La méthylation flavine-dépendante d'acides nucléiques : aspects évolutifs, métaboliques, biochimiques et spectroscopiques

Pierre Sournia

▶ **To cite this version:**

Pierre Sournia. La méthylation flavine-dépendante d'acides nucléiques : aspects évolutifs, métaboliques, biochimiques et spectroscopiques. Biochimie, Biologie Moléculaire. Université Paris Saclay (COmUE), 2016. Français. ⟨NNT : 2016SACLX108⟩. ⟨tel-01508953⟩

HAL Id: tel-01508953

https://pastel.hal.science/tel-01508953v1

Submitted on 15 Apr 2017

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.

HAL Authorization

NNT : 2016SACLX108

Thèse de doctorat de l'Université Paris-Saclay préparée à l'École Polytechnique

Ecole doctorale n°573
Interface
Spécialité : Biochimie et biologie moléculaire

par

M. Pierre SOURNIA

La méthylation flavine-dépendante d'acides nucléiques : aspects évolutifs, métaboliques, biochimiques et spectroscopiques

Thèse présentée et soutenue à l'École polytechnique, le 14 décembre 2016.

Composition du Jury :

Prof.	Chahrazade EL AMRI	UPMC	(Présidente du jury)
Prof.	Jaunius URBONAVIČIUS	VGTU	(Rapporteur)
Dr.	Laura BACIOU	Université Paris Sud	(Rapportrice)
Prof.	Anne WOISARD	UPMC	(Examinatrice)
Dr.	Ursula LIEBL	École Polytechnique	(Directrice de thèse)
Dr.	Hannu MYLLYKALLIO	École Polytechnique	(Co-Directeur de thèse)

Remerciements

J'aimerai, en premier lieu, remercier messieurs les professeurs Jean Louis Martin et Francois Hache, directeurs successifs du LOB pour m'avoir accueilli (et gardé) au sein de ce laboratoire.

Je tiens à remercier Ursula Liebl et Hannu Myllykallio pour m'avoir donné l'opportunité de réaliser ce travail. Leur méthode de management m'a permis d'acquérir une certaine autonomie et indépendance d'esprit.

Je souhaite exprimer toute ma gratitude envers les rapporteurs, Jaunius Urbonavičius et Laura Baciou pour avoir accepter d'évaluer ce manuscrit et pour leurs remarques qui ont permis de l'améliorer ainsi qu'envers les autres membres du jury, Chahrazade El Amri et Anne Woisard pour leur participation et leur suivi des corrections de ce manuscrit.

J'aimerais particulièrement remercier Marten Vos pour son aide précieuse aux différentes expériences de spectroscopie et à l'analyse de leur résultats ainsi que pour m'en avoir expliqué les principes avec patience et pédagogie.

Ma gratitude va à Latifa Bouzhir et Joelle Khun qui m'ont particulièrement aidé à a paillasse et dont la compagnie a été des plus agréables.

J'aimerais remercier Simon Dadoun d'avoir pris sur son temps libre pour avoir écrit le programme présenté en annexe de ce travail avec un certain enthousiasme.

Merci à Jean Christophe Lambry pour les différents modèles, à Sophie Bourcier pour avoir réalisé les expériences de spectrométrie de masses, à Olivier Tenaillon pour m'avoir fourni la souche $\Delta deoA$ et le plasmide pCP-20.

Merci à tous les membres du LOB pour les bons moments passés en leur compagnie.

J'aimerais aussi remercier la promotion 2012 du cours Pasteur de biochimie des protéines pour les longues heures d'encouragements mutuels à nos auto apitoiements respectifs.

Enfin, je remercie mes parents et mes amis qui m'ont soutenu tout au long de ce travail et dont certains ont beaucoup aidé à la rédaction de ce manuscrit.

Sommaire

1 Introduction **1**

1.1 Évolution du support génétique 1

1.1.1 Monde ARN 2

1.1.2 Le monde RNP 4

1.1.3 Passage de l'ARN à l'ADN 7

1.1.4 Passage de l'ADN U à l'ADN T 8

1.2 La lettre "T" dans l'ADN : Les thymidylate synthases 10

1.2.1 Les thymidylates synthases ThyA 10

1.2.2 Les thymidylate synthases flavine-dépendantes ThyX 12

1.3 La lettre "T" dans l'ARN : modifications post-transcriptionelles . . . 16

1.3.1 Les différentes méthylations post-transcriptionnelles des uridines 16

1.3.2 Les enzymes de méthylation post-transcriptionelles des uridines 18

1.3.3 La méthylaminoéthylation de l'uridine 34 d'ARNt 21

1.4 Les flavoprotéines 24

1.4.1 Propriétés (photo)chimiques du noyau isoalloxazine 24

1.4.2 Caractéristiques spectrales 26

1.4.3 Monooxygénases et déshydrogénases 27

1.4.4 Les biosenseurs flavoprotéiques 28

1.5 Problématique 29

2 Matériels et méthodes **31**

2.1 Techniques de biologie moléculaire 31

2.1.1 Liste des souches d'*E.coli* 31

2.1.2 Composition des milieux de culture 32

2.1.3 Cartes des plasmides utilisés 33

2.1.4 Transformation bacterienne 35

2.1.5 Préparation de cellules d'*Escherichia coli* électrocompétentes . . . 36
2.1.6 Inactivation du gène *thyA* chez la souche Δ*deoA* . . . 37
2.1.7 Vérification des génotypes Δ*deoA* et *thyA*::*aphA-3* . . . 38
2.1.8 Évolution dirigée de TrmFO . . . 38
2.1.9 Liste des amorces PCR . . . 39
2.1.10 Mutagénèse dirigée . . . 40
2.1.11 Liste des amorces utiliséés pour la mutagénèse dirigée . . . 41
2.1.12 Séquence du gène de synthèse codant pour TrmFO de *Thermus thermophilus* . . . 42
2.2 Techniques d'études biochimiques . . . 42
2.2.1 Expression et purification de TrmFO et de ses mutants . . . 42
2.3 Approches spectroscopiques . . . 43
2.3.1 Anisotropie de fluorescence . . . 43
2.3.2 Spectroscopie d'absorption . . . 44
2.3.3 Photoconversion de la protéine TrmFO et de ses différents mutants . . . 44
2.3.4 Chromatographie en phase liquide couplée à un spectromètre de masses . . . 44
2.3.5 Fluorescence résolue en temps . . . 45
2.4 Informatique . . . 45

3 Métabolisme des pyrimidines chez *E.coli* 46
3.1 Voie de complémentation du phénotype Δ*thyA* par des analogues nucléosidiques . . . 47
3.1.1 Complémentation du phénotype Δ*thyA* par la ribothymidine (5mU) et la méthylcytidine (5mC) . . . 47
3.1.2 Construction du double mutant Δ*deoA*Δ*thyA* . . . 48
3.1.3 Tests de complémentation des différents mutants par des analogues nucléosidiques . . . 53
3.2 Évolution dirigée de *trmFO* . . . 54
3.2.1 Présentation de la méthode . . . 55
3.2.2 Système de sélection . . . 56
3.2.3 Banque de mutants combinatoire du gène *trmFO* de *Bacillus subtilis*, approche semi-rationnelle . . . 56

3.2.4 Mutagénèse aléatoire par *error prone* PCR 59
3.2.5 Développement d'un utilitaire d'optimisation de l'encodage des gènes pour la mutagénèse aléatoire 60
3.3 Conclusions et ouvertures 62

