j+lz_j)}$$

Ce nombre complexe peut aussi s'écrire sous forme d'un module et d'une phase où les j sont les atomes de la maille:

$$F_{hkl} = |F_{hkl}| e^{i\varphi(hkl)}$$

L'intensité diffractée I_{hkl} est proportionnelle au carré du module de F_{hkl} et s'obtient directement à partir des clichés de diffraction. Par contre les phases $\varphi(hkl)$ ne peuvent pas être mesurées. On en obtient des valeurs approximatives en utilisant le modèle de la MetRS WT (voir plus loin)

La densité électronique en chaque point du cristal s'obtient alors par transformée de Fourier inverse, fonction qui prend la forme suivante :

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} F_{hkl} e^{-2i\pi(hx+ky+lz)}$$

II. 5. 3 Acquisition et traitement des données de diffraction

La densité électronique d'un cristal est l'image, par transformé de Fourier, des facteurs de structure F_{hkl} qui caractérisent l'espace réciproque. Il est donc nécessaire de cartographier l'ensemble de cet espace réciproque pour pouvoir reconstruire la densité électronique dans l'espace réel. La méthode couramment utilisée en radiocristallographie est la méthode dite de rotation du cristal qui consiste à faire tourner le cristal et à prendre un cliché de diffraction tous les $0,1^\circ$. Un jeu complet de données est obtenu en faisant tourner le cristal suivant un angle total qui dépend de la symétrie. Les détecteurs modernes à détection directe permettent d'obtenir un jeu de données redondant après une rotation de 360° .

Lorsqu'un jeu de données complet est enregistré, celui-ci est traité de manière à convertir les clichés de diffraction en intensités intégrées. Ce traitement se fait en trois étapes :

- La détermination des paramètres de maille, de la mosaïcité et de l'orientation du cristal qui est faite à partir d'un cliché de diffraction. Ces différents paramètres ainsi que ceux relatifs à l'acquisition des données sont ensuite affinés sur l'ensemble des clichés de diffraction.

- A partir de ces paramètres, la position des réflexions est alors prédite. Cette prédiction permet de mesurer l'intensité de chaque tache. Cette méthode permet d'améliorer la qualité des mesures, surtout dans le cas où les réflexions sont de faible intensité.

- L'intensité de chaque réflexion indépendante est calculée à partir des mesures de toutes les réflexions équivalentes enregistrées sur l'ensemble des clichés. Des constantes d'échelle sont calculées pour chaque cliché de manière à homogénéiser les données.

La qualité du jeu de données est donnée par diverses valeurs statistiques dont, le facteur R_{sym} :

$$R_{sym} = \frac{\sum_{hkl} \sum_i |\langle I_{hkl} \rangle - I_{hkl}^i|}{\sum_{hkl} \sum_i I_{i,hkl}}$$

Où I_{hkl} : la $i^{ème}$ mesure de la réflexion hkl

Et $\langle I_{hkl} \rangle$: la valeur moyenne des N réflexions équivalentes

Nous avons également utilisé le coefficient de corrélation entre deux moitiés de jeu de données ($CC_{1/2}$)

II. 5. 4 Résolution de la structure

La résolution des structures a été effectuée en utilisant la méthode de remplacement moléculaire qui suppose que l'on connaît la structure d'une molécule proche.

La structure de la MetRS, sans eau a été utilisée comme modèle (PDB 3H9C, Schmitt *et al.* , 2009) pour le remplacement moléculaire. On utilise les coordonnées et les phases obtenues pour démarrer l'affinement.

Les coordonnées et les facteurs B ont été affinés par plusieurs cycles d'ajustement manuel en utilisant le programme COOT (Emsley *et al.* , 2010) et de minimisation avec le logiciel PHENIX (Adams *et al.* , 2010).

Tous les cristaux appartiennent au même groupe d'espace. Pour les cristaux qui diffractaient à mieux que 1,5Å de résolution, les facteurs B ont été affinés de manière anisotrope pour la protéine. Lorsque la résolution était inférieure à 1,3 Å, ils ont été affinés de manière anisotrope pour la protéine et les molécules d'eau.

III Synthèse peptidique

III. 1 Synthèse peptidique

La synthèse peptidique est faite sur résine WANG. 500mg de résine sont lavées 3 fois au N,N-Diméthylformamide (DMF) (20mL) puis 3 fois au dichloromethane (DCM) (20 mL) et de nouveau 3 fois au DMF (20mL). Les groupements Fmoc sont déprotégés avec une solution de pipéridine/DMF (2/8) (10 mL/10 min puis 10 ml/10 min). Un test Kaiser (voir ci-dessous) est effectué afin de vérifier que la réaction est totale. Après cette déprotection, 2 équivalents (270 mg) de Fmoc-β méthionine sont ajoutés en présence de 2 équivalents de N,N'-dicyclohexylcarbodiimide (144 mg) et 2 équivalents de HOAt (96mg). De nouveau, un test Kaiser est effectué afin de vérifier que la réaction est totale. Avant de décrocher le dipeptide de la résine, les groupements Fmoc sont enlevés avec une solution de pipéridine/DMF (2/8) (10 mL/30 min puis 10 mL/30 min). Les billes de résines sont ensuite lavées au DMF puis au DCM et un dernier test Kaiser est effectué.

Lorsque le dipeptide est finalisé et déprotégé, les billes sont lavées deux fois au DCM avant d'être lavées de nouveau au DMF puis au DCM. Ensuite, la résine est séchée sous flux

d'air. Les billes sont ensuite agitées à température ambiante dans 10ml d'une solution de TFA/H₂O/phénol/TIS (88/5/5/2) pendant 4h à 0°C. La suspension est alors filtrée puis le TFA est évaporé jusqu'à ce qu'il ne reste que 1 ou 2 mL. 30 mL d'éther sont ensuite ajoutés. Le précipité est lavé à l'éther. La phase éther est ensuite extraite à l'eau. Le précipité est ensuite solubilisé dans l'eau. Cette phase aqueuse est ensuite lavée à l'éther. Les deux phases aqueuses sont ensuite diminuées jusqu'à un minimum d'eau. Une précipitation à l'éthanol est ensuite effectuée par ajout de 2,5 fois le volume en éthanol 100%. Après centrifugation, le surnageant est récupéré puis lyophilisé.

III. 2 Test Kaiser

Quelques billes de résine sont prélevées du milieu réactionnel à l'aide d'un capillaire. Ces billes sont rincées à l'EtOH. On ajoute 2 gouttes de phénol/EtOH (3/1), 2 gouttes de KCN/pyridine (0,0002 M), et 2 gouttes de ninhydrine dans EtOH (0,28 M). Le tube est chauffé à 100°C pendant 3 min et de nouveau 2 min.

Si la solution devient jaune/orange et que les billes sont incolores, c'est que le couplage est terminé. Si la solution prend une teinte bleu/vert et que les billes sont bleues cela signifie qu'il y a des amines primaires libres dans le milieu.

Bibliographie

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D***66**: 213–221.
- Agostini, F., Völler, J.-S., Kokschi, B., Acevedo-Rocha, C.G., Kubyshev, V., and Budisa, N. (2017) Biocatalysis with Unnatural Amino Acids: Enzymology Meets Xenobiology. *Angew Chem Int Ed***56**: 9680–9703.
- Ambrogelly, A., Palioura, S., and Söll, D. (2007) Natural expansion of the genetic code. *Nat Chem Biol***3**: 29–35.
- Amir, L., Carnally, S.A., Rayo, J., Rosenne, S., Melamed Yerushalmi, S., Schlesinger, O., *et al.* (2013) Surface Display of a Redox Enzyme and its Site-Specific Wiring to Gold Electrodes. *J Am Chem Soc***135**: 70–73.
- Amiram, M., Haimovich, A.D., Fan, C., Wang, Y.-S., Aerni, H.-R., Ntai, I., *et al.* (2015) Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat Biotechnol***33**: 1272–1279.
- Anand, R., Kaminski, P.A., and Ealick, S.E. (2004) Structures of Purine 2'-Deoxyribosyltransferase, Substrate Complexes, and the Ribosylated Enzyme Intermediate at 2.0 Å Resolution [†], [‡]. *Biochemistry***43**: 2384–2393.
- Anderson, J.C., Wu, N., Santoro, S.W., Lakshman, V., King, D.S., and Schultz, P.G. (2004) An expanded genetic code with a functional quadruplet codon. *Proc Natl Acad Sci***101**: 7566–7571.
- Appella, D.H., Christianson, L.A., Karle, I.L., Powell, D.R., and Gellman, S.H. (1996) β -Peptide Foldamers: Robust Helix Formation in a New Family of β -Amino Acid Oligomers. *J Am Chem Soc***118**: 13071–13072.
- Armstrong, S., Cook, W.J., Short, S.A., and Ealick, S.E. (1996a) Crystal structures of nucleoside 2-deoxyribosyltransferase in native and ligand-bound forms reveal architecture of the active site. *Structure***4**: 97–107.
- Armstrong, S.R., Steven, S., Steven, E., and David J.T., P. (1996b) Active Site Amino Acids That Participate in the Catalytic Mechanism of Nucleoside 2'-Deoxyribosyltransferase. *J Biol Chem***271**: 4978–4987.
- Arnez, J.G., and Moras, D. (1997) Structural and functional considerations of the aminacylation reaction. *TIBS***22**: 211–216.

- Barros-Álvarez, X., Turley, S., Ranade, R.M., Gillespie, J.R., Duster, N.A., Verlinde, C.L.M.J., *et al.* (2018) The crystal structure of the drug target *Mycobacterium tuberculosis* methionyl-tRNA synthetase in complex with a catalytic intermediate. *Acta Crystallogr Sect F Struct Biol Commun* **74**: 245–254.
- Belrhali, H., Yaremchuk, A., Tukalo, M., Berthet-Colominas, C., Rasmussen, B., Bösecke, P., *et al.* (1995) The structural basis for seryl-adenylate and Ap4A synthesis by seryl-tRNA synthetase. *Structure* **3**: 341–352.
- Berlicki, Ł., Kaske, M., Gutiérrez-Abad, R., Bernhardt, G., Illa, O., Ortuño, R.M., *et al.* (2013) Replacement of Thr32 and Gln34 in the C-Terminal Neuropeptide Y Fragment 25–36 by cis-Cyclobutane and cis-Cyclopentane β -Amino Acids Shifts Selectivity toward the Y4 Receptor. *J Med Chem* **56**: 8422–8431.
- Bhattacharjee, A., and Wallin, S. (2013) Exploring Protein-Peptide Binding Specificity through Computational Peptide Screening. *PLoS Comput Biol* **9**: e1003277.
- Blanquet, S., Crepin, T., Mechulam, Y., and Schmitt, E. (2005) Methionyl-tRNA Synthetases. In *The aminoacyl-tRNA synthetases*. Georgetown, Tex. p. 420.
- Blanquet, S., Dessen, P., and Iwatsubo, M. (1976) Antico-operative binding of bacterial and mammalian initiator tRNA^{Met} to methionyl-tRNA synthetase from *Escherichia coli*. *J Mol Biol* **103**: 765–784.
- Blanquet, S., Fayat, G., Poiret, M., and Waller, J.-P. (1975a) The mechanism of action of methionyl-tRNA synthetase from *Escherichia coli*. Inhibition by adenosine and 8-aminoadenosine of the amino-acid activation reaction. *Eur J Biochem* **51**: 567–571.
- Blanquet, S., Fayat, G., and Waller, J.P. (1974) The mechanism of action of methionyl-tRNA synthetase from *Escherichia coli*. Mechanism of the amino-acid activation reaction catalyzed by the native and the trypsin-modified enzymes. *Eur J Biochem* **44**: 343–351.
- Blanquet, S., Fayat, G., and Waller, J.P. (1975b) The amino acid activation reaction catalyzed by methionyl-transfer RNA synthetase: evidence for synergistic coupling between the sites for methionine, adenosine and pyrophosphate. *J Mol Biol* **94**: 1–15.
- Blanquet, S., Fayat, G., Waller, J.-P., and Iwatsubo, M. (1972) The Mechanism of Reaction of Methionyl-tRNA Synthetase from *Escherichia coli*. Interaction of the Enzyme with Ligands of the Amino-Acid-Activation Reaction. *Eur J Biochem* **24**: 461–469.
- Blanquet, S., Iwatsubo, M., and Waller, J.-P. (1973) The mechanism of action of methionyl-tRNA synthetase. 1. Fluorescence studies on tRNA^{Met} binding as a function of ligands, ions and pH. *Eur J Biochem* **36**: 213–226.

- Blow, D.M., Bhat, T.N., Metcalfe, A., Risler, J.L., Brunie, S., and Zelwer, C. (1983) Structural homology in the amino-terminal domains of two aminoacyl-tRNA synthetases. *J Mol Biol* **171**: 571–576.
- Bonnefond, L. (2005) Human mitochondrial TyrRS disobeys the tyrosine identity rules. *RNA* **11**: 558–562.
- Bornscheuer, U.T., Huisman, G.W., Kazlauskas, R.J., Lutz, S., Moore, J.C., and Robins, K. (2012) Engineering the third wave of biocatalysis. *Nature* **485**: 185–194.
- Brückner, A.M., Chakraborty, P., Gellman, S.H., and Diederichsen, U. (2003a) Molecular Architecture with Functionalized β -Peptide Helices. *Angew Chem Int Ed* **42**: 4395–4399.
- Brückner, A.M., Garcia, M., Marsh, A., Gellman, S.H., and Diederichsen, U. (2003b) Synthesis of Novel Nucleo- β -Amino Acids and Nucleobase-Functionalized β -Peptides. *Eur J Org Chem* **2003**: 3555–3561.
- Bruton, C.J., and Hartley, B.S. (1970) The subunits of methionyl-transfer-ribonucleic acid synthetase. *Biochem J* **117**: 18P–19P.
- Bryson, D.I., Fan, C., Guo, L.-T., Miller, C., Söll, D., and Liu, D.R. (2017) Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol* **13**: 1253–1260.
- Budisa, N. (2004) Prolegomena to Future Experimental Efforts on Genetic Code Engineering by Expanding Its Amino Acid Repertoire. *Angew Chem Int Ed* **43**: 6426–6463.
- Budisa, N., Rubini, M., Bae, J.H., Weyher, E., Wenger, W., Golbik, R., *et al.* (2002) Global Replacement of Tryptophan with Aminotryptophans Generates Non-Invasive Protein-Based Optical pH Sensors. *Angew Chem Int Ed* **41**: 4066–4069.
- Burbaum, J.J., and Schimmel, P. (1991) Structural relationships and the classification of aminoacyl-tRNA synthetases. *J Biol Chem* **266**: 16965–16968.
- Cabrele, C., Martinek, T.A., Reiser, O., and Berlicki, Ł. (2014) Peptides Containing β -Amino Acid Patterns: Challenges and Successes in Medicinal Chemistry. *J Med Chem* **57**: 9718–9739.
- Cahuzac, B., Berthonneau, E., Birlirakis, N., Guittet, E., and Mirande, M. (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. *EMBO J* **19**: 445–452.
- Cassio, D. (1968) Etude de la methionyl tRNA synthetase de Escherichia coli. 2. Inactivation selective et reversible de la capacite de charger le tRNA. *Eur J Biochem* **4**: 222–224.
- Cassio, D., Lemoine, F., Waller, J.-P., Sandrin, E., and Boissonnas, R.A. (1967) Selective Inhibition of Aminoacyl Ribonucleic Acid Synthetases by Aminoalkyl Adenylates. *Biochemistry* **6**: 827–836.

Cassio, D., and Waller, J.-P. (1971) Modification of methionyl-tRNA synthetase by proteolytic cleavage and properties of the trypsin-modified enzyme. *Eur J Biochem* **20**: 283–300.

Chapeville, F., Lipmann, F., vonEhrenstein, G., Weisblum, B., Ray, W.J.Jr., and Benzer, S. (1962) On the role of soluble ribonucleic acid in coding for aminoacids. *Proc Natl Acad Sci USA* **48**: 1086–1092.

Chen, P.R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B.H., and Schultz, P.G. (2009) A Facile System for Encoding Unnatural Amino Acids in Mammalian Cells. *Angew Chem Int Ed* **48**: 4052–4055.

Chin, J.W. (2017) Expanding and reprogramming the genetic code. *Nature* **550**: 53–60.

Clark, T.D., Buehler, L.K., and Ghadiri, M.R. (1998) Self-Assembling Cyclic β^3 -Peptide Nanotubes as Artificial Transmembrane Ion Channels. *J Am Chem Soc* **120**: 651–656.

Coulondre, C., and Miller, J.H. (1977) Genetic studies of the lac repressor. *J Mol Biol* **117**: 525–567.

Cowie, D.B., and Cohen, G.N. (1957) Biosynthesis by *Escherichia coli* of active altered proteins containing selenium instead of sulfur. *Biochim Biophys Acta* **26**: 252–261.

Crécy-Lagard, V.A. de, Bellalou, J., Mutzel, R., and Marlière, P. (2001) Long term adaptation of a microbial population to a permanent metabolic constraint: overcoming thymineless death by experimental evolution of *Escherichia coli*. *BMC Biotechnol* **7**.

Crepin, T., Schmitt, E., Blanquet, S., and Mechulam, Y. (2002) Structure and function of the C-terminal domain of methionyl-tRNA synthetase. *Biochemistry* **41**: 13003–11.

Crepin, T., Schmitt, E., Blanquet, S., and Mechulam, Y. (2004) Three-dimensional structure of methionyl-tRNA synthetase from *Pyrococcus abyssi*. *Biochemistry* **43**: 2635–44.

Crepin, T., Schmitt, E., Mechulam, Y., Sampson, P.B., Vaughan, M.D., Honek, J.F., and Blanquet, S. (2003) Use of Analogues of Methionine and Methionyl Adenylate to Sample Conformational Changes During Catalysis in *Escherichia coli* Methionyl-tRNA Synthetase. *J Mol Biol* **332**: 59–72.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 Å. *Nature* **347**: 249–255.

Dahiyat, B.I., and Mayo, S.L. (2008) Protein design automation. *Protein Sci* **5**: 895–903.

Dale, T., Sanderson, L.E., and Uhlenbeck, O.C. (2004) The Affinity of Elongation Factor Tu for an Aminoacyl-tRNA Is Modulated by the Esterified Amino Acid †. *Biochemistry* **43**: 6159–6166.

Daniels, D.S., Petersson, E.J., Qiu, J.X., and Schepartz, A. (2007) High-Resolution Structure of a β -Peptide Bundle. *J Am Chem Soc* **129**: 1532–1533.

Daura, X., Gademann, K., Schäfer, H., Jaun, B., Seebach, D., and Gunsteren, W.F. van (2001) The β -Peptide Hairpin in Solution: Conformational Study of a β -Hexapeptide in Methanol by NMR Spectroscopy and MD Simulation. *J Am Chem Soc* **123**: 2393–2404.

Daura, X., Gunsteren, W.F. van, Rigo, D., Jaun, B., and Seebach, D. (1997) Studying the Stability of a Helical β -Heptapeptide by Molecular Dynamics Simulations. *Chem - Eur J* **3**: 1410–1417.

Dedkova, L.M., Fahmi, N.E., Paul, R., Rosario, M. del, Zhang, L., Chen, S., *et al.* (2012) β -Puromycin Selection of Modified Ribosomes for *in Vitro* Incorporation of β -Amino Acids. *Biochemistry* **51**: 401–415.

Del Borgo, M.P., Mechler, A.I., Traore, D., Forsyth, C., Wilce, J.A., Wilce, M.C.J., *et al.* (2013) Supramolecular Self-Assembly of *N*-Acetyl-Capped β -Peptides Leads to Nano- to Macroscale Fiber Formation. *Angew Chem Int Ed* **52**: 8266–8270.

Dennis, P.P., and Bremer, H. (1974) Macromolecular Composition During Steady-State Growth of *Escherichia coli* B/r. *J BACTERIOL* **119**: 12.

Driggers, E.M., Hale, S.P., Lee, J., and Terrett, N.K. (2008) The exploration of macrocycles for drug discovery — an underexploited structural class. *Nat Rev Drug Discov* **7**: 608–624.

Ellefson, J.W., Meyer, A.J., Hughes, R.A., Cannon, J.R., Brodbelt, J.S., and Ellington, A.D. (2014) Directed evolution of genetic parts and circuits by compartmentalized partnered replication. *Nat Biotechnol* **32**: 97–101.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr* **D66**: 486–501.

Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**: 203–206.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011) A system for the continuous directed evolution of biomolecules. *Nature* **472**: 499–503.

Fan, C., Xiong, H., Reynolds, N.M., and Söll, D. (2015) Rationally evolving tRNA^{Pyl} for efficient incorporation of noncanonical amino acids. *Nucleic Acids Res* **43**: e156–e156.

Fayat, G., Blanquet, S., Dessen, P., Batelier, G., and Waller, J.-P. (1974) The molecular weight and subunit composition of phenylalanyl-tRNA synthetase from *Escherichia coli* K-12. *Biochimie* **56**: 35–41.

- Fayat, G., Fromant, M., and Blanquet, S. (1977) Couplings between the sites for methionine and adenosine 5'-triphosphate in the amino acid activation reaction catalyzed by trypsin-modified methionyl-transfer RNA synthetase from *Escherichia coli*. *Biochemistry* **16**: 2570–2579.
- Fersht, A.R. (1988) Relationships between apparent binding energies measured in site-directed mutagenesis experiments and energetics of binding and catalysis. *Biochemistry* **27**: 1577–1580.
- Fersht, A.R., and Dingwall, C. (1979a) Cysteinyl-tRNA synthetase from *Escherichia coli* does not need an editing mechanism to reject serine and alanine. High binding energy of small groups in specific molecular interactions. *Biochemistry* **18**: 1245–1249.
- Fersht, A.R., and Dingwall, C. (1979b) Establishing the misaminoacylation/deacylation of the tRNA pathway for the editing mechanism of prokaryotic and eukaryotic valyl-tRNA synthetases. *Biochemistry* **18**: 1238–1245.
- First, E.A., and Fersht, A.R. (1993a) Involvement of threonine 234 in catalysis of tyrosyl adenylate formation by tyrosyl-tRNA synthetase. *Biochemistry* **32**: 13644–13650.
- First, E.A., and Fersht, A.R. (1993b) Mutation of lysine 233 to alanine introduces positive cooperativity into tyrosyl-tRNA synthetase. *Biochemistry* **32**: 13651–13657.
- First, E.A., and Fersht, A.R. (1993c) Mutational and kinetic analysis of a mobile loop in tyrosyl-tRNA synthetase. *Biochemistry* **32**: 13658–13663.
- Forster, A.C., Tan, Z., Nalam, M.N.L., Lin, H., Qu, H., Cornish, V.W., and Blacklow, S.C. (2003) Programming peptidomimetic syntheses by translating genetic codes designed de novo. *Proc Natl Acad Sci* **100**: 6353–6357.
- Fourmy, D. (1993) Etude du centre catalytique de la methionyl-arnt synthetase d'e. Col. .
- Fourmy, D., Dardel, F., and Blanquet, S. (1993a) Methionyl-tRNA Synthetase Zinc Binding Domain. Three-dimensional Structure and Homology with Rubredoxine and gag Retroviral Proteins. *JMolBiol* **231**: 1078–1089.
- Fourmy, D., Mechulam, Y., and Blanquet, S. (1995) Crucial role of an idiosyncratic insertion in the Rossmann fold of class1 aminoacyl-tRNA synthetases: The case of methionyl-tRNA synthetase. *Biochemistry* **34**: 15681–15688.
- Fourmy, D., Meinnel, T., Mechulam, Y., and Blanquet, S. (1993b) Mapping of the Zinc Binding Domain of *Escherichia coli* Methionyl-tRNA Synthetase. *JMolBiol* **231**: 1068–1077.
- Frost, D.V., and Lish, P.M. (1975) Selenium in Biology. *Annu Rev Pharmacol* **15**: 259–284.

Fujino, T., Goto, Y., Suga, H., and Murakami, H. (2016) Ribosomal Synthesis of Peptides with Multiple β -Amino Acids. *J Am Chem Soc* **138**: 1962–1969.

Gademann, K., Ernst, M., Hoyer, D., and Seebach, D. (1999) Synthesis and Biological Evaluation of a Cyclo--tetrapeptide as a Somatostatin Analogue. *Angew Chem Int Ed* **38**: 1223–1226.

Gademann, K., Kimmerlin, T., Hoyer, D., and Seebach, D. (2001) Peptide Folding Induces High and Selective Affinity of a Linear and Small β -Peptide to the Human Somatostatin Receptor 4. *J Med Chem* **44**: 2460–2468.

Gale, A.J., Shi, J.P., and Schimmel, P. (1996) Evidence that specificity of microhelix charging by a class I tRNA synthetase occurs in the transition state of catalysis. *Biochemistry* **35**: 608–615.

Gan, R., Perez, J.G., Carlson, E.D., Ntai, I., Isaacs, F.J., Kelleher, N.L., and Jewett, M.C. (2017) Translation system engineering in *Escherichia coli* enhances non-canonical amino acid incorporation into proteins: Translation system engineering in *Escherichia coli*. *Biotechnol Bioeng* **114**: 1074–1086.

Gautier, A., Deiters, A., and Chin, J.W. (2011) Light-Activated Kinases Enable Temporal Dissection of Signaling Networks in Living Cells. *J Am Chem Soc* **133**: 2124–2127.

Gautier, A., and Hinner, M.J. (eds) (2015) *Site-specific protein labeling: methods and protocols*. Humana Pr, New York, NY.

Gelman, M.A., Richter, S., Cao, H., Umezawa, N., Gellman, S.H., and Rana, T.M. (2003) Selective Binding of TAR RNA by a Tat-Derived β -Peptide. *Org Lett* **5**: 3563–3565.

Ghosh, G., Kim, H.Y., Demaret, J.-P., Brunie, S., and Schulman, L.H. (1991) Arginine 395 is required for efficient in vivo and in vitro aminoacylation of tRNAs by *Escherichia coli* methionyl-tRNA synthetase. *Biochemistry* **30**: 11567–11574.

Ghosh, G., Pelka, H., and Schulman, L.H. (1990) Identification of the tRNA anticodon recognition site of *Escherichia coli* methionyl-tRNA synthetase. *Biochemistry* **29**: 2220–2225.

Giege, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res* **26**: 5017–5035.

Giegé, R., and Springer, M. (2016) Aminoacyl-tRNA Synthetases in the Bacterial World. *EcoSal* *Plus* **7**
<http://www.asmscience.org/content/journal/ecosalplus/10.1128/ecosalplus.ESP-0002-2016>. Accessed May 9, 2019.

- Giuliano, M.W., Horne, W.S., and Gellman, S.H. (2009) An α/β -Peptide Helix Bundle with a Pure β^3 -Amino Acid Core and a Distinctive Quaternary Structure. *J Am Chem Soc* **131**: 9860–9861.
- Gopalan, R.D., Del Borgo, M.P., Mechler, A.I., Perlmutter, P., and Aguilar, M.-I. (2015) Geometrically Precise Building Blocks: the Self-Assembly of β -Peptides. *Chem Biol* **22**: 1417–1423.
- Griffith, O.W. (1986) β -Amino Acids: Mammalian Metabolism and Utility as α -Amino Acid Analogues. *Annu Rev Biochem* **55**: 855–878.
- Guillon, L., Schmitt, E., Blanquet, S., and Mechulam, Y. (2005) Initiator tRNA binding by e/aIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. *Biochemistry* **44**: 15594–601.
- Guo, J., Melançon, C.E., Lee, H.S., Groff, D., and Schultz, P.G. (2009) Evolution of Amber Suppressor tRNAs for Efficient Bacterial Production of Proteins Containing Nonnatural Amino Acids. *Angew Chem Int Ed* **48**: 9148–9151.
- Hancock, S.M., Uprety, R., Deiters, A., and Chin, J.W. (2010) Expanding the Genetic Code of Yeast for Incorporation of Diverse Unnatural Amino Acids via a Pyrrolysyl-tRNA Synthetase/tRNA Pair. *J Am Chem Soc* **132**: 14819–14824.
- Hartman, M.C.T., Josephson, K., Lin, C.-W., and Szostak, J.W. (2007) An Expanded Set of Amino Acid Analogs for the Ribosomal Translation of Unnatural Peptides. *PLoS ONE* **2**: e972.
- Hartman, M.C.T., Josephson, K., and Szostak, J.W. (2006) Enzymatic aminoacylation of tRNA with unnatural amino acids. *Proc Natl Acad Sci* **103**: 4356–4361.
- Hauser, C.A.E., Deng, R., Mishra, A., Loo, Y., Khoe, U., Zhuang, F., *et al.* (2011) Natural tri- to hexapeptides self-assemble in water to amyloid β -type fiber aggregates by unexpected α -helical intermediate structures. *PNAS* **108**: 1361–66.
- Hecht, S., Alford, B., Kuroda, Y., and Kitano, S. (1978) Chemical Aminoacylation” of tRNA’s. *J Biol Chem* **10**: 4517–4520.
- Heinrikson, R.L., and Hartley, B.S. (1967) Purification and properties of methionyl-transfer-ribonucleic acid synthetase from *Escherichia coli*. *Biochem J* **105**: 17–24.
- Hendrickson, T.L., Nomanbhoy, T.K., Crécy-Lagard, V. de, Fukai, S., Nureki, O., Yokoyama, S., and Schimmel, P. (2002) Mutational Separation of Two Pathways for Editing by a Class I tRNA Synthetase. *Mol Cell* **9**: 353–362.
- Hendrickson, W.A. (1991) Determination of Macromolecular Structures from Anomalous-Diffraction of Synchrotron Radiation. *Science* **254**: 51–58.

- Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**: 1665–1672.
- Hest, J.C.M. van, Kiick, K.L., and Tirrell, D.A. (2000) Efficient Incorporation of Unsaturated Methionine Analogues into Proteins in Vivo. *J Am Chem Soc* **122**: 1282–1288.
- Horne, W.S., Price, J.L., Keck, J.L., and Gellman, S.H. (2007) Helix Bundle Quaternary Structure from α/β -Peptide Foldamers. *J Am Chem Soc* **129**: 4178–4180.
- Hou, Y.M., and Schimmel, P. (1989) Evidence that a major determinant for the identity of a transfer RNA is conserved in evolution. *Biochemistry* **28**: 6800–6804.
- Hountondji, C., Blanquet, S., and Lederer, F. (1985) Methionyl-tRNA synthetase from *Escherichia coli*. Primary structure at the binding site for the 3'-end of tRNA^{fMet}. *Biochemistry* **24**: 1175–1180.
- Hountondji, C., Dessen, P., and Blanquet, S. (1986) Sequence similarities among the family of aminoacyl-tRNA synthetases. *Biochimie* **68**: 1071–1078.
- Hunt, S. (1985) The Non-Protein Amino Acids. In *Chemistry and Biochemistry of the Amino Acids*. Springer, Dordrecht, pp. 55–138.
- Hyafil, F., Jacques, Y., Fayat, G., Fromant, M., Dessen, P., and Blanquet, S. (1976) Methionyl-tRNA synthetase from *Escherichia coli*: active stoichiometry and stopped-flow analysis of methionyl adenylate formation. *Biochemistry* **15**: 3678–3685.
- Ibba, M. (1999) Quality Control Mechanisms During Translation. *Science* **286**: 1893–1897.
- Ibba, M., Bono, J.L., Rosa, P.A., and Soll, D. (1997) Archaeal-type lysyl-tRNA synthetase in the Lyme disease spirochete *Borrelia burgdorferi*. *Proc Natl Acad Sci USA* **94**: 14383–8.
- Ibba, M., and Soll, D. (2001) The renaissance of aminoacyl-tRNA synthesis. *EMBO Rep* **2**: 382–387.
- Itoh, Y., Sekine, S. -i., Suetsugu, S., and Yokoyama, S. (2013) Tertiary structure of bacterial selenocysteine tRNA. *Nucleic Acids Res* **41**: 6729–6738.
- Iwane, Y., Hitomi, A., Murakami, H., Katoh, T., Goto, Y., and Suga, H. (2016) Expanding the amino acid repertoire of ribosomal polypeptide synthesis via the artificial division of codon boxes. *Nat Chem* **8**: 317–325.
- Jakubowski, H. (1991) Proofreading in vivo: editing of homocysteine by methionyl-tRNA synthetase in the yeast *Saccharomyces cerevisiae*. *EMBO J* **10**: 593–598.
- Jakubowski, H. (1999) Protein homocysteinylation: possible mechanism underlying pathological consequences of elevated homocysteine levels. *FASEB J* **13**: 2277–2283.

Jakubowski, H. (2011) Quality control in tRNA charging -- editing of homocysteine. *Acta Biochim Pol* **58** <http://ojs.ptbioch.edu.pl/index.php/abp/article/view/2259>. Accessed May 27, 2019.

Jakubowski, H., and Fersht, A.R. (1981) Alternative pathways for editing non-cognate amino acids by aminoacyl-tRNA synthetases. *Nucleic Acids Res* **9**: 3105–3117.

Jakubowski, H., and Goldman, E. (1992) Editing of errors in selection of amino acids for protein synthesis. *Microbiol Rev* **56**: 412–429.

Johnson, D.B.F., Xu, J., Shen, Z., Takimoto, J.K., Schultz, M.D., Schmitz, R.J., *et al.* (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat Chem Biol* **7**: 779–786.

Josephson, K., Hartman, M.C.T., and Szostak, J.W. (2005) Ribosomal Synthesis of Unnatural Peptides. *J Am Chem Soc* **127**: 11727–11735.

Kaiser, F., Bittrich, S., Salentin, S., Leberecht, C., Haupt, V.J., Krautwurst, S., *et al.* (2018) Backbone Brackets and Arginine Tweezers delineate Class I and Class II aminoacyl tRNA synthetases. *PLOS Comput Biol* **14**: e1006101.

Kaminska, M., Shalak, V., and Mirande, M. (2001) The Appended C-Domain of Human Methionyl-tRNA Synthetase Has a tRNA-Sequestering Function [†]. *Biochemistry* **40**: 14309–14316.