4 Caractérisation biochimique de la flavoenzyme TrmFO de *Thermus thermophilus* 66
4.1 Fixation du substrat 69
4.1.1 Principe de l'anisotropie de fluorescence 71
4.1.2 Résidus impliqués dans la fixation de TrmFO à son substrat . 73
4.2 Caractérisation spectrale de TrmFO de *Thermus thermophilus* 75
4.2.1 Spectroscopie d'absorption UV-visible 76
4.2.2 Identification de l'espèce chimique absorbant à 400 nm 78
4.2.3 Spectroscopie d'émission de fluorescence à l'état stationnaire . 80
4.3 Spectroscopie ultra-rapide 82
4.3.1 Principe 82
4.3.2 Identification d'un second *quencher* de fluorescence 83
4.4 Chromatographie en phase liquide couplée à un spectromètre de masses 85
4.4.1 Principe 85
4.4.2 Indication d'un intermédiaire covalent FAD-ARNt par LC-MS 87
4.5 Conclusions et ouvertures 88

5 Conclusions générales et perspectives 92

Bibliographie 96

Liste des tableaux 107

Liste des figures 108

Annexes 111
A Folate-Dependent Thymidylate-Forming Enzymes : "All roads lead to T" 111
B Présentation de posters aux différents congrès 126
1. Anti-microbial compounds targeting the alternative thymidylate synthase ThyX. 127
2. Active site dynamics of flavin-dependent methylases. 128

3. Ultrafast electron transfer and visualisation of active site dynamics in the flavoenzyme TrmFO. 129
C Code source de l'utilitaire d'optimisation d'encodage des gènes pour la mutagénèse aléatoire 130

Liste d'abréviations

ADN : Acide désoxyribonucléique
Amp : Ampicilline
ADN : Acide désoxyribonucléique
Amp : Ampicilline
aphA-3 : Gène de résistance à la kanamycine
ARN : Acide ribonucléique
ARNm : Acide ribonucléique messager
ARNr : Acide ribonucléique ribosomique
ARNt : Acide ribonucléique de transfert
BET : Bromure d'éthidium
bla : Gène codant pour la beta-lactamase
BSA : Albumine sérique bovine
Cm : Chloramphénicol
COG : *Cluster of Orthologous Groups of proteins*
deoA : Gène codant pour la thymidine phosphorylase
dNTP : Désoxyribonucléotide
DO : Densité optique
dTMP : Désoxythymidine monophosphate
dUMP : Désoxyuridine monophosphate
ESI : *Electro Spray Ionisation*
ETF : Flavoprotéine de transfert d'électrons
FAD : Flavine adénine dinucléotide
FLP : Flippase
FMN : Flavine mononucléotide
folA : Gène codant pour la dihydrofolate réductase
FRT : *Flippase recognition target*

FTICR : *Fourrier Transform Ion Cyclotron Resonnance*
HEPES : Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
IPTG : Isopropyl β-D-1-thiogalactopyranoside
Kan : Kanamycine
LB : Lysogeny broth ("Luria Bertani")
LEW : *Lysis equilibrium wash*
MALDI : *Matrix Assisted Laser Desorption Ionisation*
MCS : *Multiple cloning site*
MTHF : 5,10-Méthylènetétrahydrofolate
NADH : Nicotinamide adénine dinucléotide réduit
NADPH : Nicotinamide adénine dinucléotide phosphate
PDB : *Protein Data Bank*
Ni-TED : Nickel-tris(carboxymethyl)ethylene diamine
PCR : Réaction de polymérisation en chaîne
RNP : Ribonucléoprotéine
SAM : S-adénosyl méthionine
SOB : *Super Optimal Broth*
SOC : *Super Optimal Broth* avec glucose 20 mM
ThyA : Thymidylate synthase A
ThyX : Thymidylate synthase X (flavine-dépendante)
TrmA : tRNA-(m5U54) méthyltransférase
TOF : *Time of Flight*
TrmFO : Méthylènetetrahydrofolate-tRNA-(uracil-5-)-méthyltransférase
UNG : Uracil-ADN glycosylase

Chapitre 1

Introduction

1.1 Évolution du support génétique

Une des caractéristiques fondamentales du vivant est le programme selon lequel les organismes peuvent s'adapter à leur milieu, y puiser les ressources nécessaires à leur homéostasie et *in fine* à leur reproduction [1]. Chez les cellules vivantes, ce programme, dit génétique, a pour support de stockage ubiquitaire l'ADN (acide 2'désoxyribonucléique), un polymère de quatre nucléotides différents formant entre eux des liaisons phosphodiester ; ces polymère s'organisent en deux brins antiparallèles et complémentaires, formant une double hélice [2]. Cette structure permet la réplication de l'ADN lors de la division cellulaire en vue de la transmission héréditaire du programme génétique (Fig. 1.1).

L'exécution de ce programme génétique se fait selon un processus paradigmatique de transcription de l'ADN en ARN messagers (ARNm) puis de traduction de ces ARN messagers en protéines, des polymères d'acides aminés formant des liaisons peptidiques. La transcription est catalysée par une machinerie essentiellement protéique, tandis que la traduction est assurée par une machinerie nucléoprotéique, composée (pour la partie nucléique) d'ARN ribosomiques (ARNr) et d'ARN de transfert (ARNt).

Les deux différences principales entre ADN et ARN résident d'une part en un groupement hydroxyle sur le carbone 2' du ribose présent dans l'ARN et absent de l'ADN, d'autre part, en un remplacement de l'uridine dans l'ARN par la thymidine dans l'ADN (Fig. 1.2).

Ces observations du vivant suggèrent deux étapes notables dans l'évolution du

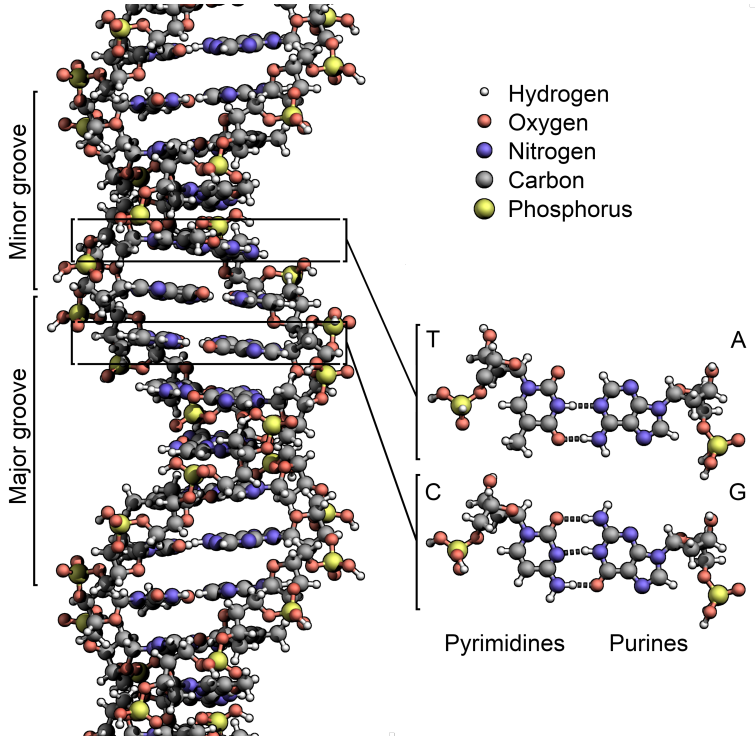

FIGURE 1.1 – **Schéma représentant la structure de la double hélice d'ADN.** Les deux brins sont antiparallèles et complémentaires l'un de l'autre par des interactions Watson-Crick A-T et G-C (images.huffingtonpost.com/2013-12-17-fig1_sm.jpg).

programme génétique ayant mené à l'ADN des organismes modernes : le passage d'un monde ARN (A,U,G,C) (Fig. 1.3) à l'ADN (dA, dU, dG, dC) dans un premier temps, et celui d'un ADN contenant de l'uridine (dU) à de l'ADN contenant de la thymidine (dT) dans un deuxième.