Katoh, T., and Suga, H. (2018) Ribosomal Incorporation of Consecutive β -Amino Acids. *J Am Chem Soc* **140**: 12159–12167.

Katoh, T., Tajima, K., and Suga, H. (2017) Consecutive Elongation of D-Amino Acids in Translation. *Cell Chem Biol* **24**: 46–54.

Katoh, T., Wohlgemuth, I., Nagano, M., Rodnina, M.V., and Suga, H. (2016) Essential structural elements in tRNA^{Pro} for EF-P-mediated alleviation of translation stalling. *Nat Commun* **7** <http://www.nature.com/articles/ncomms11657>. Accessed June 3, 2019.

Kawaguchi, S. -i. (2001) The crystal structure of the ttCsaA protein: an export-related chaperone from *Thermus thermophilus*. *EMBO J* **20**: 562–569.

Kiick, K.L., and Tirrell, D.A. (2000) Protein Engineering by In Vivo Incorporation of Non-Natural Amino Acids: Control of Incorporation of Methionine Analogues by Methionyl-tRNA Synthetase. *Tetrahedron* **56**: 9487–9493.

Kiick, K.L., Weberskirch, R., and Tirrell, D.A. (2001) Identification of an expanded set of translationally active methionine analogues in *Escherichia coli*. *FEBS Lett* **502**: 25–30.

- Kim, H.Y., Pelka, H., Brunie, S., and Schulman, L.H. (1993) Two separate peptides in *Escherichia coli* methionyl-tRNA synthetase form the anticodon binding site for methionine tRNA. *Biochemistry* **32**: 10506–10511.
- Kim, J., Kwon, S., Kim, S.H., Lee, C.-K., Lee, J.-H., Cho, S.J., *et al.* (2012) Microtubes with Rectangular Cross-Section by Self-Assembly of a Short β -Peptide Foldamer. *J Am Chem Soc* **134**: 20573–20576.
- Kingsbury, W.D., Boehm, J.C., Perry, D., and Gilvarg, C. (1984) Portage of various compounds into bacteria by attachment to glycine residues in peptides. *Proc Natl Acad Sci* **81**: 4573–4576.
- Kisumi, M., Sugiura, M., and Chibata, I. (1976a) Biosynthesis of Norvaline, Norleucine, and Homoisoleucine in *Serratia marcescens*. *J Biochem (Tokyo)* **80**: 333–339.
- Kisumi, M., Sugiura, M., Kato, J., and Chibata, I. (1976b) L-Norvaline and L-Homoisoleucine Formation by *Serratia marcescens*. *J Biochem (Tokyo)* **79**: 1021–1028.
- Koglin, N., Zorn, C., Beumer, R., Cabrele, C., Bubert, C., Sewald, N., *et al.* (2003) Analogues of Neuropeptide Y Containing β -Aminocyclopropane Carboxylic Acids are the Shortest Linear Peptides That Are Selective for the Y1 Receptor. *Angew Chem Int Ed* **42**: 202–205.
- Köhler, F., Zimmermann, A., Hager, M., and Sippel, A.E. (2004) A genetic, non-transcriptional assay for nuclear receptor ligand binding in yeast. *Gene* **337**: 113–119.
- Kraus, J., Soll, D., and Brooks Low, K. (1979) Glutamyl- γ -methyl ester acts as a methionine analogue in *Escherichia coli*: analogue resistant mutants map at the *metJ* and *metK* loci. *Genet Res* **33**: 49.
- Kudo, F., Miyanaga, A., and Eguchi, T. (2014) Biosynthesis of natural products containing β -amino acids. *Nat Prod Rep* **31**: 1056–1073.
- Kwon, S., Jeon, A., Yoo, S.H., Chung, I.S., and Lee, H.-S. (2010) Unprecedented Molecular Architectures by the Controlled Self-Assembly of a β -Peptide Foldamer. *Angew Chem Int Ed* **49**: 8232–8236.
- Lajoie, M.J., Rovner, A.J., Goodman, D.B., Aerni, H.-R., Haimovich, A.D., Kuznetsov, G., *et al.* (2013) Genomically Recoded Organisms Expand Biological Functions. *Science* **342**: 357–360.
- LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294**: 165–8.
- Larson, E.T., Kim, J.E., Zucker, F.H., Kelley, A., Mueller, N., Napuli, A.J., *et al.* (2011) Structure of *Leishmania* major methionyl-tRNA synthetase in complex with intermediate products methionyladenylate and pyrophosphate. *Biochimie* **93**: 570–582.

- Leatherbarrow, R.J., Fersht, A.R., and Winter, G. (1985) Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. *Proc Natl Acad Sci USA* **82**: 7840–7844.
- Ledoux, S., and Uhlenbeck, O.C. (2008) [3'-32P]-labeling tRNA with nucleotidyltransferase for assaying aminoacylation and peptide bond formation. *Methods* **44**: 74–80.
- Leduc, S. (1912) *La biologie synthétique*. A. Poinat.,.
- Lemeignan, B., Sonigo, P., and Marlière, P. (1993) Phenotypic Suppression by Incorporation of an Alien Amino Acid. *J Mol Biol* **231**: 161–166.
- Leon, O., and Schulman, L.H. (1987) tRNA recognition site of Escherichia coli methionyl-tRNA synthetase. *Biochemistry* **26**: 5416–5422.
- Leveque, F., Plateau, P., Dessen, P., and Blanquet, S. (1990) Homology of lysS and lysU, the two Escherichia coli genes encoding distinct lysyl-tRNA synthetase species. *Nucleic Acids Res* **8**: 305–312.
- Li, T., Vandesquille, M., Koukouli, F., Dudeffant, C., Youssef, I., Lenormand, P., *et al.* (2016) Camelid single-domain antibodies: A versatile tool for in vivo imaging of extracellular and intracellular brain targets. *J Controlled Release* **243**: 1–10.
- Ling, J., Reynolds, N., and Ibba, M. (2009) Aminoacyl-tRNA Synthesis and Translational Quality Control. *Annu Rev Microbiol* **63**: 61–78.
- Link, A.J., Mock, M.L., and Tirrell, D.A. (2003) Non-canonical amino acids in protein engineering. *Curr Opin Biotechnol* **14**: 603–609.
- Link, A.J., Vink, M.K.S., Agard, N.J., Prescher, J.A., Bertozzi, C.R., and Tirrell, D.A. (2006) Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc Natl Acad Sci* **103**: 10180–10185.
- Liu, C.C., and Schultz, P.G. (2010) Adding New Chemistries to the Genetic Code. *Annu Rev Biochem* **79**: 413–444.
- Mahmoud, Z.N., Gunnoo, S.B., Thomson, A.R., Fletcher, J.M., and Woolfson, D.N. (2011) Bioorthogonal dual functionalization of self-assembling peptide fibers. *Biomaterials* **32**: 3712–3720.
- Maini, R., Chowdhury, S.R., Dedkova, L.M., Roy, B., Daskalova, S.M., Paul, R., *et al.* (2015) Protein Synthesis with Ribosomes Selected for the Incorporation of β -Amino Acids. *Biochemistry* **54**: 3694–3706.
- Marlière, P., Patrouix, J., Döring, V., Herdewijn, P., Tricot, S., Cruveiller, S., *et al.* (2011) Chemical Evolution of a Bacterium's Genome. *Angew Chem Int Ed* **50**: 7109–7114.

- Mascarenhas, A.P., An, S., Rosen, A.E., Martinis, S.A., and Musier-Forsyth, K. (2009) Fidelity Mechanisms of the Aminoacyl-tRNA Synthetases. In *Protein Engineering*. Köhrer, C., and RajBhandary, U.L. (eds). Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 155–203 http://link.springer.com/10.1007/978-3-540-70941-1_6. Accessed July 15, 2019.
- Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., *et al.* (1999) Crystal structure of escherichia coli methionyl-tRNA synthetase highlights species-specific features. *J Mol Biol* **294**: 1287–97.
- Mechulam, Y., Schmitt, E., Panvert, M., Schmitter, J.M., Lapadat-Tapolsky, M., Meinnel, T., *et al.* (1991) Methionyl-tRNA synthetase from *Bacillus stearothermophilus*: structural and functional identities with the *Escherichia coli* enzyme. *Nucleic Acids Res* **19**: 3673–81.
- Meinnel, T., Mechulam, Y., Blanquet, S., and Fayat, G. (1991a) Binding of the anticodon domain of tRNA^{Met} to *Escherichia coli* methionyl-tRNA synthetase. *J Mol Biol* **220**: 205–208.
- Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S., and Fayat, G. (1991b) Selection of suppressor methionyl-tRNA synthetases: mapping the tRNA anticodon binding site. *Proc Natl Acad Sci U S A* **88**: 291–5.
- Mellot, P., Mechulam, Y., Le Corre, D., Blanquet, S., and Fayat, G. (1989) Identification of an amino acid region supporting specific methionyl-tRNA synthetase: tRNA recognition. *J Mol Biol* **208**: 429–43.
- Melo Czekster, C., Robertson, W.E., Walker, A.S., Söll, D., and Schepartz, A. (2016) *In Vivo* Biosynthesis of a β -Amino Acid-Containing Protein. *J Am Chem Soc* **138**: 5194–5197.
- Moor, N., Kotik-Kogan, O., Tworowski, D., Sukhanova, M., and Safo, M. (2006) The Crystal Structure of the Ternary Complex of Phenylalanyl-tRNA Synthetase with tRNA^{Phe} and a Phenylalanyl-Adenylate Analogue Reveals a Conformational Switch of the CCA End †. *Biochemistry* **45**: 10572–10583.
- Morales, A.J. (1999) Structure-specific tRNA-binding protein from the extreme thermophile *Aquifex aeolicus*. *EMBO J* **18**: 3475–3483.
- Morinaka, B.I., Lakis, E., Verest, M., Helf, M.J., Scalvenzi, T., Vagstad, A.L., *et al.* (2018) Natural noncanonical protein splicing yields products with diverse β -amino acid residues. *Science* **359**: 779–782.
- Mukai, T., Hayashi, A., Iraha, F., Sato, A., Ohtake, K., Yokoyama, S., and Sakamoto, K. (2010) Codon reassignment in the *Escherichia coli* genetic code. *Nucleic Acids Res* **38**: 8188–8195.

- Mukai, T., Hoshi, H., Ohtake, K., Takahashi, M., Yamaguchi, A., Hayashi, A., *et al.* (2015) Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. *Sci Rep* **5** <http://www.nature.com/articles/srep09699>. Accessed May 31, 2019.
- Murakami, H., Kourouklis, D., and Suga, H. (2003a) Using a Solid-Phase Ribozyme Aminoacylation System to Reprogram the Genetic Code. *Chem Biol* **10**: 1077–1084.
- Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat Methods* **3**: 357–359.
- Murakami, H., Saito, H., and Suga, H. (2003b) A Versatile tRNA Aminoacylation Catalyst Based on RNA. *Chem Biol* **10**: 655–662.
- Mutzel, R., Marlière, P., and Mazel, D. (2007) METHOD FOR OBTAINING NEW LIFE FORMS. .
- Nakanishi, K., Ogiso, Y., Nakama, T., Fukai, S., and Nureki, O. (2005) Structural basis for anticodon recognition by methionyl-tRNA synthetase. *Nat Struct Mol Biol* **12**: 931–932.
- Neumann, H. (2012) Rewiring translation - Genetic code expansion and its applications. *FEBS Lett* **586**: 2057–2064.
- Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008) Genetically encoding Nε-acetyllysine in recombinant proteins. *Nat Chem Biol* **4**: 232–234.
- Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J.W. (2010) Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**: 441–444.
- Nomanbhoy, T., Morales, A.J., Abraham, A.T., Vörtler, C.S., Giegé, R., and Schimmel, P. (2001) Simultaneous binding of two proteins to opposite sides of a single transfer RNA. *Nat Struct Biol* **8**: 5.
- Nomanbhoy, T.K., Hendrickson, T.L., and Schimmel, P. (1999) Transfer RNA-Dependent Translocation of Misactivated Amino Acids to Prevent Errors in Protein Synthesis. *Mol Cell* **4**: 519–528.
- Nozawa, K., O'Donoghue, P., Gundllapalli, S., Arais, Y., Ishitani, R., Umehara, T., *et al.* (2009) Pyrrolysyl-tRNA synthetase-tRNA^{Pyl} structure reveals the molecular basis of orthogonality. *Nature* **457**: 1163–1167.
- Nureki, O., Kohno, T., Sakamoto, K., Miyazawa, T., and Yokoyama, S. (1993) Chemical modification and mutagenesis studies on zinc binding of aminoacyl-tRNA synthetases. *J Biol Chem* **268**: 15368–73.

Nureki, O., Vassilyev, D., Katayanagi, K., Shimizu, T., Sekine, S., Kigawa, T., *et al.* (1995) Architectures of class-defining and specific domains of glutamyl-tRNA synthetase. *Science* **267**: 1958–1965.

O'Donoghue, P., and Luthey-Schulten, Z. (2003) On the Evolution of Structure in Aminoacyl-tRNA Synthetases. *Microbiol Mol Biol Rev* **67**: 550–573.

Ohtsuki, T., Yamamoto, H., Doi, Y., and Sisido, M. (2010) Use of EF-Tu mutants for determining and improving aminoacylation efficiency and for purifying aminoacyl tRNAs with non-natural amino acids. *J Biochem (Tokyo)* **148**: 239–246.

Ohuchi, M., Murakami, H., and Suga, H. (2007) The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Curr Opin Chem Biol* **11**: 537–542.

Oki, K., Sakamoto, K., Kobayashi, T., Sasaki, H.M., and Yokoyama, S. (2008) Transplantation of a tyrosine editing domain into a tyrosyl-tRNA synthetase variant enhances its specificity for a tyrosine analog. *Proc Natl Acad Sci* **105**: 13298–13303.

Orelle, C., Carlson, E.D., Szal, T., Florin, T., Jewett, M.C., and Mankin, A.S. (2015) Protein synthesis by ribosomes with tethered subunits. *Nature* **524**: 119–124.

Packer, M.S., and Liu, D.R. (2015) Methods for the directed evolution of proteins. *Nat Rev Genet* **16**: 379–394.

Pauling, L. (1957) The probability of errors in the process of synthesis of protein molecules. In *Festschrift für Pr. Dr. Arthur Stoll*. Birkhauser Verlag, Basel. pp. 597–602.

Perona, J.J., and Gruic-Sovulj, I. (2013) Synthetic and Editing Mechanisms of Aminoacyl-tRNA Synthetases. In *Aminoacyl-tRNA Synthetases in Biology and Medicine*. Kim, S. (ed.). Springer Netherlands, Dordrecht. pp. 1–41 http://link.springer.com/10.1007/128_2013_456. Accessed July 15, 2019.

Perryt, D., and Gilvarg, C. (1984) Spectrophotometric Determination of Affinities of Peptides for Their Transport Systems in Escherichia coli. *J BACTERIOL* **160**: 6.

Pizzey, C.L., Pomerantz, W.C., Sung, B.-J., Yuwono, V.M., Gellman, S.H., Hartgerink, J.D., *et al.* (2008) Characterization of nanofibers formed by self-assembly of β -peptide oligomers using small angle x-ray scattering. *J Chem Phys* **129**: 095103.

Pomerantz, W.C., Yuwono, V.M., Pizzey, C.L., Hartgerink, J.D., Abbott, N.L., and Gellman, S.H. (2008) Nanofibers and Lyotropic Liquid Crystals from a Class of Self-Assembling β -Peptides. *Angew Chem Int Ed* **47**: 1241–1244.

Porter, D.J.T., Merrill, B.M., and Short, S.A. (1995) Identification of the Active Site Nucleophile in Nucleoside 2-Deoxyribosyltransferase as Glutamic Acid 98. *J Biol Chem* **270**: 15551–15556.

- Prescher, J.A., and Bertozzi, C.R. (2005) Chemistry in living systems. *Nat Chem Biol* **1**: 13–21.
- Putney, S.D., and Schimmel, P. (1981) An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription. *Nature* **291**: 632–635.
- Raguse, T.L., Lai, J.R., LePlae, P.R., and Gellman, S.H. (2001) Toward β -Peptide Tertiary Structure: Self-Association of an Amphiphilic 14-Helix in Aqueous Solution. *Org Lett* **3**: 3963–3966.
- Raynal, N. (2002) Synthèse et caractérisation d'enchaînements de mimeseptidiques triés par modélisation moléculaire :application à la synthèse d'analogues du hCRF. .
- Reynolds, N.M., Lazazzera, B.A., and Ibba, M. (2010) Cellular mechanisms that control mistranslation. *Nat Rev Microbiol* **8**: 849–856.
- Richmond, M.H. (1962) The effect of amino acid analogues on growth and protein synthesis in microorganisms. *Bacteriol Rev* **26**: 398.
- Rossmann, M.G., Moras, D., and Olsen, K.W. (1974) Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**: 194–199.
- Rould, M.A., Perona, J.J., Söll, D., and Steitz, T.A. (1989) Structure of E.coli glutamyl-tRNA synthetase complexed with tRNAGln and ATP at 2.8 Å resolution. *Science* **246**: 1135–1142.
- Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., *et al.* (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^{Asp}. *Science* **252**: 1682–1689.
- Safro, M., Mosyak, L., Reshetnikova, L., Delarue, M., and Dauter, Z. (1993) X-Ray Studies on the Phenylalanine-tRNA Synthetase from *Thermus thermophilus*. *Cap Agde 15th Int TRNA Workshop* .
- Sanni, A., Walter, P., Boulanger, Y., Ebel, J.-P., and Fasiolo, F. (1991) Evolution of aminoacyl-tRNA synthetase quaternary structure and activity: *Saccharomyces cerevisiae* mitochondrial phenylalanyl-tRNA synthetase. *Proc Natl Acad Sci USA* **88**: 8387–8391.
- Schmitt, E., Meinnel, T., Blanquet, S., and Mechulam, Y. (1994) Methionyl-tRNA synthetase needs an intact and mobile 332KMSKS336 motif in catalysis of methionyl adenylate formation. *J Mol Biol* **242**: 566–577.
- Schmitt, E., Meinnel, T., Panvert, M., Mechulam, Y., and Blanquet, S. (1993) Two acidic residues of *Escherichia coli* methionyl-tRNA synthetase are negative discriminants towards the binding of non-cognate tRNA anticodons. *J Mol Biol* **233**: 615–628.

Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1995) Transition state stabilization by the “high” motif of class I aminoacyl- tRNA synthetases: the case of *Escherichia coli* methionyl-tRNA synthetase. *Nucleic Acids Res* **23**: 4793–8.

Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1998) Crystal structure of methionyl-tRNA^{Met} transformylase complexed with the initiator formyl-methionyl-tRNA^{Met}. *EMBO J* **17**: 6819–6826.

Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (2005) Structural basis for tRNA-dependent amidotransferase function. *Structure* **13**: 1421–33.

Schmitt, E., Panvert, M., Mechulam, Y., and Blanquet, S. (1997) General structure/function properties of microbial methionyl-tRNA synthetases. *Eur J Biochem* **246**: 539–47.

Schmitt, E., Tanrikulu, I.C., Yoo, T.H., Panvert, M., Tirrell, D.A., and Mechulam, Y. (2009) Switching from an induced-fit to a lock-and-key mechanism in an aminoacyl-tRNA synthetase with modified specificity. *J Mol Biol* **394**: 843–51.

Schön, A., Hottinger, H., and Söll, D. (1988a) Misaminoacylation and transamidation are required for protein biosynthesis in *Lactobacillus bulgaricus*. *Biochimie* **70**: 391–394.

Schön, A., Kannangara, C.G., Gough, S., and Söll, D. (1988b) Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* **331**: 187–190.

Schrader, J.M., Chapman, S.J., and Uhlenbeck, O.C. (2011) Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc Natl Acad Sci* **108**: 5215–5220.

Sedlacek, L., and Smith, L. (1988) BIOTRANSFORMATIONS OF STEROIDS. *Crit Rev Biotechnol* **7**: 187–236.

Seebach, D., Abele, S., Gademann, K., and Jaun, B. (1999) Pleated Sheets and Turns of β -Peptides with Proteinogenic Side Chains. *Angew Chem Int Ed* **38**: 1595–1597.

Seebach, D., Beck, A.K., and Bierbaum, D.J. (2004) The World of β - and γ -Peptides Comprised of Homologated Proteinogenic Amino Acids and Other Components. *Chem Biodivers* **1**: 1111–1239.

Seebach, D., Dubost, E., Mathad, R.I., Jaun, B., Limbach, M., Löweneck, M., *et al.* (2008) New Open-Chain and Cyclic Tetrapeptides, Consisting of α -, β^2 -, and β^3 -Amino-Acid Residues, as Somatostatin Mimics - A Survey. *Helv Chim Acta* **91**: 1736–1786.

Seebach, D., and Gardiner, J. (2008) β -Peptidic Peptidomimetics. *Acc Chem Res* **41**: 1366–1375.

Seebach, D., Hook, D.F., and Glättli, A. (2006) Helices and other secondary structures of β - and γ -peptides. *Biopolymers* **84**: 23–37.

Seebach, D., Overhand, M., Kühnle, F.N.M., Martinoni, B., Oberer, L., Hommel, U., and Widmer, H. (1996) β -Peptides: Synthesis by Arndt-Eistert homologation with concomitant peptide coupling. Structure determination by NMR and CD spectroscopy and by X-ray crystallography. Helical secondary structure of a β -hexapeptide in solution and its stability towards pe. *Helv Chim Acta* **79**: 913–941.

Seebach, D., Rueping, M., Arvidsson, P.I., Kimmerlin, T., Micuch, P., Noti, C., *et al.* (2001) Linear, Peptidase-Resistant β 2/ β 3-Di- and α / β 3-Tetrapeptide Derivatives with Nanomolar Affinities to a Human Somatostatin Receptor, Preliminary Communication. *Helv Chim Acta* **84**: 3503–3510.

Senger, B., Despons, L., Walter, P., Jakubowski, H., and Fasiolo, F. (2001) Yeast cytoplasmic and mitochondrial methionyl-tRNA synthetases: two structural frameworks for identical functions. *J Mol Biol* **311**: 205–216.

Serre, L., Verdon, G., Choinowski, T., Hervouet, N., Risler, J.-L., and Zelwer, C. (2001) How methionyl-tRNA synthetase creates its amino acid recognition pocket upon l-methionine binding. *J Mol Biol* **306**: 863–876.

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**: 751–755.

Shimizu, Y., Kuruma, Y., Ying, B.-W., Umekage, S., and Ueda, T. (2006) Cell-free translation systems for protein engineering. *FEBS J* **273**: 4133–4140.

Simonson, T., Gaillard, T., Mignon, D., Schmidt am Busch, M., Lopes, A., Amara, N., *et al.* (2013) Computational protein design: The proteus software and selected applications. *J Comput Chem* **34**: 2472–2484.

Sissler, M., Lorber, B., Messmer, M., Schaller, A., Pütz, J., and Florentz, C. (2008) Handling mammalian mitochondrial tRNAs and aminoacyl-tRNA synthetases for functional and structural characterization. *Methods* **44**: 176–189.

Spicer, C.D., and Davis, B.G. (2014) Selective chemical protein modification. *Nat Commun* **5** <http://www.nature.com/articles/ncomms5740>. Accessed May 28, 2019.

Sugiura, I., Nureki, O., Ugaji-Yoshikawa, Y., Kuwabara, S., Shimada, A., Tateno, M., *et al.* (2000) The 2.0 Å crystal structure of *Thermus thermophilus* methionyl-tRNA synthetase reveals two RNA-binding modules. *Structure* **8**: 197–208.

- Suzuki, T., Miller, C., Guo, L.-T., Ho, J.M.L., Bryson, D.I., Wang, Y.-S., *et al.* (2017) Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat Chem Biol* **13**: 1261–1266.
- Swairjo, M.A. (2000) Crystal structure of Trbp111: a structure-specific tRNA-binding protein. *EMBO J* **19**: 6287–6298.
- Tanrikulu, I.C., Schmitt, E., Mechulam, Y., Goddard, W.A., 3rd, and Tirrell, D.A. (2009) Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *Proc Natl Acad Sci U A* **106**: 15285–90.
- Terada, T., Nureki, O., Ishitani, R., Ambrogelly, A., Ibba, M., Söll, D., and Yokoyama, S. (2002) Functional convergence of two lysyl-tRNA synthetases with unrelated topologies. *Nat Struct Biol* **9**: 257–262.
- Tuffery, P., Etchebest, C., Hazout, S., and Lavery, R. (1991) A New Approach to the Rapid Determination of Protein Side Chain Conformations. *J Biomol Struct Dyn* **8**: 1267–1289.
- Tzima, E., and Schimmel, P. (2006) Inhibition of tumor angiogenesis by a natural fragment of a tRNA synthetase. *Trends Biochem Sci* **31**: 7–10.
- Uhlenbeck, O.C., and Schrader, J.M. (2018) Evolutionary tuning impacts the design of bacterial tRNAs for the incorporation of unnatural amino acids by ribosomes. *Curr Opin Chem Biol* **46**: 138–145.
- Villa, F., Panel, N., Chen, X., and Simonson, T. (2018) Adaptive landscape flattening in amino acid sequence space for the computational design of protein:peptide binding. *J Chem Phys* **149**: 072302.
- Waller, J.-P. (1963) The NH₂-terminal residues of the proteins from cell-free extracts of *E. coli*. *J Mol Biol* **7**: 483–IN1.
- Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., and Church, G.M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**: 894–898.
- Wang, K., Sachdeva, A., Cox, D.J., Wilf, N.M., Lang, K., Wallace, S., *et al.* (2014) Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat Chem* **6**: 393–403.
- Wang, L., and Schultz, P.G. (2005) Expanding the Genetic Code. *Angew Chem Int Ed* **44**: 34–66.
- Wang, P., Tang, Y., and Tirrell, D.A. (2003) Incorporation of Trifluoroisoleucine into Proteins in Vivo. *J Am Chem Soc* **125**: 6900–6906.

Wang, P.S.P., and Schepartz, A. (2016) β -Peptide bundles: Design. Build. Analyze. Biosynthesize. *Chem Commun* **52**: 7420–7432.

Ward, W.H.J., Holdgate, G.A., Rowsell, S., McLean, E.G., Pauptit, R.A., Clayton, E., *et al.* (1999) Kinetic and Structural Characteristics of the Inhibition of Enoyl (Acyl Carrier Protein) Reductase by Triclosan ‡. *Biochemistry* **38**: 12514–12525.

Wernisch, L., Hery, S., and Wodak, S.J. (2000) Automatic protein design with all atom force-fields by exact and heuristic optimization 1 Edited by J. Thornton. *J Mol Biol* **301**: 713–736.

Wolf, Y.I., Aravind, L., Grishin, N.V., and Koonin, E.V. (1999) Evolution of Aminoacyl-tRNA Synthetases—Analysis of Unique Domain Architectures and Phylogenetic Trees Reveals a Complex History of Horizontal Gene Transfer Events. *Genome Res* 689–710.

Wu, I.-L., Patterson, M.A., Carpenter Desai, H.E., Mehl, R.A., Giorgi, G., and Conticello, V.P. (2013) Multiple Site-Selective Insertions of Noncanonical Amino Acids into Sequence-Repetitive Polypeptides. *ChemBioChem* **14**: 968–978.

Yasumoto, T., and Satake, M. (1998) Bioactive Compounds from Marine Microalgae. *Chim Int J Chem* **52**: 63–68.

Young, F.S., and Furano, A.V. (1981) Regulation of the synthesis of *E. coli* elongation factor TU. *Cell* **24**: 695–706.

Young, T.S., and Schultz, P.G. (2010) Beyond the Canonical 20 Amino Acids: Expanding the Genetic Lexicon. *J Biol Chem* **285**: 11039–11044.

Zhang, M.S., Brunner, S.F., Huguenin-Dezot, N., Liang, A.D., Schmied, W.H., Rogerson, D.T., and Chin, J.W. (2017a) Biosynthesis and genetic encoding of phosphothreonine through parallel selection and deep sequencing. *Nat Methods* **14**: 729–736.

Zhang, Y., Ptacin, J.L., Fischer, E.C., Aerni, H.R., Caffaro, C.E., San Jose, K., *et al.* (2017b) A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551**: 644–647.

Annexe

RESEARCH ARTICLE

Adaptive landscape flattening allows the design of both enzyme: Substrate binding and catalytic power

Vaitea Opuu, Giuliano Nigro, Thomas Gaillard, Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson*

Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, Palaiseau, France

* thomas.simonson@polytechnique.fr



Abstract

Designed enzymes are of fundamental and technological interest. Experimental directed evolution still has significant limitations, and computational approaches are a complementary route. A designed enzyme should satisfy multiple criteria: stability, substrate binding, transition state binding. Such multi-objective design is computationally challenging. Two recent studies used adaptive importance sampling Monte Carlo to redesign proteins for ligand binding. By first flattening the energy landscape of the apo protein, they obtained positive design for the bound state and negative design for the unbound. We have now extended the method to design an enzyme for specific transition state binding, *i.e.*, for its catalytic power. We considered methionyl-tRNA synthetase (MetRS), which attaches methionine (Met) to its cognate tRNA, establishing codon identity. Previously, MetRS and other synthetases have been redesigned by experimental directed evolution to accept noncanonical amino acids as substrates, leading to genetic code expansion. Here, we have redesigned MetRS computationally to bind several ligands: the Met analog azidonorleucine, methionyl-adenylate (MetAMP), and the activated ligands that form the transition state for MetAMP production. Enzyme mutants known to have azidonorleucine activity were recovered by the design calculations, and 17 mutants predicted to bind MetAMP were characterized experimentally and all found to be active. Mutants predicted to have low activation free energies for MetAMP production were found to be active and the predicted reaction rates agreed well with the experimental values. We suggest the present method should become the paradigm for computational enzyme design.

OPEN ACCESS

Citation: Opuu V, Nigro G, Gaillard T, Schmitt E, Mechulam Y, Simonson T (2020) Adaptive landscape flattening allows the design of both enzyme: Substrate binding and catalytic power. PLoS Comput Biol 16(1): e1007600. <https://doi.org/10.1371/journal.pcbi.1007600>

Editor: Alexey Onufriev, Virginia Tech, UNITED STATES

Received: September 20, 2019

Accepted: December 11, 2019

Published: January 9, 2020

Copyright: © 2020 Opuu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Designed enzymes are of major interest. Experimental directed evolution still has significant limitations, and computational approaches are another route. Enzymes must be stable, bind substrates, and be powerful catalysts. It is challenging to design for all these properties. A method to design substrate binding was proposed recently. It used an adaptive Monte Carlo method to explore mutations of a few amino acids near the substrate. A

bias energy was gradually “learned” such that, in the absence of the ligand, the simulation visited most of the possible protein mutations with comparable probabilities. Remarkably, a simulation of the protein:ligand complex, including the bias, will then preferentially sample tight-binding sequences. We generalized the method to design binding *specificity*. We tested it for the methionyl-tRNA synthetase enzyme, which has been engineered in order to expand the genetic code. We redesigned the enzyme to obtain variants with low activation free energies for the catalytic step. The variants proposed by the simulations were shown experimentally to be active, and the predicted activation free energies were in reasonable agreement with the experimental values. We expect the new method will become the paradigm for computational enzyme design.

Introduction

One of the most important challenges in computational protein design (CPD) is to modify a protein so that it will bind a given ligand [1–4]. This is essential for problems like enzyme design, biosensor design, and building tailored protein assemblies. To design ligand binding means optimizing a free energy difference between bound and unbound states. This two-state optimization is not directly tractable by the most common CPD methods, such as simulated annealing, plain Monte Carlo (MC), or simple branch-and-bound and dead end elimination methods [4, 5]. Rather, most studies have used either heuristic methods that optimize the *bound state* energy [1–4, 6], or enumeration methods that are rigorous but expensive and explore a limited free energy range [7–10].

Recently, a new approach was proposed, using Monte Carlo simulation and importance sampling. The energy landscape in sequence space is flattened adaptively over the course of a simulation, thanks to a bias potential [11]. Flattening can be done for the bound state, the unbound state, or both [12]. Remarkably, this leads to a situation where sequence variants are sampled according to a Boltzmann distribution controlled by the *binding free energy*, exactly the quantity we want to select for. Several variations have been employed, including one that used molecular dynamics instead of MC [13]. The method allows sequences to be designed for binding affinity, but also binding *specificity*. This is especially important for enzyme design, since catalytic power is directly related to the enzyme’s ability to preferentially stabilize the transition state [14]. We apply the method here to an enzyme of biological and technological importance, methionyl-tRNA synthetase (MetRS). We demonstrate that the method can be used to design an enzyme for its catalytic power.

Each aminoacyl-tRNA synthetase (aaRS) attaches a specific amino acid to a tRNA that carries the corresponding anticodon, establishing the genetic code [15]. Two reactions are catalyzed. In the first, the amino acid reacts with ATP to give aaAMP and pyrophosphate. In the second, tRNA reacts with aaAMP. For MetRS, the first reaction does not require tRNA. Several aaRSs have been engineered experimentally to bind noncanonical amino acids (ncAAs) [16–20]. Obtaining an aaRS that binds an ncAA and uses it as a substrate is a key step to allow the ncAA to become part of an expanded genetic code [17, 20, 21]. The ncAA can then be genetically encoded and incorporated into proteins by the cellular machinery. Several MetRS variants that accepted the ncAA azidonorleucine (AnL) as a substrate were obtained earlier by experimental directed evolution [22]. The AnL azide group can be used for protein labeling and imaging.

The design procedure has two stages. First, a bias potential is optimized adaptively over the course of a MC simulation of the apo protein. The adaptation method is closely analogous to

the Wang-Landau and metadynamics approaches [23, 24]. The bias is chosen so that all the allowed residue types achieve comparable probabilities at all mutating positions. This implies that the free energy landscape in sequence space has been flattened, and the bias of each sequence is approximately the opposite of its apo free energy. In the second stage, the holo state is simulated. The bias is included in the energy function, “subtracting out” the apo free energy. Thus, the method achieves positive design for the bound state and negative design for the unbound. The sequences sampled in the second stage are distributed according to their binding free energies, with tight binders exponentially enriched.