1.1.1 Monde ARN

L'hypothèse du monde ARN a été formulée pour la première fois par Walter Gilbert [3] dès la découverte d'ARN ayant des activités catalytiques. Jusqu'alors, les premiers systèmes réplicatifs étaient supposés être composés à la fois de protéines et d'ARN. La découverte d'ARN catalytiques (ribozymes), au nombre de deux à l'époque, laissait déjà entrevoir la possibilité de multiples réactions catalysables par des ARN. Des réactions qui se sont vérifiées plus tard [1], permettant au modèle de s'affranchir des protéines, le rendant plus simple et donc plus probable. Les mécanismes

1. notamment par la méthode SELEX [4] qui permet la sélection d'ARN ayant des propriétés catalytiques parmi une banque d'oligonucléotides combinatoire

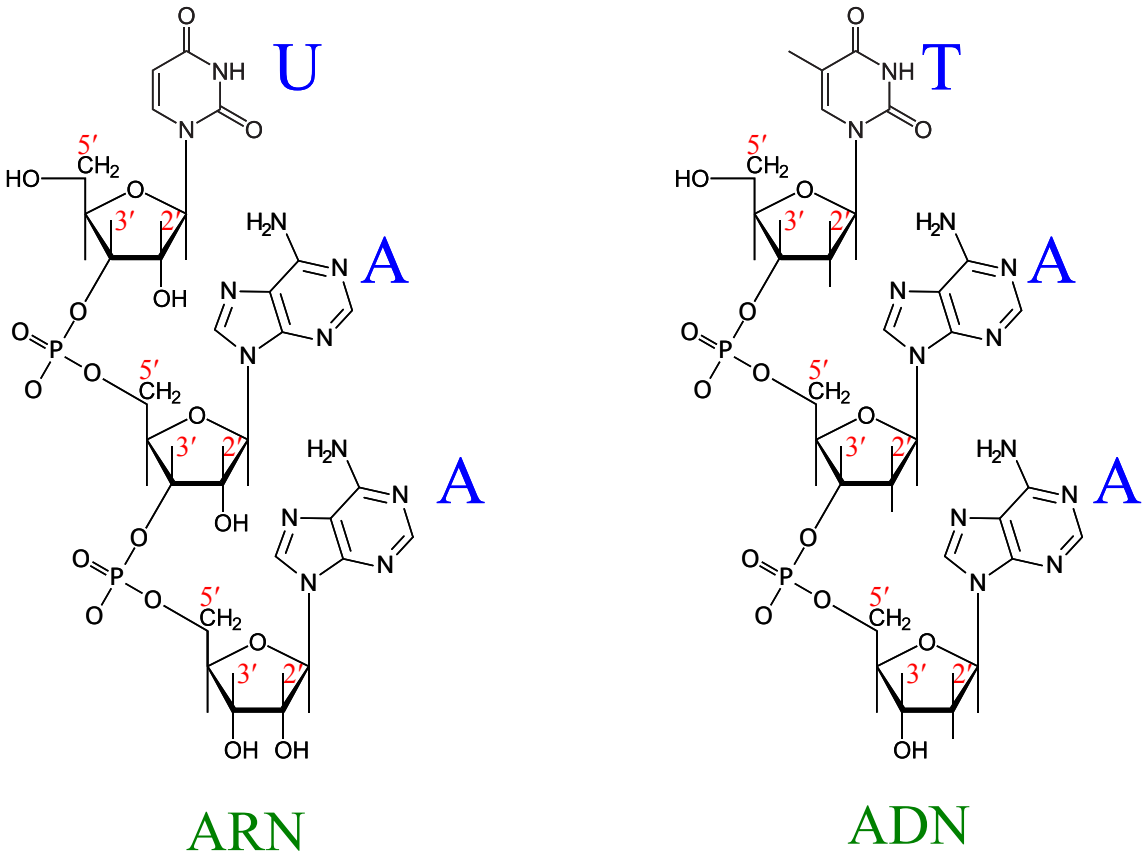

FIGURE 1.2 – **Comparaison des structures de l'ADN et de l'ARN** : les nucléotides sont liés entre eux par des liaisons 5'-3' phosphodiester. La principale différence réside en la présence d'un groupement hydroxyle en 2' présent sur les riboses de l'ARN et absent des désoxyriboses de l'ADN. Une autre différence notable est la lettre "U" de l'ARN correspondant à la lettre "T" de l'ADN.

d'auto-épissage chez les ARN leur conférant un important potentiel évolutif constituait un argument supplémentaire.

Cette hypothèse du monde ARN a ensuite été renforcée avec la découverte du pouvoir de réplication non enzymatique des ARN [5]. Les protéines seraient apparues plus tard et sélectionnées pour leurs activités catalytiques plus efficaces. Le fait que les machineries d'expression de protéines des organismes vivants contemporains passent par des ARN (le ribosome traduit les ARNm à l'aide des ARNt) est un bon indicateur du passage de l'ARN aux protéines comme catalyseur.

On peut ainsi distinguer le deuxième âge du monde ARN du premier âge. Le premier âge correspondant à l'étape où les catalyseurs biologiques étaient exclusivement ribonucléiques, et le deuxième âge (aussi appelé monde RNP pour ribo-

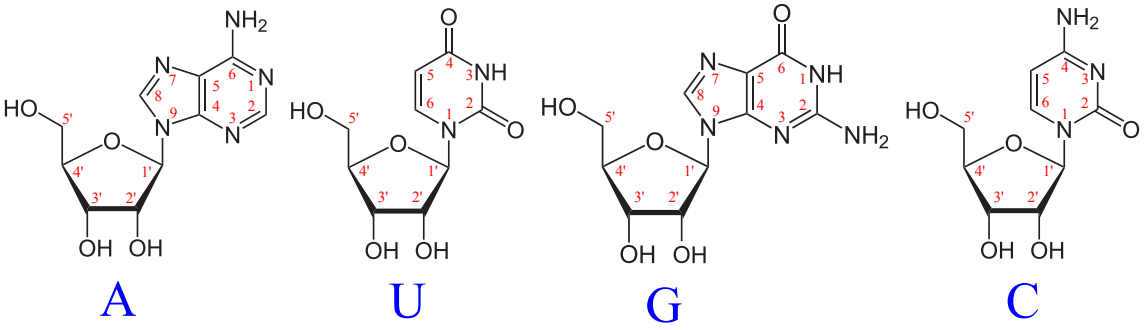

FIGURE 1.3 – **Structure des différents ribonucléosides, constituants de l'ARN** et numérotation de leurs carbones et azotes (en rouge). En bleu les lettres correspondantes : A (adénosine) U (uridine) G (guanosine) et C (cytidine).

nucléoprotéique) correspondant à une ère de systèmes réplicatifs composés d'ARN et de protéines.