In an analogous procedure, a bias potential can be optimized for the protein bound to one ligand, say L . Then a complex with another ligand is simulated, say, L' , including the bias. The sequences sampled preferentially in the second simulation are those with a strong binding free energy difference between the two ligands, *i.e.*, the most *specific* binders. Importantly, L' can be an activated, transition state ligand, while L is the non-activated substrate. In this case, the first simulation flattens the ground state landscape, while the second preferentially samples sequences that stabilize the transition state, relative to the ground state. Thus, the method can be used to select directly for low activation free energies. It is then straightforward to rank the sampled sequences based on their catalytic efficiency, the ratio between the rate constant for the catalytic step, k_{cat} and the Michaelis constant K_M .

Here, we report CPD calculations that aim to increase the binding of several ligands by MetRS. We first considered AnL. Three residues in the active site were allowed to mutate. The CPD method was tested for its ability to recover the known experimental variants [22]. The top six experimental variants were visited by the MC simulations and were highly ranked among the predicted sequences. We next considered the natural ligand methionyl adenylate (MetAMP). Another set of three residues near the ligand side chain were allowed to mutate. The wildtype sequence was highly ranked by the computational design. 17 other sequences among the top 40 predictions were tested experimentally and all found to be active. The computed binding free energy differences between variants were mostly in good agreement with the experimental values, obtained from kinetic measurements of the enzyme reaction. Next, we predicted MetRS variants that were specifically designed to bind the transition state for the enzymatic reaction $\text{Met} + \text{ATP} \rightarrow \text{MetAMP} + \text{PP}_i$. The wildtype enzyme was highly ranked among 5832 possible variants, and for 20 variants that were characterized experimentally, the transition state binding free energies from the simulations were in good agreement with the values deduced from the experimental reaction rates. These calculations represent the first time an enzyme is specifically designed to optimize its transition state binding free energy relative to ground state binding, *i.e.*, its catalytic power. We expect the method will become the paradigm for computational design of enzymes.

Materials and methods

Theoretical approach: Designing for ligand binding

Stage 1: Adaptive apo simulation. We consider a polypeptide, with or without a bound ligand. Below, we will use a fixed backbone geometry, but the method is valid with a flexible backbone. Side chains can explore a few discrete conformations, or rotamers, and a few selected positions are allowed to mutate. In a first stage, we perform a MC exploration of the protein with no ligand, using the usual Metropolis-Hastings scheme [25–27]. We gradually increment a bias potential until all the side chain types at the mutating positions have roughly equal populations, thus flattening the free energy landscape. We number the mutating positions arbitrarily

1, ..., p . The bias E^B at time t has the form:

$$E^B(s_1(t), s_2(t), \dots, s_p(t); t) = \sum_i E_i^B(s_i(t); t) + \sum_{i < j} E_{ij}^B(s_i(t), s_j(t); t) \quad (1)$$

Here, $s_i(t)$ represents the side chain type at position i . The first sum is over single amino acid positions; the second is over pairs. The individual terms are updated at regular intervals of length T . At each update, whichever sequence variant ($s_1(t), s_2(t), \dots, s_p(t)$) is populated is penalized by adding an increment $e_i^B(s_i(t); t)$ or $e_{ij}^B(s_i(t), s_j(t); t)$ to each corresponding term in the bias. The increments have the form:

$$e_i^B(s_i(t); t) = e_0 \exp[-E_i^B(s_i(t); t)/E_0] \quad (2)$$

$$e_{ij}^B(s_i(t), s_j(t); t) = e_0 \exp[-E_{ij}^B(s_i(t), s_j(t); t)/E_0] \quad (3)$$

where e_0 and E_0 are constant energies. Thus, the increments decrease exponentially as the bias increases. This scheme is adapted from well-tempered metadynamics [24, 28, 29]. The individual bias terms depend on the system history, and can be written:

$$E_i^B(s; t) = \sum_{n; nT < t} e_i^B(s; nT) \delta_{s, s_i(nT)} \quad (4)$$

$$E_{ij}^B(s, s'; t) = \sum_{n; nT < t} e_{ij}^B(s, s'; nT) \delta_{s, s_i(nT)} \delta_{s', s_j(nT)} \quad (5)$$

where $\delta_{a,b}$ is the Kronecker delta. Over time, the bias for the most probable states grows until it pushes the system into other regions of sequence space. Two-position biases were implemented in the Proteus software [30, 31] during this work.

Stage 2: Biased holo simulation. In the second stage, the protein:ligand complex is simulated using the bias potential from stage 1. The sampled population of a sequence S is normalized to give a probability, denoted $\tilde{p}_H(S)$, where the subscript means “holo” and the tilde indicates that the bias is present. The apo state probability $\tilde{p}_A(S)$ was obtained in stage 1. Both probabilities can be converted into free energies $\tilde{G}$:

$$\begin{aligned} \tilde{p}_X(S) &= \frac{1}{Z_X} \exp(-\tilde{G}_X(S)/kT) \\ \tilde{G}_X(S) &= -kT \ln \tilde{p}_X(S) - kT \ln Z_X \end{aligned} \quad (6)$$

where $X = A$ or H and Z_X is a normalization factor that depends on X but not S . We also have a relation between the free energies with and without the bias:

$$\tilde{G}_X(S) = G_X(S) + E^B(S) \quad (7)$$

whose (straightforward) derivation is given in the Supporting Appendix (S1 File). Note that if the apo state flattening were ideal, $\tilde{p}_A(S)$ would be a constant, so that (from Eqs 6 and 7) $E^B(S) = -G_A(S)$, up to an additive constant. Thus, the ideal bias is the opposite of the apo free energy.

The binding free energy relative to a reference sequence R can be deduced from the populations. We have:

$$\begin{aligned}\Delta\tilde{G}(S) &= (\tilde{G}_H(S) - \tilde{G}_A(S)) - (\tilde{G}_H(R) - \tilde{G}_A(R)) \\ &= -kT \ln \frac{\tilde{p}_H(S)}{\tilde{p}_H(R)} + kT \ln \frac{\tilde{p}_A(S)}{\tilde{p}_A(R)} \\ &= \Delta\Delta G(S)\end{aligned}\quad (8)$$

Since the bias is the same in the bound and unbound states, it cancels out from $\Delta\tilde{G}(S)$, which is equal to the relative binding free energy *in the absence of bias*, $\Delta\Delta G(S)$. While the bias does not appear explicitly in (8), it is essential for accurate sampling. Perfect flattening, however, is not usually achieved, nor is it needed.

In the holo state, the probability of a sequence S (with bias) is:

$$\tilde{p}_H(S) \propto \exp\left(-\frac{\tilde{G}_H(S)}{kT}\right) = \exp\left(-\frac{G_H(S) + E^B(S)}{kT}\right) \approx \exp\left(-\frac{G_H(S) - G_A(S)}{kT}\right) \quad (9)$$

Thus, holo sampling follows a Boltzmann distribution governed by $G_H(S) + E^B(S)$, which is approximately the binding free energy $G_H(S) - G_A(S)$. This is exactly the quantity we want to design for. If the apo state is well-flattened, the biased holo simulation will be exponentially enriched in tight binders.

Energy function and matrix

The energy was computed using either an MMGBLK or an MMGBSA function (“molecular mechanics + Generalized Born + Lazaridis-Karplus” or “Surface Area”):

$$E = E_{MM} + E_{GB} + E_{LK|SA} \quad (10)$$

The MM term used the Amber ff99SB force field [30, 32]. The SA term was described earlier [33–35]. The LK term and its parameterization were described earlier [35]. The GB term corresponds to a variant very similar to the one used in Amber, detailed in previous articles [33, 36, 37]. To make the calculation efficient, we compared two strategies. The first used a Native Environment Approximation (NEA), where the GB solvation radii for a given side chain were computed with the rest of the system in its native conformation [36, 38]. The second used a “Fluctuating Dielectric Boundary” (FDB) method, where the GB interaction between two residues I, J was expressed as a polynomial function of their solvation radii [39]. These were kept up to date over the course of the MC simulation, so the GB interaction could be deduced with little additional calculation [37, 39]. The solvent dielectric constant was 80; the protein one was 4.0 with the GBSA variants and 6.8 with GBLK [35]. Each solvent model is referred to by its GB variant and nonpolar term; for example, the FDBLK model combines FDB with LK.

To allow very fast MC simulations, we precomputed an energy matrix for each system [34, 40]. For each pair of residues I, J and all their allowed types and rotamers, we performed a short energy minimization (15 conjugate gradient steps) [30]. The backbone was fixed (in its crystal geometry) and the energy only included interactions between the two side chains and with the backbone. At the end of the minimization, we computed the interaction energy between the two side chains. Side chain–backbone interaction energies were computed similarly (and formed the matrix diagonal) [30].

Structural models

MetRS:AnL and MetRS:MetAMP complexes. For MetRS:AnL, we started from the crystal structure of a complex between a triple mutant of *E. coli* MetRS and AnL (PDB code 3H9B) [41]. The protein mutations were L13S, Y260L, H301L. We refer to this mutant as SLL. The protein backbone was held fixed. Side chains more than 20 Å from the ligand were held fixed. The other side chains were allowed to explore rotamers, taken from the Tuffery library, augmented to allow multiple orientations for certain hydrogen atoms [42, 43]. Side chains 13 and 301 were allowed to mutate into the following 14 types: ACDEHIKLMNQSTV; position 260 was allowed to mutate into the same types, except that Tyr replaced Asp. Thus, there were $14^3 = 2744$ possible sequences in all. Histidine protonation states at non-mutating positions were assigned by visual inspection of the 3D structure. System preparation was done using the protX module of the Proteus design software [31].

For MetRS:MetAMP, we started from a crystal complex (PDB code 1PG0) between *E. coli* MetRS and a methionyl adenylate (MetAMP) analogue [44]. The protein backbone was held fixed. Side chains more than 20 Å from the ligand were held fixed. The other side chains were allowed to explore rotamers [42, 43]. Side chains 13, 256 and 297 were allowed to mutate into all types except Gly or Pro, for a total of 5832 possible sequences in all. Histidine protonation states at non-mutating positions were assigned by visual inspection of the 3D structure.

Unfolded state. The unfolded state energy was estimated with a tri-peptide model [45]. For each mutating position, side chain type, and rotamer, we computed the interaction between the side chain and the tri-peptide it forms with the two adjacent backbone and C_β groups. Then, for each allowed type, we computed the energy of the best rotamer and averaged over mutating positions. The mean energy for each type was taken to be its contribution to the unfolded state energy. The contributions of the mutating positions were summed to give the total unfolded energy.

Ligand force field and rotamers

Force field. For the AnL azido group, we used atomic charges and van der Waals parameters obtained earlier for azidophenylalanine [46]. Parameters for the implicit solvent energy terms were assigned by analogy to existing groups. For methionyl adenylate (MetAMP), we mostly used existing Met and AMP parameters. For atoms close to the Met:AMP junction, we used atomic charges computed earlier for ThrAMP (G. Monard, personal communication) from *ab initio* quantum chemistry, in a manner consistent with the rest of the Amber force field [32]. Van der Waals parameters for atoms near the junction were assigned the same types as in Met or AMP. Parameters for bond lengths, angles and dihedrals involving junction atoms were taken from the experimental geometry of MetAMP. Stiffness parameters were assigned by analogy to existing parameters. The complete set of parameters for AnL and MetAMP is in Supplementary Material (S2 and S3 Files, respectively).

Rotamers. AnL was positioned in the protein complex so that its backbone had the position occupied in the MetRS:AnL crystal structure [41]. The ligand's side chain was allowed to explore rotamers. These were defined by the usual side chain rotamers of Met [42, 43]. We started by positioning Met in the pocket by superimposing it on AnL in the mutant MetRS:AnL crystal complex (PDB code 3H9B). We then positioned the 17 Met side chain rotamers from the Tuffery library. We extracted the AnL side chain from the experimental complex and superimposed it on each of the 17 Met rotamers, producing 17 AnL conformers. Finally, for each one, we performed a short energy minimization with the AnL backbone held fixed. The 17 minimized conformers defined the AnL rotamers. Notice that with this procedure, the azido group always had the same orientation relative to the aliphatic part of the AnL side

chain. For MetAMP, we allowed the Met rotamers from the Tuffery library, with the rest of the ligand held fixed. The ϕ and ψ dihedral angles around the MetAMP C_α were not allowed to rotate and the whole AMP moiety stayed fixed.

Modeling the MetRS transition state complex

MetRS catalyzes two reactions. In the first, Met reacts with ATP to give MetAMP and pyrophosphate. In the second, tRNA reacts with MetAMP. Here, we considered the first reaction, which occurs in the absence of tRNA. A model for the ground state ligands Met + ATP was first obtained, starting from the crystal complex between MetAMP and PP_i (PDB code 3KFL). The covalent structure was reset to that of Met + ATP and the geometry was adjusted by a short energy minimization. The complex included a magnesium ion. Next, a model for the activated ligand $[Met:ATP]^\ddagger$ was obtained, starting from the Met + ATP complex. First, a phosphate and carboxylate fragment were positioned in a geometry close to the expected pentacoordinate transition state arrangement [44, 47–49] and an *ab initio* energy minimization was done, including planarity constraints for the phosphorus and three oxygens. This led to a length of 2.4 Å for the P–O bonds perpendicular to the plane. Next, the molecular mechanics model was constructed. A covalent bond was introduced between the reacting Met carboxylate oxygen and the α phosphorus atom. The lengths for this bond and the symmetric one on the other side of the phosphorus were set to 2.4 Å. Planarity restraints were imposed on the phosphorus and the three α phosphate oxygens. A short energy minimization was done (with molecular mechanics). This led to an α phosphate geometry with three oxygens in plane and two perpendicular (Fig 1), as expected for in-line attack of the Met carboxylate on the

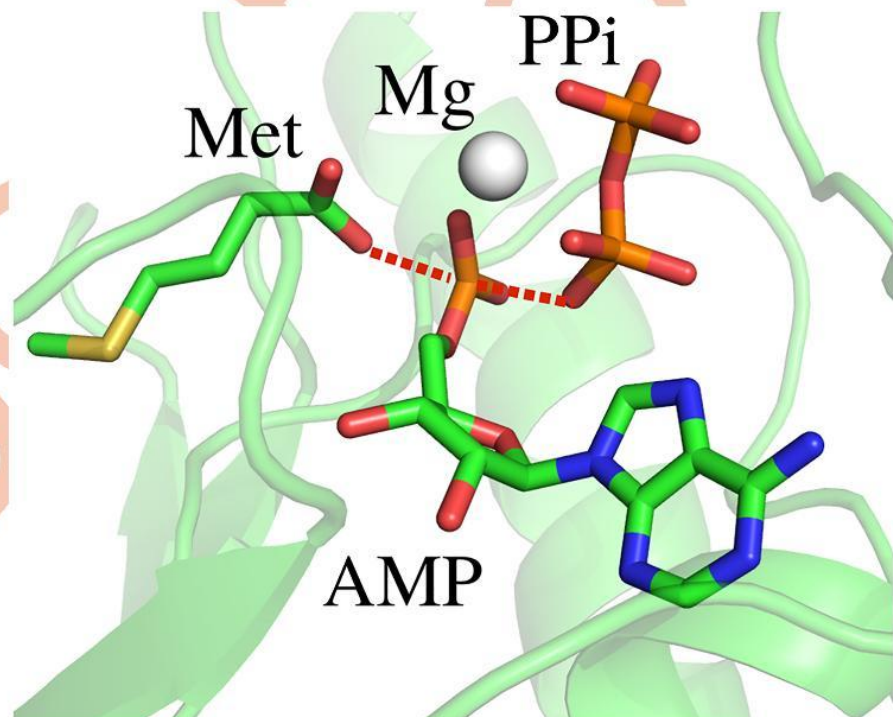


Fig 1. MetRS transition state for MetAMP formation. Closeup of the ligands.

<https://doi.org/10.1371/journal.pcbi.1007600.g001>

phosphate [44, 47–49]. *Ab initio* atomic charges were then computed for the entire activated ligand in this geometry, from a Merz-Kollman population analysis of the HF/6-31G* wavefunction [32], using Gaussian 9.0. The magnesium ion, which bridges the α , β and γ phosphates, was included in the calculation. The resulting charges were applied to atoms close to the α phosphate group, while other atoms kept their usual Met or ATP charges. Small manual adjustments were made to establish the correct total charge of -4. The final Mg charge was +1.5. Charges are in Supplementary Material (S3 File).

The geometry of the protein around the ligands was relaxed slightly by performing a short, restrained molecular dynamics simulation, with the ligands held fixed. The entire system was placed in a large box of explicit TIP3P water [50]. Harmonic restraints were applied to nonhydrogen atoms, with force constants that decreased gradually from 5 to 0.5 kcal/mol/Å² over 575 ps of dynamics, performed with the NAMD program [51]. The final protein geometry was used for the design calculations.