1.1.2 Le monde RNP

Les protéines ont, dans le vivant actuel, des rôles multiples tels que la catalyse de réactions chimiques du métabolisme, la structure de la cellule [6], le transport intra-cellulaire [7] et inter-cellulaire [8] de molécules etc... Les protéines sont des macromolécules composées d'une ou plusieurs chaînes peptidiques, des hétéropolymères d'acides aminés reliés entre eux par des liaisons peptidiques. La formation de la liaison peptidique est catalysée par le ribosome au cours du processus de traduction des ARN messagers selon le code génétique. On dénombre 20 acides aminés protéinogènes, c'est-à-dire pouvant être incorporé à la chaîne peptidique en cours d'élongation pendant la traduction [2]. Tout au long de cet exposé, les différents acides aminés seront référencés selon la nomenclature habituelle donnée par le tableau 1.1.

L'ARNm est donc composé de quatre nucléotides différents et les protéines, de vingt acides aminés différents, il en découle naturellement qu'un acide aminé donné doit être encodé par un ensemble d'au moins trois nucléotides, ensemble appelé codon. Quatre possibilités pour trois emplacements donne $4^3 = 64$ possibilités. Il en résulte que le code génétique est dégénéré, c'est-à-dire qu'un acide aminé peut être encodé par plusieurs codons différents (Fig. 1.4) en plus d'être universel (identique

2. On en compte en réalité deux de plus, la sélénocystéine [9] et la pyrolysine [10] [11], que l'on peut considérer comme des cas particuliers et dont je ne parlerai pas dans cet exposé.

TABLEAU 1.1 – **Liste des vingt acides aminés standards et leur nomenclature correspondante**.

Glycine	G	Gly		Aspartate	D	Asp
Alanine	A	Ala		Glutamine	Q	Gln
Valine	V	Val		Asparagine	N	Asn
Leucine	L	Leu		Sérine	S	Ser
Isoleucine	I	Ile		Thréonine	T	Thr
Tyrosine	Y	Tyr		Cystéine	C	Cys
Phénylalanine	F	Phe		Lysine	K	Lys
Tryptophane	W	Trp		Arginine	R	Arg
Glutamate	E	Glu		Méthionine	M	Met
Proline	P	Pro		Histidine	H	His

pour la plupart des êtres vivants [3]).

Le déchiffrage de chaque codon se fait via un ARNt chargé de son acide aminé et comptant dans sa structure primaire l'anticodon (séquence de trois nucléotides complémentaire du codon correspondant). La formation de la liaison peptidique est ensuite catalysée par une partie essentiellement nucléique [14] du ribosome [4].

Le code génétique contemporain s'est probablement établi progressivement et de manière séquentielle. En 1993, Baumann et Oro ont proposé un modèle d'évolution du code génétique en trois étapes [15]. D'abord, un code génétique primitif [16] ayant des codons de deux bases permettait l'expression de protéines contenant douze acides aminés différents comme décrit par la figure 1.5. La provenance des différents acides aminés peut, quand à elle, s'expliquer par une approche *bottom up* de chimie pré-biotique [17].

Au cours d'une deuxième étape, le code aurait évolué et permis l'incorporation, dans les protéines, d'acides aminés chimiquement plus complexes tels que l'histidine, la cystéine, la phénylalanine et la tyrosine ainsi que celle de l'asparagine et du glutamate. Ce ne serait qu'au cours d'une troisième et dernière étape que la méthionine et le tryptophane auraient été intégrés au code.

Le fait que la machinerie d'expression protéique soit essentiellement ribonucléique permet de concevoir une forme de vie composée d'ARN et de protéines qui serait

3. Il existe quelques exceptions [13] telles que chez les mitochondries des vertébrés, les plastes des plantes, certaines bactéries et archées...

4. Nous reviendrons plus en détails sur ce processus de traduction dans une partie ultérieure sur les modifications post-transcriptionelles des ARNt et des ARNr.

1ère base	2nde base U	C	A	G	3ème base
U	F F L L	S S S S	Y Y STOP STOP	C C STOP W	U C A G
C	L L L L	P P P P	H H Q Q	R R R R	U C A G
A	I I I M	T T T T	N N K K	S S R R	U C A G
G	V V V V	A A A A	D D E E	G G G G	U C A G

FIGURE 1.4 – **Tableau matérialisant le code génétique** [12].

1ère base \ 2ème base	U	C	A	G
U	Leu	Ser	STOP	STOP
C	Leu	Pro	Gln	Arg
A	Ile	Thr	Lys	? Ser Arg
G	Val	Ala	Asp	Gly

FIGURE 1.5 – **Postulat d'un code génétique primitif** (selon [15]).

antérieure à l'apparition [5] de l'ADN. En outre, dans le vivant, les désoxyribonucléotides sont issus de la réduction des ribonucléotides par les ribonucléotide réductases via un mécanisme réactionnel radicalaire (Fig. 1.6) ; mécanisme auquel les ARN seraient sensibles (sensibilité liée à leur instabilité, détaillée dans la section suivante). Cette sensibilité suggère la nécessité de catalyseurs protéiques pour l'apparition de l'ADN [18] [19].

On peut remarquer ici que le mécanisme des ribonucléotide réductases implique un radical cystéinyl et que le code génétique primitif postulé ne permettait pas l'encodage de cet acide aminé. On peut donc supposer que l'ADN n'a pu apparaître

5. Il convient d'entendre, par apparition, l'aboutissement d'un processus stochastique.

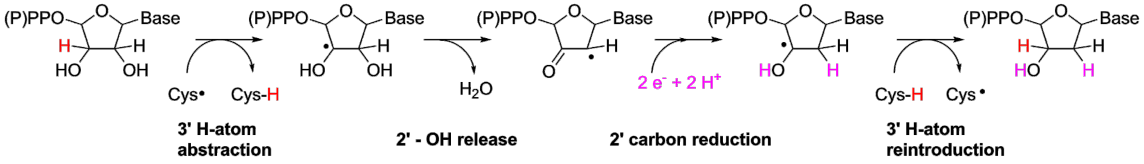

FIGURE 1.6 – **Mécanisme général des ribonucléotides réductases** [18]. La réduction des ribonucléotides se fait en quatre principales étapes : l'extraction de l'atome d'hydrogène en 3' permet l'activation du substrat dont la conséquence est la libération du groupement 2'OH en une molécule d'eau.

que lors de la deuxième étape de l'évolution du code génétique. L'intégration de la cystéine dans le code aurait été une étape limitant l'évolution du support génétique de l'ARN en ADN.

1.1.3 Passage de l'ARN à l'ADN

Une des limites de l'ARN comme molécule du support génétique est son instabilité intrinsèque en conditions alcalines. La présence de fonctions alcool en 2' permet l'attaque nucléophile d'un ion hydroxyle sur son hydrogène ; s'ensuit une attaque d'un doublet non liant de l'oxygène sur le phosphate en 3' pour finir par une attaque de l'oxygène 5' du nucléotide précédent sur une molécule d'eau, libérant une extrémité 5'OH et une autre 3' 2' phosphate cyclique (Fig. 1.7).

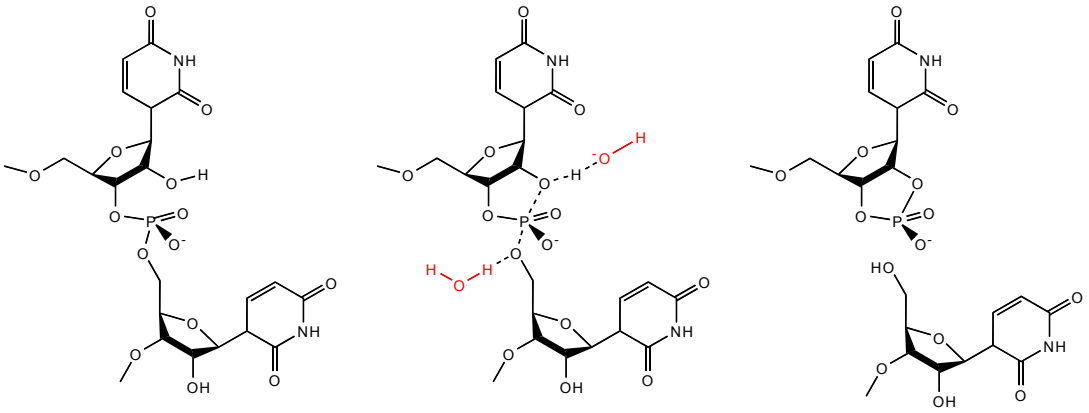

FIGURE 1.7 – **Instabilité intrinsèque de la liaison phosphodiester des ARN et mécanisme de son hydrolyse en condition alcaline.** Un ion hydroxyle attaque l'hydrogène du 2'OH dont l'oxygène attaque le phosphate en 3', le libérant de sa liaison avec l'extrémité 5' du nucléotide précédent. Enfin, l'oxygène 5' récupère son proton par une attaque sur une molécule d'eau pour laisser une extrémité 5'OH et une autre 2' 3' phosphate cyclique.

Pour assurer une adaptabilité pérenne des organismes à leurs environnements, le support physique du programme génétique, qu'ils conserveront tout au long de leur vie, se doit d'être stable. La sélection des ribonucléotides réductases, qui catalysent l'excision du groupement hydroxyle en 2' des ribonucléotides, a permis d'augmenter cette stabilité en empêchant l'hydrolyse de la liaison phosphodiester, marquant ainsi le passage de l'ARN à l'ADN, support ubiquitaire des organismes vivants actuels.

1.1.4 Passage de l'ADN U à l'ADN T

L'ADN permettait donc une meilleure stabilité structurale du support génétique, mais comptait encore une instabilité du message liée à la désamination spontanée des cytidines (Fig. 1.8). L'ADN alors composé des bases dA, dU, dG, et dC aurait été sujet à une transformation du dC vers le dU comme le montre la figure 1.8.

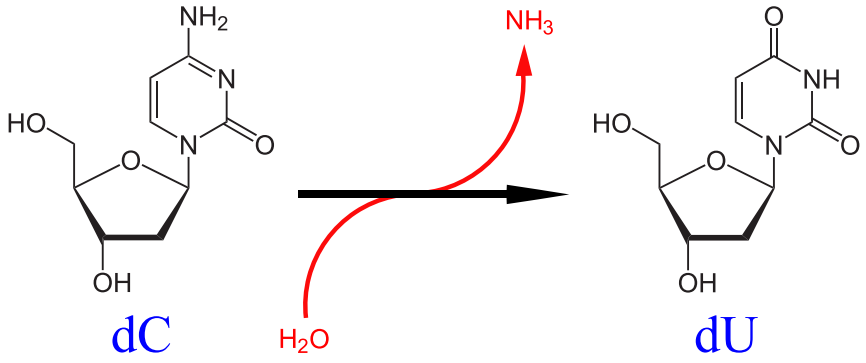

FIGURE 1.8 – **Réaction de désamination spontanée des cytidines.** La 2'désoxyribocytidine (dC) se transforme en 2'désoxyribouridine (dU), engendrant des erreurs dans le maintient du programme génétique.

Cette réaction spontanée était probablement délétère dans la mesure où elle peut introduire des erreurs dans le programme génétique en transformant une base en une autre. On peut facilement imaginer que si tous les "a" de ce manuscrit pouvaient se transformer en "e" de manière aléatoire et spontanée, le message risquerait de s'en trouver fortement altéré. Une solution pour pallier ce problème serait de dessiner un nouveau caractère pour remplacer le "e" et de lire tous les "e" comme des "a" (chose que le lecteur de ce manuscrit est invité à ne pas faire). Ce désavantage aurait mené à la sélection, dans un premier temps, des uracile-N-glycosylases qui détectent les mésappariements G-U dans l'ADN double brin pour remplacer le U par un C, et dans un deuxième temps, la sélection des thymidylate synthases [20].

Les thymidylate synthases catalysent la méthylation des uridines sur le carbone 5 (Fig. 1.9), permettant aux machineries cellulaires de discriminer entre, d'une part, les uridines issues de la désamination des cytidines, et d'autre part les uridines incorporées.

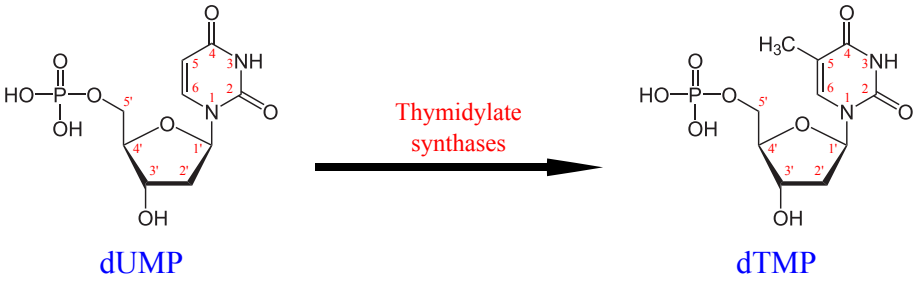

FIGURE 1.9 – **Réaction catalysée par les thymidylate synthases**. Les carbones et azotes du thymidylate (dTMP) et de l'uridylate (dUMP) sont numérotés en rouge.

Le thymidylate est ainsi devenu un précurseur indispensable et ubiquitaire de la synthèse d'ADN des organismes vivants contemporains. La première famille d'enzymes identifiée comme catalysant la synthèse *de novo* du thymidylate fut ThyA, dont fait partie la thymidylate synthase humaine. Cette biosynthèse s'effectue via un mécanisme de réduction du 5,10-méthylènetetrahydrofolate (MTHF) en dihydrofolate. Plus récemment, une nouvelle famille d'enzymes a été découverte par analyse informatique des séquences d'organismes desquels le gène *thyA* était absent [21]. Les enzymes de cette famille, nommée ThyX, catalysent la synthèse *de novo* du thymidylate de manière significativement différente : elles utilisent le MTHF comme donneur de méthyle qui sera transformé en tétrahydrofolate, et le nicotinamide adénine dinucléotide phosphate (NADPH) comme agent réducteur qui permet de finaliser la réaction, par l'intermédiaire d'un cofacteur FAD.

Les thymidylate synthases ThyA sont globalement plus efficaces et les organismes qui les utilisent ont des génomes statistiquement plus grands [22]. Cependant l'utilisation de ThyA va de pair avec la nécessité d'exprimer la dihydrofolate réductase codée par le gène *folA*, gène pouvant être absent de certains organismes qui utilisent ThyX exclusivement [23].

Les deux familles d'enzymes catalysent donc la synthèse du thymidylate via des mécanismes différents (qui seront détaillés dans la partie suivante) et n'ont pas d'origine évolutive commune, ce qui implique que la lettre "T" de l'ADN constitue une convergence évolutive et aurait été sélectionnée deux fois, indépendamment, au cours de l'évolution.