Monte Carlo simulations

To optimize the bias potential, we performed MC simulations of the apo state with bias updates every $T = 1000$ steps, with $e_0 = 0.2$ kcal/mol and $E_0 = 50$ kcal/mol [12]. During the first 10⁸ MC steps, we optimized a bias potential including only single-position terms. There were $p = 3$ mutating positions, which all contributed to the bias. In the second stage, we ran MC or (in one case: MetAMP complex with the FDBSA solvent model) Replica Exchange MC (REMC) simulations of 5.10⁸ MC steps [27], using 8 replicas with thermal energies (kcal/mol) of 0.17, 0.26, 0.39, 0.59, 0.88, 1.33, 2.0 and 3.0. Temperature swaps were attempted every 500 steps. All the replicas experienced the same bias potential. Both stages used 1- and 2-position moves.

Experimental mutagenesis and kinetic assays

Purification of wildtype and mutant MetRS. Throughout this study, we used a His-tagged M547 monomeric version of *E. coli* MetRS, fully active, both in vitro and in vivo [41]. The gene encoding M547 MetRS from pBSM547+ [52, 53] was subcloned into pET15bplp [54] to overproduce the His-tagged enzyme in *E. coli* (Nigro et al., submitted). Site-directed mutations were generated using the QuickChange method [55], and the whole mutated genes verified by DNA sequencing. The enzyme and its variants were produced in BLR(DE3) *E. coli* cells. Transformed cells were grown overnight at 37°C in 0.25 L of TBAl autoinducible medium containing 50 µg/ml ampicillin. They were harvested by centrifugation and resuspended in 20 ml of buffer A (10 mM Hepes-HCl pH 7.0, 3 mM 2-mercaptoethanol, 500 mM NaCl). They were disrupted by sonication (5 min, 0°C), and debris was removed by centrifugation (15,300 G, 15 min). The supernatant was applied on a Talon affinity column (10 ml; Clontech) equilibrated in buffer A. The column was washed with buffer A plus 10 mM imidazole and eluted with 125 mM imidazole in buffer A. Fractions containing tagged MetRS were pooled and diluted ten-fold in 10 mM Hepes-HCl pH 7.0, 10 mM 2-mercaptoethanol (buffer B). These solutions were applied on an ion exchange column Q Hiload (16 mL, GE-Healthcare), equilibrated in buffer B containing 50 mM NaCl. The column was washed with buffer B and eluted with a linear gradient from 5 to 500 mM NaCl in buffer B (2 ml/min, 10 mM/min). Fractions containing tagged MetRS were pooled, dialyzed against a 10 mM Hepes-HCl buffer (pH 7.0) containing 55% glycerol, and stored at -20°C. The homogeneity of the purified MetRS was estimated by SDS-PAGE to be higher than 95%.

Measurement of ATP-PPI exchange activity. Prior to activity measurements, MetRS was diluted in a 20 mM Tris-HCl buffer (pH 7.6) containing 0.2 mg/ml bovine serum albumin

(Aldrich) if the concentration after dilution was less than 1 μ M. Initial rates of ATP-PPi exchange activity were measured at 25°C as described [56]. In brief, the 100 μ l reaction mixture contained Tris-HCl (20 mM, pH 7.6), MgCl₂ (7 mM), ATP (2 mM), [³²P]PPi (1800–3700 Bq, 2 mM) and various concentrations (0–16 mM) of the Met amino acid. The exchange reaction was started by adding catalytic amounts of MetRS (20 μ l). After quenching the reaction, ³²P-labeled ATP was adsorbed on charcoal, filtered, and measured by scintillation counting. *k*_{cat} and *K*_M values were derived from iterative nonlinear fits of the theoretical equation to the experimental values using either MC-fit [57] or Origin (Origin Lab).

Results

Designing MetRS to bind azidonorleucine

As a first test, we searched for MetRS variants with strong azidonorleucine (AnL) binding. Positions 13, 260 and 301 were allowed to mutate, for comparison to the earlier experimental data [22]. 14 types were allowed at each position (see [Methods](#)), for a total of 2744 possible sequences. We compared three variants of the solvent model, which gave similar results. The first stage was to optimize a bias potential that flattened the free energy landscape in sequence space for apo MetRS. We used a bias potential including single-position terms only. After the adaptation period, we ran a further simulation of 10⁸ MC steps to determine the biased populations. With the FDBLK solvent model, 2099 sequences were visited at least 1000 times, thanks to the adaptive bias. The second stage was to simulate the MetRS:AnL complex in the presence of the bias. 1957 sequences were visited at least 1000 times in both the first and second stages. For these, we used the sampled populations to deduce the AnL binding free energy (Eq 8), relative to the X-ray sequence. The overall computation time for system setup, energy matrix precalculation and both MC stages was about one day (per solvent model). Sequences sampled with and without the bias and ligand are shown in [Fig 2](#) as logos.

Experimental directed evolution had revealed 21 active variants [22]. 13 of them were sampled by the computations and are listed in [Table 1](#). 8 others either were not sampled or were predicted to have low stability. Each variant is referred to by the sequence of the three mutating positions; for example, the X-ray variant is SLL. The top six experimental sequences are the ones that were observed in multiple clones. The others were seen in just one clone [22]. The top six were all sampled by the computations and had good predicted stabilities and affinities ([Table 1](#)). SML was ranked the highest, 17th. The five others had lower ranks, between 45 and 104, but they were all within 1.4 kcal/mol of the top predicted variant (which was HMS). Other predicted variants may also be active, even though they were not revealed by the directed evolution experiments. For the SLL variant, the predicted rotamers for binding site residues were in good agreement with the X-ray structure (Supplementary Material; [S1 File](#)). The results in [Table 1](#) were obtained with the FDBLK solvent treatment. The FDBSA solvent model gave similar results, while NEASA was slightly poorer (not shown), probably due to its simpler GB treatment [37].

We also searched for MetRS variants that maximized the AnL binding *specificity*, relative to Met. A bias potential was adaptively optimized for the MetRS:Met complex, then used in a simulation of the MetRS:AnL complex. The mutating positions and allowed types were the same as above. Specificity ranks are included in [Table 1](#). Three of the top six experimental variants had high specificity ranks. The top experimental variant NLL was 36th, the next-best experimental variant SLL was 2nd, AQL was 18th, and CLL was 3rd. Thus, among the top 40 specificity ranks, there were 4 sequences that are known to be active. Evidently, selecting for specificity can help reveal active variants.

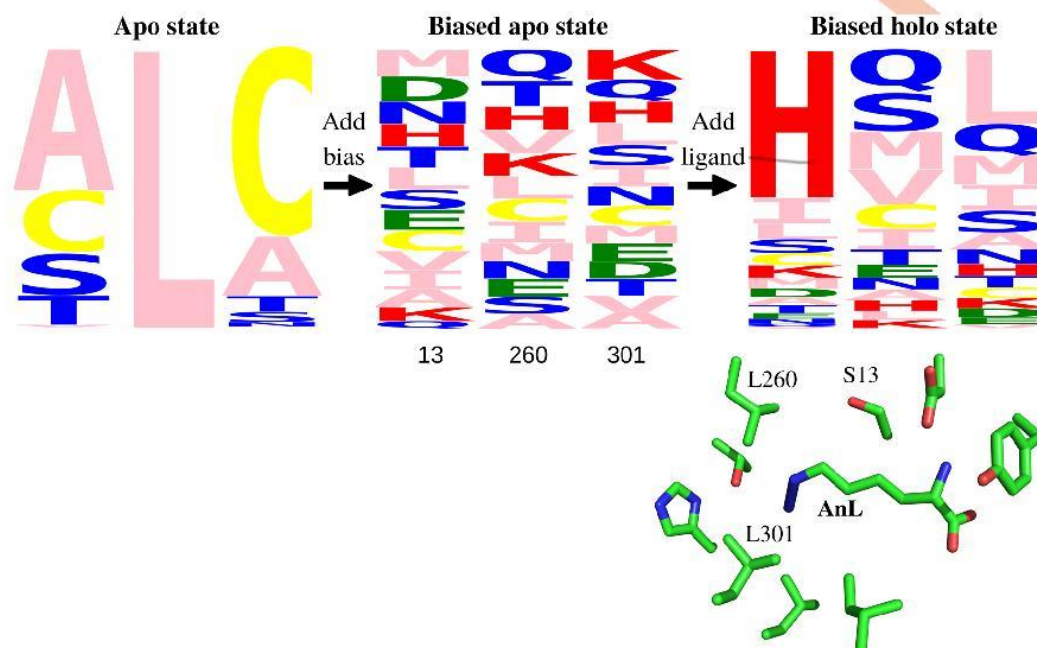


Fig 2. MetRS sequence logos. Sequences sampled without and with the AnL ligand (FDBLK solvent model) are shown in the form of logos, including the three mutating positions, 13, 260, 301. The logos represent the apo state (left), the biased apo state (middle), and the biased holo state (right). The height of each letter measures the frequency of its type. The 3D view below is a closeup of azidonorleucine (AnL) in the binding pocket, with selected side chains.

<https://doi.org/10.1371/journal.pcbi.1007600.g002>

Table 1. MetRS redesigned for AnL binding affinity or specificity.

<i>a</i> seq.	<i>b</i> pop.	<i>c</i> fold	<i>d</i> bind	<i>e</i> rank	<i>f</i> spec.	<i>g</i> rank	<i>a</i> seq.	<i>b</i> pop.	<i>c</i> fold	<i>d</i> bind	<i>f</i> rank	<i>e</i> spec.	<i>g</i> rank
NLL	62	6.7	0.3	104	5.7	36	CVL	1	6.9	-0.4	23	10.9	164
SLL	12	0.0	0.0	55	0.0	2	ACL	1	5.0	0.2	86	11.0	175
SML	4	4.6	-0.5	17	8.3	74	SCM	1	-0.9	0.4	123	18.8	589
AVL	3	6.7	-0.1	45	11.0	165	SLV	1	-2.3	1.4	688	7.4	57
AQL	2	4.2	0.0	57	3.3	18	SNL ^h	1	7.6	0.0	–	10.2	–
CLL	2	-0.6	0.1	73	1.0	3	SSL ^h	1	7.2	-0.1	–	10.3	–
							STL ^h	1	7.2	0.6	–	10.2	–

^aSequence at the designed positions 13, 260, 301, ranked by

^bpopulation among the experimental clones.

^cFolding and

^dbinding free energies (kcal/mol) relative to the X-ray sequence SLL.

^eRank based on affinity or

^fspecificity.

^gSpecificity, defined by the binding free energy difference between AnL and Met (relative to SLL).

^hNot ranked, since folding free energy is above the 7 kcal/mol threshold.

Calculations used the FDBLK solvent.

<https://doi.org/10.1371/journal.pcbi.1007600.t001>

Redesigning MetRS to bind MetAMP

As a second test, we searched for MetRS variants with a high affinity for the natural ligand methionyl adenylate (MetAMP). These should include the wildtype (WT) sequence and close homologs. Three positions close to the Met side chain (Fig 3), 13, 256, and 297 were allowed to mutate into all types except Gly or Pro, for a total of 5832 possible sequences. The first stage was to optimize a bias potential that flattened the free energy landscape in sequence space for apo MetRS. We performed calculations with both the FDBSA and the FDBLK variants of the solvent model, which gave similar results. We report the FDBSA results, since they were obtained first and were the basis for choosing which sequences to test experimentally. Selected FDBLK results are also reported. With the FDBSA solvent model, using Replica Exchange MC, 4178 variants were visited at least 1000 times.

The second stage was to simulate the MetRS:MetAMP complex in the presence of the bias potential. For sequences visited at least 1000 times in both stages (528 sequences), we used the sampled populations to deduce the MetAMP binding free energy (Eq 8), relative to the wild-type (WT) sequence LAI. The folding energy of each variant was also estimated (see Methods) and sequences less stable than WT by 5 kcal/mol or more were discarded. The top 20 remaining sequences, with the largest binding free energies, are shown in Table 2. The top sequence, CDV, had an Asp at position 256, positioned to form a salt bridge with the MetAMP ammonium group. Its binding free energy, relative to WT, was -1.4 kcal/mol. The next 19 variants had types similar to WT. Their computed binding free energies were close to WT, with relative values between -0.2 and 0.6 kcal/mol. The WT sequence was sixth overall. Among the top 40 variants, 17 mutants were produced experimentally. They were representative of the computational variants, while providing ease of construction (see Methods). CDV was left out, as the A256D mutation, selected for binding, might reduce the catalytic activity. All 17 tested variants had detectable activity, a 100% success rate for the design procedure. One other sequence, SAI, was tested experimentally and found to be active, but did not show up in the MC simulation. Thus the method produced one false negative along with 17 true positives.

Going further, we made a quantitative comparison between the computed and experimental binding free energies. The experimental dissociation constants were estimated from the Michaelis constants K_M . In the experimental conditions (excess ATP) and under the usual Michaelis-Menten assumptions [14, 58], K_M represents the dissociation constant for Met binding in the presence of bound ATP. Here, we computed relative binding free energies for binding MetAMP, not Met. Nevertheless, we expect that these MetAMP binding free energy changes can be compared to the experimental Met binding free energy changes; *i.e.*, we make the additional assumption that the relative effects of the mutations will be conserved going from MetAMP to Met+ATP.

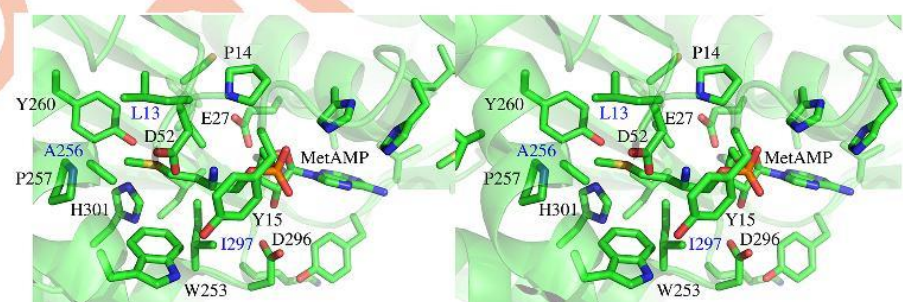


Fig 3. MetRS:MetAMP complex. Binding site closeup (stereo). Mutating side chains are 13, 256, 297.

<https://doi.org/10.1371/journal.pcbi.1007600.g003>

Table 2. MetRS redesigned for MetAMP binding by mutating positions 13, 256, 297.

rank	variant	^a folding	binding		rank	variant	^a folding	binding	
			^b comp.	^b exp.				^b comp.	^b exp.
1	CDV	4.5	-1.36		11	LAC	-0.3	0.25	1.8
2	MAV	1.3	-0.23	1.8	12	MAT	4.6	0.28	2.4
3	MAI	2.5	-0.20		13	LSV	0.4	0.29	
4	LAV	-1.3	-0.16	1.8	14	LAA	-0.6	0.31	3.8
5	MAC	2.3	-0.09	2.3	15	CAV	-8.8	0.34	2.8
6	LAI	0.0	0.00	0.0	16	CAI	-7.4	0.37	1.2
7	MAA	2.0	0.02		17	MSC	4.1	0.45	
8	MSV	3.1	0.11	3.4	18	MCV	1.0	0.46	
9	MSI	4.4	0.15	2.2	19	MCI	2.3	0.48	
10	LSI	1.6	0.20		20	MSA	3.8	0.56	
					21	LAT	1.8	0.59	2.2
26	CAC	-7.9	0.69	3.0	28	SAI	-3.5	0.72	1.2
51	SAC	-4.0	1.11	3.0	68	LAS	1.3	1.34	3.4
70	SSI	-1.9	1.35	2.2	81	SSC	-2.2	1.45	3.6
	MST	6.2	0.98	3.5		MSS	5.8	1.64	3.4

Calculations with the FDBSA solvent model.

^aFolding and

^bMetAMP binding free energies (kcal/mol) from computations and experiment, relative to the WT sequence LAI.

<https://doi.org/10.1371/journal.pcbi.1007600.t002>

Certain mutations at position 297 involved significant changes in the side chain volume, where the largest type, Ile or the smallest type, Ala was introduced or removed. For these, the computed binding free energies departed significantly from the experimental ones. However, if these two types were excluded, there were 25 point mutations between experimental variants, and for these, agreement was very good. The computed binding free energy differences had an rms error of just 0.52 kcal/mol and a mean unsigned error (mue) of 0.43 kcal/mol. The correlation between the experimental and computed sets was 0.52. Fig 4 shows the binding free energy changes. Note that the good agreement supports the assumption that the experimental K_M values are good proxies for the relative MetAMP binding free energies.

With the FDBLK solvent model, results were similar. The WT variant was ranked slightly lower, 20th. The top sequence was SAN, with a binding free energy of -1.3 kcal/mol relative to the WT. 7 of the 17 experimental sequences were ranked among the top 20 predictions. The computed and experimental binding free energy changes associated with point mutations are shown in Supplementary Material (Figure B in S1 File). Excluding (as above) mutations involving the types Ile or Ala at position 297, the mue and rms error were 0.76 and 0.98 kcal/mol, respectively, only slightly larger than with FDBSA.

Redesigning MetRS for catalytic power

For enzyme design, it is of great interest to select for a low activation free energy [14]. Therefore, we considered a model of the transition state complex (Fig 1). The ATP α phosphorus was bound to five oxygens: three coplanar and two perpendicular, corresponding to in-line attack of the Met carboxylate. In the first stage, we simulated a competing, ground state complex between MetRS, Met, and ATP. The same three binding pocket residues as above, 13, 256, and 297 were allowed to mutate into all types except Gly, Pro. We used the FDBLK solvent model. We optimized a bias potential during the MC simulation, flattening the free energy

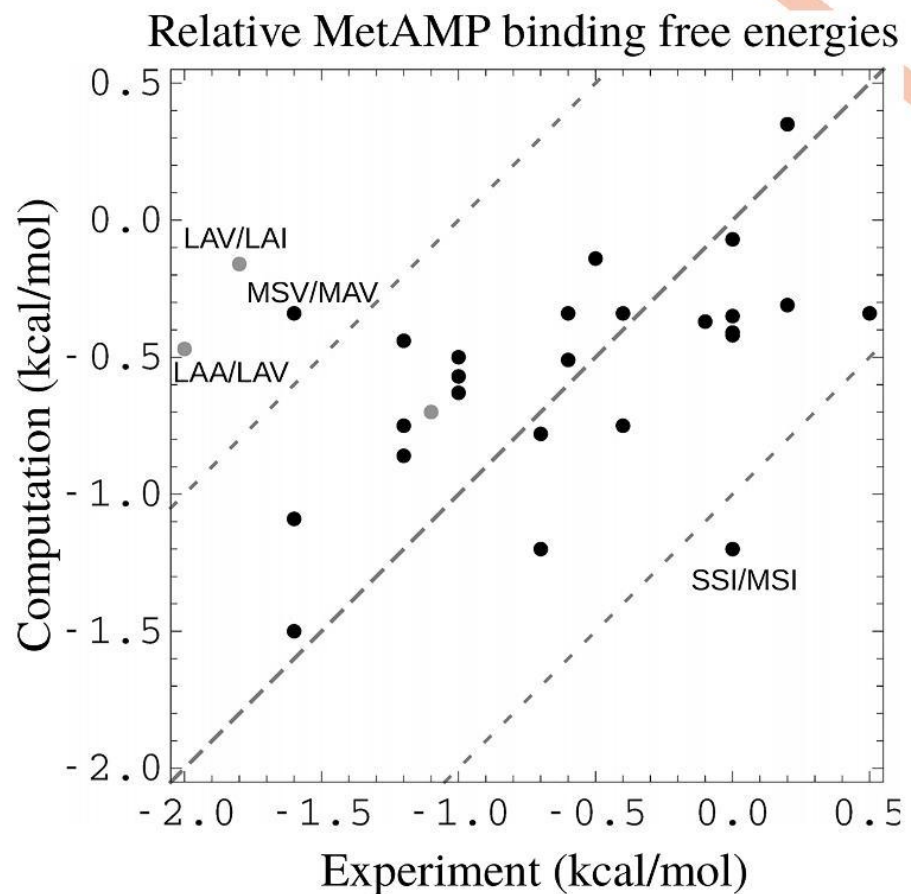


Fig 4. MetRS:MetAMP binding free energies, relative to the wildtype protein (WT). Shown are data for 28 point mutations. 3 gray points correspond to two mutations at position 297 (labeled) that change the side chain volume, plus one involving a variant (MST) that was predicted to be weakly stable (above our 5 kcal/mol threshold, see text) but was produced and measured experimentally nevertheless. Two other mutations with sizable errors are labeled.

<https://doi.org/10.1371/journal.pcbi.1007600.g004>

surface in sequence space. In the second stage, we simulated the transition state complex, with the bias included. All the variants that had been tested experimentally (Table 2) were sampled (WT and 19 variants, including five that Proteus had predicted (with FDBLK) to be above our 5 kcal/mol instability threshold). For each one, from the sampled populations, we deduced the free energy difference (Eq 8) between its ground state and transition state complexes, *i.e.*, its activation free energy. From transition state theory [14], this difference can be identified with the log of the catalytic reaction rate, k_{cat} . We also computed the Met dissociation free energies for the ground state complexes, which can be identified with the Met Michaelis constants, K_M . We first simulated the ground state complex with ATP but no Met, flattening its free energy surface with an adaptive bias. We then simulated the MetRS+Met+ATP complex, including the bias. From the sampled populations, we deduced the Met binding free energy of each variant, relative to WT (Eq 8). The overall protocol is schematized in Fig 5.

Fig 6 compares the k_{cat}/K_M ratios from experiment and simulations. We refer to them as catalytic efficiencies. We recall that they represent the 2nd order rate constant for the reaction of Met with the MetRS:ATP complex. Fig 6 shows the quantities $kT \log (k_{\text{cat}}/K_M) / (k_{\text{cat}}/K_M)_{\text{WT}}$.

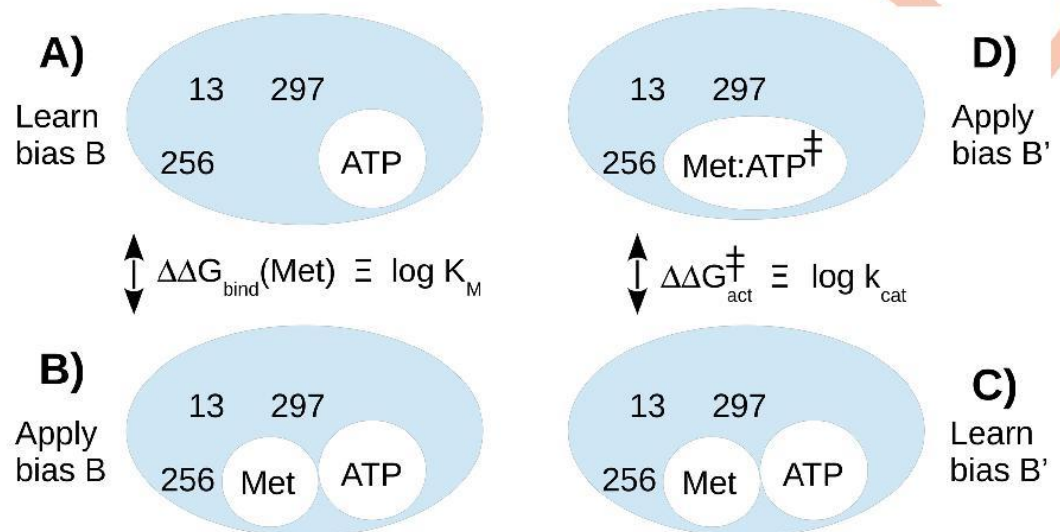


Fig 5. Computational scheme used to obtain the catalytic efficiencies k_{cat}/K_M . A) A bias B is optimized to flatten the sequence landscape of the enzyme without the Met ligand. Mutating positions are 13, 256, 297. B) The same bias B is used to simulate the complex including Met. Sequences are populated according to their Met binding affinities. C) A bias B' is optimized to flatten the sequence landscape of the complex including Met. D) B' is used to simulate the transition state complex. Sequences are populated according to their activation free energies. The lefthand simulations yield the predicted K_M values. The righthand simulations yield the predicted k_{cat} values.

<https://doi.org/10.1371/journal.pcbi.1007600.g005>

which express the catalytic efficiencies on a log scale, in thermal energy units, relative to the WT value. The figure includes WT and 19 other experimental variants. 5 of these had low predicted stabilities and are shown as gray points. WT defines the origin. For the other 14 points, agreement between calculations and experiment is quite good, with a correlation of 0.73 and mean errors of 1.36 kcal/mol (rms) and 1.18 kcal/mol (mue). Experimentally, WT has the largest efficiency. Computationally, two variants are predicted to be slightly better, by 0.9 and 0.2 kcal/mol, respectively, which is less than the mean error. Overall, by designing directly for a low activation free energy, we retrieve all the experimental variants and reproduce the catalytic efficiencies semi-quantitatively.

Discussion

Adaptive importance sampling solves the design problem for ligand binding and specificity. It applies positive design to one state (say, bound) and negative design to the other (unbound). It provides quantitative values for relative binding free energies or activation free energies. Variants sampled for one criterion, such as activation free energy (k_{cat}), can be reranked *a posteriori* based on another criterion, such as k_{cat}/K_M . A *a posteriori* reranking or filtering does not leave out any important solutions; rather, the initial selection brings in too many solutions (e.g., unstable variants), which are then filtered out at very little cost. In the first stage of the procedure, the sampling is very aggressive, if not exhaustive. In the second stage, it does not need to be exhaustive, since the best designs are exponentially enriched, and the unsampled variants are the ones with poor affinities or specificities. If one wants to reveal weak binders or perform reranking on another property, one can also flatten the energy landscape in the second stage. One can also use a more aggressive bias in one or both stages, including two-position biases. Replica Exchange MC can also be used to increase sampling. Using plain MC, one-position

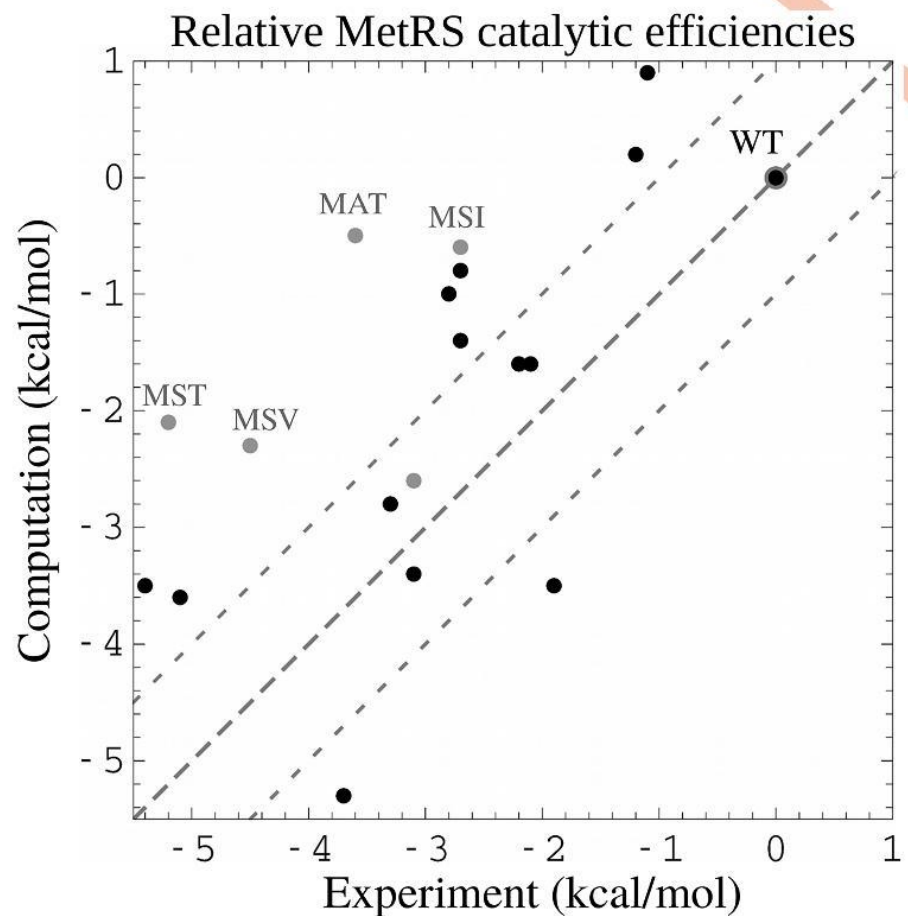


Fig 6. MetRS catalytic efficiencies $kT \log (k_{cat}/K_M) / (k_{cat}/K_M)_{WT}$ relative to the wildtype (kcal/mol). Four gray points correspond to variants that were predicted to be weakly stable but were produced and measured experimentally nevertheless. Results obtained with the FDBLK solvent model.

<https://doi.org/10.1371/journal.pcbi.1007600.g006>

biases and no flattening of the holo state, our simulations produced 200 MetRS variants, enriched in tight binders, spanning a 7–8 kcal/mol range of binding free energies.

A difficulty when designing ligand binding is to choose one or more poses for the ligand. Here, we redesigned MetRS in cases where the ligand pose was known from an X-ray structure for one sequence: the SLL sequence in the AnL case and the wildtype sequence in the MetAMP case. For these ligands, we used the experimental ligand pose and protein backbone conformation. Three residues close to the ligand were then allowed to mutate. Not surprisingly, the calculations produced designed sequences that were homologous to the X-ray sequence. The experimental binding free energies in the MetAMP case were well-reproduced (the AnL values are not known). It is likely that other poses exist that would be compatible with other mutations, and would possibly lead to even stronger binding. The exploration of such alternate poses was left aside in this work. For the transition state complex, the position of the ligands could also be inferred with some confidence, since the enzyme achieves catalysis with little reorganization or motion of the substrates [15], and the modeled transition state geometry of the α phosphate was intermediate between that seen in two Met RS X-ray structures: the

MetRS product (adenylate) complex and the reactant (ATP) complex (PDB 4QRE). By designing the protein to stabilize this ligand pose, we may have biased the results towards native-like solutions. Here, too, the experimental relative activation free energies were well-reproduced, supporting the structural model.

Another important model component is the implicit solvent model. Here, we used a carefully-parameterized Generalized Born variant [33], a physically-plausible value of the protein dielectric constant and an “FDB” computational scheme that maintains the many-body nature of the GB model. The simpler, NEA scheme gave somewhat poorer results, similar to another recent study [35]. For nonpolar contributions to solvation, we compared a Surface Area (SA) treatment and a Lazaridis-Karplus (LK) treatment, which gave similar results. In the reported calculations, no water molecules were modelled explicitly. We also tested a model where three waters in the MetRS active site were explicitly represented: those that directly coordinate the Mg^{2+} ion in the substrate and transition state complexes for MetAMP formation. With both the FDBSA and FDBLK treatments, their explicit representation led to k_{cat} values well within the mean error of the calculations (relative to experiment). Most $kT \log(k_{cat}/K_M) / (k_{cat}/K_M)_{WT}$ values were with 0.2–0.3 kcal/mol of those reported above. Overall, the results were reasonably robust with respect to model details, with FDB giving improved performance.

Agreement with experiment was very good for three MetRS redesign test problems: redesign to bind the AnL ncAA, redesign to bind the natural intermediate MetAMP, and redesign for catalytic power for the reaction that produces MetAMP. Except for the earlier AnL data [22], the experiments were done in this work. Transition state modeling was done simply, by combining two X-ray structures and running a standard quantum chemistry protocol for atomic charges, consistent with the usual Amber force field [32]. All the procedures were carried out with the Proteus software, which is freely available to academics (<https://proteus.polytechnique.fr>). An entire calculation (setup, matrix calculation, MC simulations, postprocessing) lasted around one day on a 16-core desktop computer. We expect the present adaptive MC method will become the paradigm for computational enzyme design in the future.

Supporting information

S1 File. Supplementary appendix. This file includes a short theoretical derivation, some explanation of force field parameters, atomic charges for the MetRS transition state ligands, a figure showing the MetRS:AnL complex structure, and MetRS:MetAMP binding free energy results obtained with the FDBLK solvent model.
(PDF)

S2 File. Azidonorleucine force field information. This file contains the “topology” or 2D structure of AnL, including the atomic charges, followed by energy parameters for covalent bonds, angles, dihedrals, impropers, van der Waals terms and Generalized Born.
(FF)

S3 File. MetAMP force field information. This file contains the “topology” or 2D structure of MetAMP, including the atomic charges, followed by energy parameters for covalent bonds, angles, dihedrals, impropers, van der Waals terms and Generalized Born.
(FF)

Acknowledgments

We thank Christine Lazennec-Schurdevin for technical assistance, Alexandrine Daniel for preliminary MetRS:AnL calculations and Francesco Villa and David Mignon for many helpful discussions.

Author Contributions

Conceptualization: Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Data curation: Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Formal analysis: Vaitea Opuu, Thomas Simonson.

Investigation: Vaitea Opuu, Giuliano Nigro, Thomas Gaillard, Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Methodology: Vaitea Opuu, Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Project administration: Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Software: Vaitea Opuu, Thomas Simonson.

Supervision: Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

Writing – original draft: Thomas Simonson.

Writing – review & editing: Emmanuelle Schmitt, Yves Mechulam, Thomas Simonson.