1.2 La lettre "T" dans l'ADN : Les thymidylate synthases

1.2.1 Les thymidylates synthases ThyA

Les enzymes ThyA [24] constituent, *a priori*, la seule famille de thymidylate synthases présente chez les eucaryotes mais sont présentes dans les trois domaines du vivant. Elles se présentent sous forme d'enzymes homodimériques dont les sous-unités ont des masses moléculaires allant de 30 à 35 kDa (Fig. 1.10). Les enzymes ThyA, depuis leur découverte, ont été très largement étudiées dans la mesure où la synthèse *de novo* du thymidylate est une étape nécessaire à la réplication de l'ADN ; elles constituent une cible thérapeutique de choix.

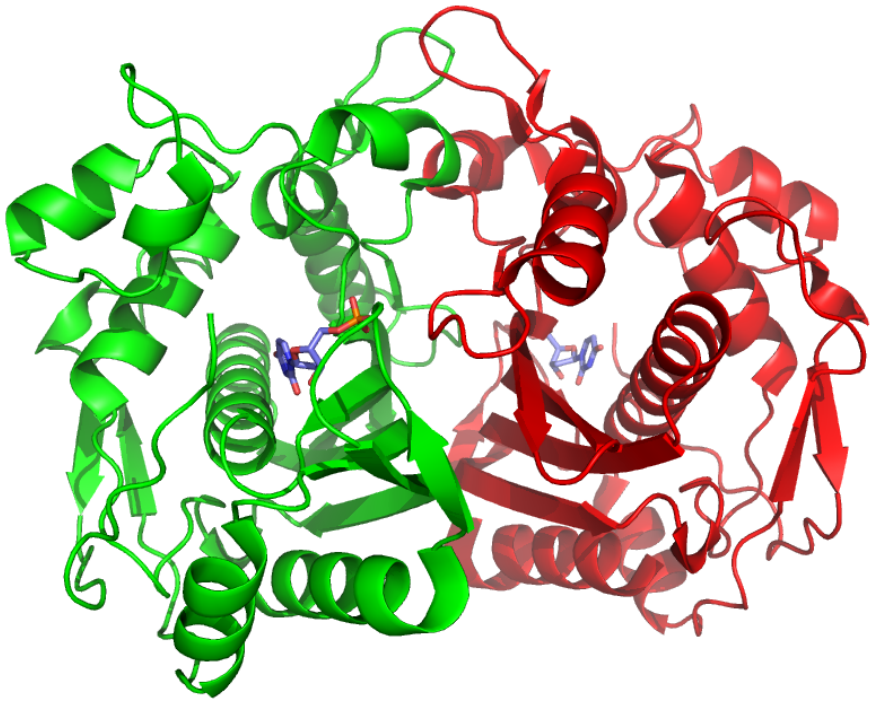

FIGURE 1.10 – **Structure de l'enzyme ThyA de *Bacillus subtilis*** en représentation PyMOL cartoon (code PDB : 1BSP). Les deux sous-unité sont matérialisées par des couleurs différentes (rouge et verte). Le dUMP (en bleu foncé) indique le site actif de chaque sous unités.

Les thymidylate synthases de type ThyA catalysent la synthèse du dTMP de manière séquentielle et ordonnée en six étapes notables (Fig. 1.11). Après fixation du dUMP dans le site actif, la première étape consiste en une attaque nucléophile

d'un groupement thiol sur le carbone 6 du dUMP pour former un intermédiaire covalent instable. Au cours d'une deuxième étape, un intermédiaire énolate se forme par transfert d'électrons vers le méthylènetétrahydrofolate, qui est alors activé en un ion iminium, lequel fixera ensuite covalemment l'énolate. Au cours de la troisième étape, l'hydrogène du carbone 5 de l'énolate sera capté par une base, entraînant une quatrième étape de β-élimination du tétrahydrofolate pour former un intermédiaire méthylène exocyclique. La cinquième étape de la réaction consiste en une réduction par le tétrahydrofolate du groupement méthylène transféré en un groupement méthyle, entraînant la β-élimination de l'enzyme par clivage de la liaison covalente entre la cystéine et le dTMP néoformé. Enfin, les produits de la réaction dTMP et dihydrofolate seront libérés au cours d'une sixième et dernière étape.

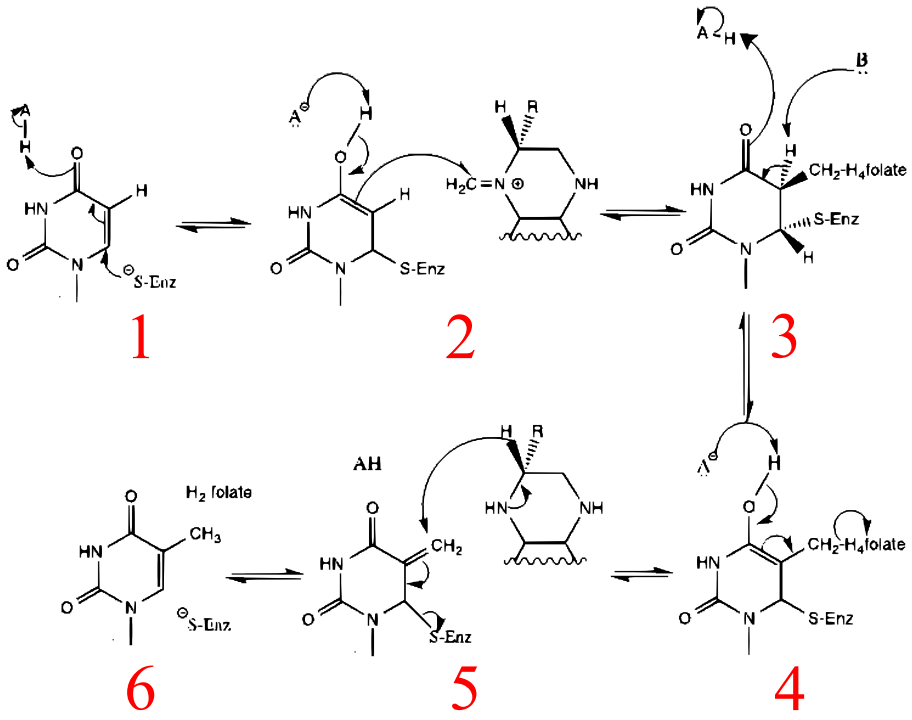

FIGURE 1.11 – **Mécanisme catalytique des enzymes de la famille ThyA [25].** 1 : Attaque nucléophile d'un groupement thiol sur le carbone 6 du dUMP. 2 : activation du méthylènetetrahydrofolate en ion iminium. 3 : recaptage de l'hydrogène du carbone. 5 de l'enolate exocyclique. 4 : β-élimination du tetrahydrofolate et formation de l'intermédiaire méthylène exocyclique. 5 : réduction du groupement méthylène en groupement méthyle par le tétrahydrofolate et β-élimination de l'enzyme. 6 : libération des produits du dihydrofolate et du dTMP, produits de la réaction.

1.2.2 Les thymidylate synthases flavine-dépendantes ThyX

La famille des thymidylate synthases ThyX, que l'on trouve exclusivement chez certaines bactéries, archées et quelques virus [6], a été découverte plus récemment [21] et les données concernant ces enymes sont moins nombreuses que celles concernant les ThyA. Elles ont cependant suscité un vif intérêt depuis leur découverte au vue de leur potentiel comme cible thérapeutique [27] [28]. Outre le fait que les thymidylate synthases catalysent la synthèse d'un précurseur *sine qua non* du vivant, de nombreuses bactéries pathogènes pour l'homme tels que *Chlamydia trachomatis*, *Mycobacterium tuberculosis* [7], *Borrelia burgdorferi* (facteur étiologique de la maladie de Lyme) ont une thymidylate de la famille des ThyX (Fig. 1.12).

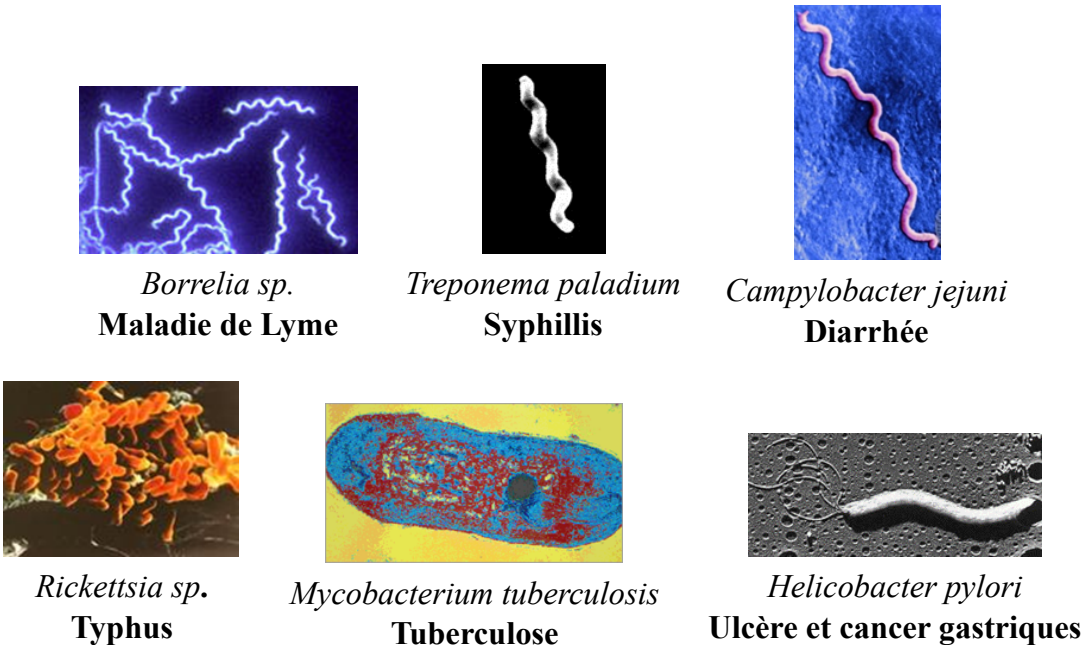

FIGURE 1.12 – **Différentes bactéries pathogènes utilisant une thymidylate synthase ThyX.**

Enfin, les enzymes ThyX étant significativement différentes, en termes de séquences, de structures (Fig. 1.13) et de mécanisme réactionnel, des enzymes ThyA dont fait partie la thymidylate synthase humaine, la découverte d'inhibiteurs spécifiques de ThyX permettrait de cibler un pathogène sans impacter son hôte.

6. Là encore, au moins une exception existe, le gène *thyX* a été trouvé chez *Dictyostelium discoideum* [26], un eucaryote.

7. *Mycobacterium tuberculosis* a la particularité de porter sur son chromosome *thyA* et *thyX* ; ce dernier est essentiel. L'expression de chaque enzyme est fonction de l'état de la bactérie au cours du cycle d'infection [29][30].

Les thymidylate synthases ThyX se présentent (pour la plupart) sous forme d'homotétramères dont chaque sous-unité a une masse moléculaire comprise entre 26 et 30 kDa (Fig. 1.13).

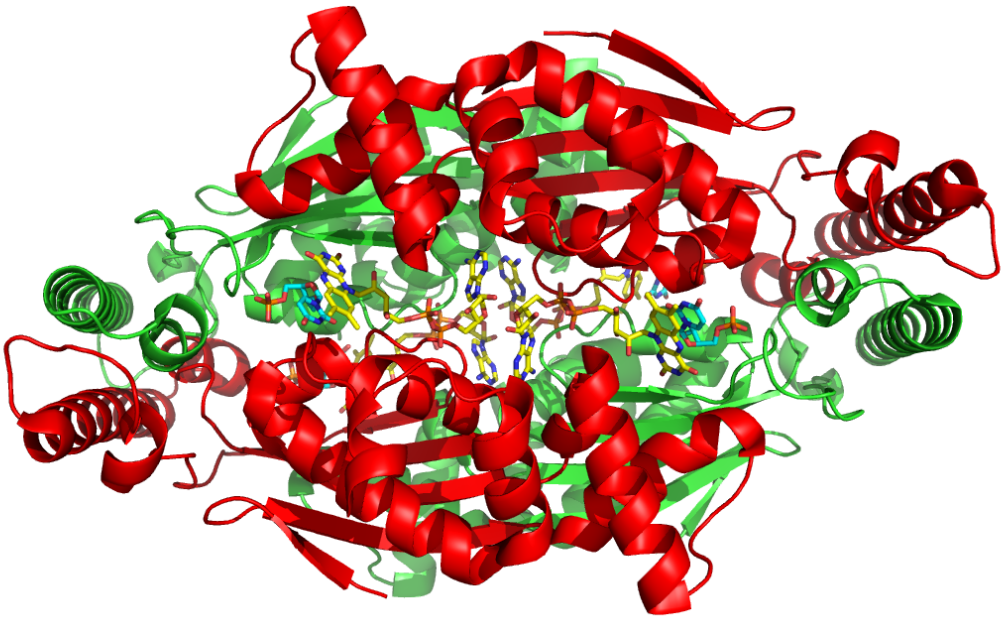

FIGURE 1.13 – **Structure de ThyX de *Thermotoga maritima* co-cristallisée avec FAD et dUMP** (code PDB : 1O26). L'enzyme est homotétramérique et chacun des quatre cofacteurs FAD se fixe à l'interface de deux sous unités successives, formant quatre sites actifs.

A la différence des enzymes ThyA qui utilisent le méthylènetétrahydrofolate comme donneur de carbone et comme agent réducteur, les thymidylate synthases ThyX ne l'utilisent que comme donneur de méthyle. L'étape de réduction est assurée par un troisième co-substrat, le NADPH.

La détection d'un intermédiaire covalent entre une thymidylate synthase ThyA et le 5 fluoro-dUMP à permis d'identifier une cystéine catalytique (numérotée 146 chez *E.coli*) qui initie la réaction par une attaque nucléophile sur le carbone 6 du dUMP (Fig. 1.11) tandis qu'aucun intermédiaire de ce type n'a été identifié chez ThyX. La recherche, chez ThyX, d'un nucléophile similaire à la cystéine catalytique de ThyA est restée infructueuse, tant et si bien qu'un nouveau mécanisme ne nécessitant pas d'activation préalable a été proposé [31], qui met en jeu la polarisation du dUMP par son environnement électrostatique au sein de l'enzyme (Fig. 1.14).

Chez les thymidylate synthases ThyX, le dUMP adopte une conformation assez inhabituelle, comme replié sur lui-même, qui pourrait être moteur de la catalyse.