References

1. Malisi C, Schumann M, Toussaint NC, Kageyama J, Kohlbacher O, Höcker B. Binding Pocket Optimization by Computational Protein Design. *PLoS One*. 2012; 7:e52505. <https://doi.org/10.1371/journal.pone.0052505> PMID: 23300688
2. Feldmeier K, Hoecker B. Computational protein design of ligand binding and catalysis. *Curr Opin Chem Biol*. 2013; 17:929–933. <https://doi.org/10.1016/j.cbpa.2013.10.002> PMID: 24466576
3. Tinberg CE, Khare SD, Dou J, Doyle L, Nelson JW, Schena A, et al. Computational design of ligand-binding proteins with high affinity and selectivity. *Nature*. 2013; 501:212–218. <https://doi.org/10.1038/nature12443> PMID: 24005320
4. Stoddard B, editor. *Methods in Molecular Biology: Design and Creation of Ligand Binding Proteins*. Springer Verlag, New York; 2016.
5. Samish I, MacDermaid CM, Perez-Aguilar JM, Saven JG. Theoretical and computational protein design. *Ann Rev Phys Chem*. 2011; 62:129–149. <https://doi.org/10.1146/annurev-physchem-032210-103509>
6. Simonson T, Ye-Lehmann S, Palmi Z, Amara N, Bigan E, Wydau S, et al. Redesigning the stereospecificity of tyrosyl-tRNA synthetase. *Proteins*. 2016; 84:240–253. <https://doi.org/10.1002/prot.24972> PMID: 26676967
7. Shen Q, Tian H, Tang D, Yao W, Gao X. Ligand-K* sequence elimination: a novel algorithm for ensemble-based redesign of receptor-ligand binding. *Trans Comp Biol Bioinf*. 2014; 11:573–578. <https://doi.org/10.1109/TCBB.2014.2302795>
8. Viricel C, Simoncini D, Allouche D, de Givry S, Barbe S, Schiex T. Approximate counting with deterministic guarantees for affinity computation. In: LeThi HA, Dinh TP, Nguyen NT, editors. *Adv. Intell. Syst. Comput.* vol. 360. Springer, New York; 2015. p. 165–176.
9. Hallen MA, Donald BR. COMETS (Constrained Optimization of Multistate Energies by Tree Search): A provable and efficient protein design algorithm to optimize binding affinity and specificity with respect to sequence. *J Comp Biol*. 2016; 23:311–321. <https://doi.org/10.1089/cmb.2015.0188>
10. Karimi M, Shen Y. iCFN: an efficient exact algorithm for multistate protein design. *Bioinf*. 2018; 34:i811–820. <https://doi.org/10.1093/bioinformatics/bty564>
11. Bhattacharjee A, Wallin S. Exploring protein-peptide binding specificity through computational peptide screening. *PLoS Comp Biol*. 2013; 7:e1003277. <https://doi.org/10.1371/journal.pcbi.1003277>
12. Villa F, Panel N, Chen X, Simonson T. Adaptive landscape flattening in amino acid sequence space for the computational design of protein:peptide binding. *J Chem Phys*. 2018; 149:072302. <https://doi.org/10.1063/1.5022249> PMID: 30134674
13. Hayes RL, Armacost KA, Vilseck JZ, Brooks CL III. Adaptive landscape flattening accelerates sampling of alchemical space in multisite lambda dynamics. *J Phys Chem B*. 2017; 121:3626–3635. <https://doi.org/10.1021/acs.jpcc.6b09656> PMID: 28112940
14. Jencks WP. *Catalysis in chemistry and enzymology*. Dover, New York; 1986.

15. Ibba M, Francklyn C, Cusack S, editors. Aminoacyl-tRNA Synthetases. Landes Bioscience, Georgetown; 2005.
16. Xie J, Schultz PG. A chemical toolkit for proteins: an expanded genetic code. *Nat Rev Molec Cell Biol*. 2006; 7:775–782. <https://doi.org/10.1038/nrm2005>
17. Young TS, Schultz PG. Beyond the canonical twenty amino acids: expanding the genetic lexicon. *J Biol Chem*. 2010; 285:11039–11044. <https://doi.org/10.1074/jbc.R109.091306> PMID: 20147747
18. Liu CC, Schultz PG. Adding new chemistries to the genetic code. *Ann Rev Biochem*. 2010; 79:413–444. <https://doi.org/10.1146/annurev.biochem.052308.105824> PMID: 20307192
19. Neumann-Staubitz P, Neumann H. The use of unnatural amino acids to study and engineer protein function. *Curr Opin Struct Biol*. 2016; 38:119–128. <https://doi.org/10.1016/j.sbi.2016.06.006> PMID: 27318816
20. Chin JW. Expanding and reprogramming the genetic code. *Nature*. 2017; 550:53–60. <https://doi.org/10.1038/nature24031> PMID: 28980641
21. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of *Escherichia coli*. *Science*. 2001; 292:498–500. <https://doi.org/10.1126/science.1060077> PMID: 11313494
22. Tanrikulu IC, Schmitt E, Mechulam Y, Goddard W III, Tirrell DA. Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *Proc Natl Acad Sci USA*. 2009; 106:15285–15290. <https://doi.org/10.1073/pnas.0905735106> PMID: 19706454
23. Wang FG, Landau DP. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett*. 2001; 86:2050–2053. <https://doi.org/10.1103/PhysRevLett.86.2050> PMID: 11289852
24. Laio A, Gervasio F. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Prog Phys*. 2008; 71:art. 126601. <https://doi.org/10.1088/0034-4885/71/12/126601>
25. Frenkel D, Smit B. Understanding molecular simulation, Chapter 3. Academic Press, New York; 1996.
26. Grimmett GR, Stirzaker DR. Probability and random processes. Oxford University Press, Oxford, United Kingdom; 2001.
27. Mignon D, Simonson T. Comparing three stochastic search algorithms for computational protein design: Monte Carlo, Replica Exchange Monte Carlo, and a multistart, steepest-descent heuristic. *J Comput Chem*. 2016; 37:1781–1793. <https://doi.org/10.1002/jcc.24393> PMID: 27197555
28. Barducci A, Bussi G, Parrinello M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett*. 2008; 100:art. 020603. <https://doi.org/10.1103/PhysRevLett.100.020603> PMID: 18232845
29. Dama JF, Parrinello M, Voth GA. Well-tempered metadynamics converges asymptotically. *Phys Rev Lett*. 2014; 112:art. 240602. <https://doi.org/10.1103/PhysRevLett.112.240602> PMID: 24996077
30. Simonson T, Gaillard T, Mignon D, Schmidt am Busch M, Lopes A, Amara N, et al. Computational protein design: the Proteus software and selected applications. *J Comput Chem*. 2013; 34:2472–2484. <https://doi.org/10.1002/jcc.23418> PMID: 24037756
31. Simonson T. The Proteus software for computational protein design. Ecole Polytechnique, Paris: <https://proteus.polytechnique.fr>; 2019.
32. Cornell W, Cieplak P, Bayly C, Gould I, Merz K, Ferguson D, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc*. 1995; 117:5179–5197. <https://doi.org/10.1021/ja00124a002>
33. Lopes A, Aleksandrov A, Bathelt C, Archontis G, Simonson T. Computational sidechain placement and protein mutagenesis with implicit solvent models. *Proteins*. 2007; 67:853–867. <https://doi.org/10.1002/prot.21379> PMID: 17348031
34. Gaillard T, Simonson T. Pairwise decomposition of an MMGBSA energy function for computational protein design. *J Comput Chem*. 2014; 35:1371–1387. <https://doi.org/10.1002/jcc.23637> PMID: 24854675
35. Michael E, Polydorides S, Simonson T, Archontis G. Simple models for nonpolar solvation: parametrization and testing. *J Comput Chem*. 2017; 38:2509–2519. <https://doi.org/10.1002/jcc.24910> PMID: 28786118
36. Polydorides S, Simonson T. Monte Carlo simulations of proteins at constant pH with generalized Born solvent, flexible sidechains, and an effective dielectric boundary. *J Comput Chem*. 2013; 34:2742–2756. <https://doi.org/10.1002/jcc.23450> PMID: 24122878
37. Villa F, Mignon D, Polydorides S, Simonson T. Comparing pairwise-additive and many-body Generalized Born models for acid/base calculations and protein design. *J Comput Chem*. 2017; 38:2396–2410. <https://doi.org/10.1002/jcc.24898> PMID: 28749575
38. Polydorides S, Amara N, Aubard C, Plateau P, Simonson T, Archontis G. Computational protein design with a generalized Born solvent model: application to Asparaginyl-tRNA synthetase. *Proteins*. 2011; 79:3448–3468. <https://doi.org/10.1002/prot.23042> PMID: 21563215

39. Archontis G, Simonson T. Proton binding to proteins: a free energy component analysis using a dielectric continuum model. *Biophys J*. 2005; 88:3888–3904. <https://doi.org/10.1529/biophysj.104.055996> PMID: 15821163
40. Dahiyat BI, Mayo SL. De novo protein design: fully automated sequence selection. *Science*. 1997; 278:82–87. <https://doi.org/10.1126/science.278.5335.82> PMID: 9311930
41. Schmitt E, Tanrikulu IC, Yoo TH, Panvert M, Tirrell DA, Mechulam Y. Switching from an Induced-Fit to a Lock-and-Key Mechanism in an Aminoacyl-tRNA Synthetase with Modified Specificity. *J Mol Biol*. 2009; 394:843–851. <https://doi.org/10.1016/j.jmb.2009.10.016> PMID: 19837083
42. Tuffery P, Etchebest C, Hazout S, Lavery R. A new approach to the rapid determination of protein side chain conformations. *J Biomol Struct Dyn*. 1991; 8:1267–1289. <https://doi.org/10.1080/07391102.1991.10507882> PMID: 1892586
43. Gaillard T, Panel N, Simonson T. Protein sidechain conformation predictions with an MMGBSA energy function. *Proteins*. 2016; 84:803–819. <https://doi.org/10.1002/prot.25030> PMID: 26948696
44. Crépin T, Schmitt E, Mechulam Y, Sampson PB, Vaughan MD, Honek JF, et al. Use of analogues of methionine and methionyl adenylate to sample conformational changes during catalysis in *Escherichia coli* methionyl-tRNA synthetase. *J Mol Biol*. 2003; 332:59–72. [https://doi.org/10.1016/s0022-2836\(03\)00917-3](https://doi.org/10.1016/s0022-2836(03)00917-3) PMID: 12946347
45. Pokala N, Handel TM. Energy functions for protein design: adjustment with protein-protein complex affinities, models for the unfolded state, and negative design of solubility and specificity. *J Mol Biol*. 2005; 347:203–227. <https://doi.org/10.1016/j.jmb.2004.12.019> PMID: 15733929
46. Druart K, Palmai Z, Omarjee E, Simonson T. Protein:ligand binding free energies: a stringent test for computational protein design. *J Comput Chem*. 2016; 37:404–415. <https://doi.org/10.1002/jcc.24230> PMID: 26503829
47. Arnez JG, Augustine JG, Moras D, Francklyn CS. The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proc Natl Acad Sci USA*. 1997; 94:7144–7149. <https://doi.org/10.1073/pnas.94.14.7144> PMID: 9207058
48. Zurek J, Bowman A, Sokalski W, Mulholland A. MM and QM/MM modeling of threonyl-tRNA synthetase: Model testing and simulations. *Struct Chem*. 2004; 15:405–414. <https://doi.org/10.1023/B:STUC.0000037896.80027.2c>
49. Banik S, Nandi N. Aminoacylation Reaction in the Histidyl-tRNA Synthetase: Fidelity Mechanism of the Activation Step. *J Phys Chem B*. 2010; 114:12301–2311. <https://doi.org/10.1021/jp910730s>
50. Jorgensen WL, Chandrasekar J, Madura J, Impey R, Klein M. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983; 79:926–935. <https://doi.org/10.1063/1.445869>
51. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26:1781–1802. <https://doi.org/10.1002/jcc.20289> PMID: 16222654
52. Mellot P, Mechulam Y, Le Corre D, Blanquet S, Fayat G. Identification of an amino acid region supporting specific methionyl-tRNA synthetase:tRNA recognition. *J Mol Biol*. 1989; 208:429–443. [https://doi.org/10.1016/0022-2836\(89\)90507-x](https://doi.org/10.1016/0022-2836(89)90507-x) PMID: 2477552
53. Schmitt E, Meinel T, Panvert M, Mechulam Y, Blanquet S. Two acidic residues of *Escherichia coli* methionyl-tRNA synthetase act as negative discriminants towards the binding of noncognate tRNA anticodons. *J Mol Biol*. 1993; 233:615–628. <https://doi.org/10.1006/jmbi.1993.1540> PMID: 8411169
54. Guillon L, Schmitt E, Blanquet S, Mechulam Y. Initiator tRNA binding by eIF5B, the eukaryotic/archaeal homologue of bacterial Initiation Factor IF2. *Biochemistry*. 2005; 44:15594–15601. <https://doi.org/10.1021/bi051514j> PMID: 16300409
55. Braman J, Papworth C, A G. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Molec Biol*. 1996; 57:31–44.
56. Schmitt E, Meinel T, Blanquet S, Mechulam Y. Methionyl-tRNA synthetase needs an intact and mobile KMSKS motif in catalysis of methionyl adenylate formation. *J Mol Biol*. 1994; 242:566–577. <https://doi.org/10.1006/jmbi.1994.1601> PMID: 7932711
57. Dardel F. *Comp App Biosci*. 1994; 10:273–275.
58. Thompson D, Plateau P, Simonson T. Free energy simulations reveal long-range electrostatic interactions and substrate-assisted specificity in an aminoacyl-tRNA synthetase. *ChemBioChem*. 2006; 7:337–344. <https://doi.org/10.1002/cbic.200500364> PMID: 16408313

Titre : Incorporation de β -méthionine dans des polypeptides

Mots clés : Méthionyl ARNt synthetase, acides aminés β , traduction, acides aminés non standard

Les cellules vivantes utilisent 20 acides α aminés canoniques lors de la synthèse ribosomale des protéines, même si dans de rares occasions, d'autres acides aminés (selenocysteine ou pyrrolysine) peuvent être utilisés. Le répertoire des acides aminés utilisables est donc assez restreint. Cela limite la possibilité de construire des protéines ayant de nouvelles propriétés. Une tendance forte dans le domaine de l'ingénierie des protéines est la construction de systèmes permettant d'incorporer des acides aminés non canoniques pendant la biosynthèse *in vivo*. Ces travaux ont connu d'importants succès, mais tous les acides aminés incorporés jusqu'aux années 2010 étaient des acides aminés α . L'incorporation d'acides aminés β à des sites spécifiques créerait une flexibilité supérieure à celle des acides aminés α dans la chaîne principale, ce qui augmenterait les possibilités de repliement de la protéine. Il a été montré récemment en utilisant un système de traduction *in vitro* de synthèse protéique qu'il était possible d'incorporer des acides aminés β à des positions spécifiques (Katoh and Suga, 2018). Pour parvenir à transposer ce système *in vivo*, la principale limitation est l'aminoacylation des ARNt avec les acides aminés β . Il a été récemment proposé que certaines aminoacyl-ARNt synthetases étaient capables d'utiliser les acides aminés β , mais cette capacité reste limitée.

Le but de cette thèse est d'incorporer la β -méthionine dans des polypeptides *in vivo*. Pour cela nous avons caractérisé la reconnaissance et l'utilisation de la L- β -homométhionine par la méthionyl-ARNt synthetase (MetRS) d'*E. coli*. Nous avons en particulier déterminé une structure cristallographique à haute résolution du complexe MetRS: β -Met. En utilisant la spectroscopie de fluorescence ainsi que la spectroscopie de masse, nous avons pu mettre en évidence l'activation de la β -Met en adénylate ainsi que son estérification à un ARNt^{Met}. Toutefois, les efficacités mesurées sont très petites par rapport à ce qui avait été publié. Une contamination de la β -Met commerciale par de la méthionine pourrait expliquer ces différences. Une étude *in silico* basée sur la structure du complexe MetRS: β -Met a été menée en collaboration avec l'équipe de Bio-informatique du laboratoire afin de rechercher des enzymes mutantes plus efficaces vis à vis des acides aminés β . Enfin, nous avons amorcé la mise en place d'une méthode d'évolution dirigée destinée à améliorer l'efficacité d'incorporation des acides aminés β *in vivo*.

Title : Incorporation of β -methionine in polypeptides

Keywords : Méthionyl tRNA synthetase, β amino acids, translation, non-standard amino acid

Living cells use 20 canonical α amino acids during ribosomal protein synthesis, although in rare cases, other amino acids such as selenocysteine or pyrrolysine can be used. The repertoire of usable amino acids is therefore quite limited. This limits the ability to build proteins having new properties. A dominant trend in the field of protein engineering is the construction of systems capable of incorporating non-canonical amino acids during *in vivo* protein synthesis. These studies have been very successful, but all amino acids incorporated until the 2010s were α -amino acids. Incorporation of β -amino acids at discrete sites would create unprecedented flexibility in the main chain, which would increase the potential for new protein folds. It has recently been shown using an *in vitro* protein synthesis system that it is possible to incorporate β -amino acids at specific positions (Katoh and Suga, 2018). In order to transpose this system *in vivo*, the main limitation is the aminoacylation of tRNAs with β amino acids. It has recently been proposed that some aminoacyl-tRNA synthetases can use β -amino acids, but this capacity remains limited.

The aim of this PhD is to incorporate β -methionine in polypeptides *in vivo*. For this purpose, we have characterized the recognition and use of L- β -homomethionine by methionyl-tRNA synthetase (MetRS) from *E. coli*. In particular, we have determined a high-resolution crystallographic structure of the MetRS: β -Met complex. Using fluorescence spectroscopy and mass spectroscopy, we were able to demonstrate the activation of β -Met into adenylate as well as its esterification onto tRNA^{Met}. However, the measured efficiencies are very small compared to what was published. Contamination of commercial β -Met with methionine may explain these differences. An *in silico* study based on the structure of the MetRS: β -Met complex was conducted in collaboration with the laboratory's bioinformatics team in order to search for mutant enzymes more efficient in the use of β -amino acids. Finally, we initiated the implementation of a directed evolution method to improve the efficiency of incorporation of amino acids *in vivo*.