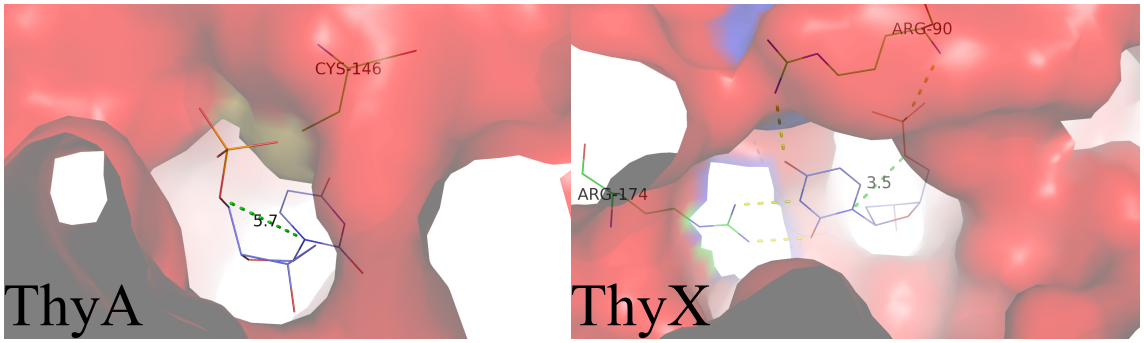

FIGURE 1.14 – **Comparaison des conformations du dUMP au sein de ThyA d'*Escherichia coli* (code PDB : 1BID) et de ThyX de *Thermotoga maritima* (code PDB : 1O25).** Pour ThyA d'*E.coli*, la cystéine 148 initie la réaction par une attaque nucléophile sur le carbone 6 du dUMP, tandis que chez ThyX, une charge négative du phosphate proche de l'azote 1 favorise sa charge, de même que les charges positives des arginines 174 et 90 favorisent une charge négative de l'oxygène lié au carbone 4.

Cette conformation permet au phosphate (chargé négativement) d'être à proximité de l'azote 1 et de favoriser une charge positive de celui-ci. En outre, deux arginines (numérotées 90 et 174 chez *Thermotoga maritima*) chargées positivement, favorisent la charge négative de l'oxygène du carbone 4 (Fig. 1.15).

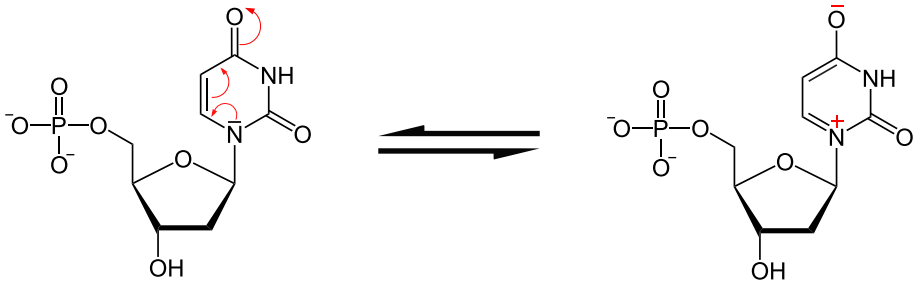

FIGURE 1.15 – **Mécanisme de polarisation du dUMP par les thymidylate synthases ThyX.** La conformation adoptée par le dUMP permet la stabilisation d'une forme chargée, stabilisée par deux arginines au sein du site actif.

Cette polarisation du dUMP permet une augmentation de la densité électronique du carbone 5 pour le rendre réactif. Ce n'est pas beaucoup plus tard que l'identification d'intermédiaires de réaction a permis d'établir un modèle de mécanisme catalytique plus complet [32] détaillé sur la figure 1.16.

Dans un premier temps se forme une base de Schiff sur le méthylènetétrahydrofolate permettant le transfert du méthylène sur l'azote 5 de la flavine (sous forme $FADH^-$). C'est grâce à la polarisation du dUMP, décrite précédemment, que se formera par la suite un pont méthylène entre le dUMP et la flavine. Le groupement méthylène

sera ensuite entièrement transféré sur le dUMP avant d'être réduit en groupement méthyle de paire avec une oxydation de la flavine en FAD. Enfin, le FAD sera réduit en $FADH^-$ grâce à l'oxydation du NADPH en $NADP^+$ pour permettre un nouveau cycle.

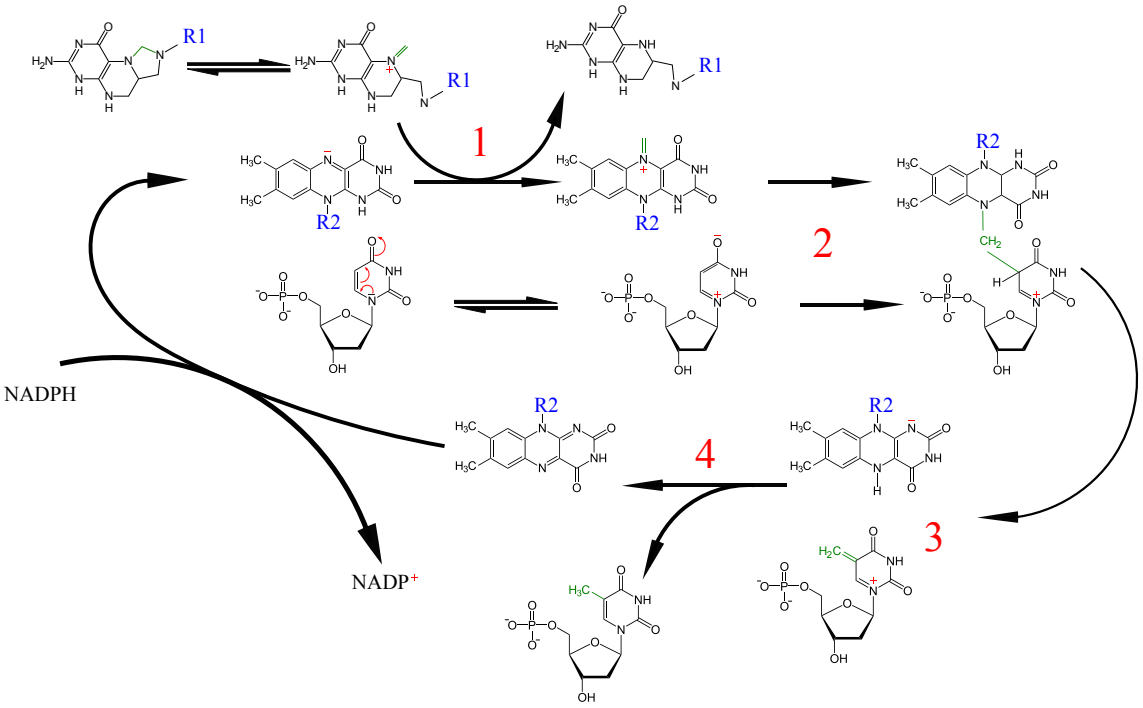

FIGURE 1.16 – **Mécanisme catalytique proposé pour les thymidylate synthases ThyX**, d'après [32]. La formation sur le MTHF d'une base de Schiff permet le transfert du méthylène sur l'azote 5 du FAD (1). Un pont méthylène peut ensuite se former entre le FAD et le dUMP grâce à la polarisation de celui-ci (2). Le méthylène est ensuite transféré au dUMP (3) avant que l'oxydation de la flavine ne permette la réduction du méthylène en méthyle (4). En fin de cycle, la flavine (FAD) sera réduite en $FADH^-$ par l'oxydation du NADPH en $NADP^+$.

La méthylation des désoxyuridines est ainsi apparue, de deux manières entièrement différentes, indépendamment au cours de l'évolution. Toujours indépendamment, et en réponse à des pressions évolutives différentes, est apparue la méthylation des uridines des ARN en ribothymidines. Contrairement à l'ADN, ces méthylations interviennent exclusivement sous forme de modifications post-transcriptionelles, c'est-à-dire sur des ARN, après leurs polymérisations.

1.3 La lettre "T" dans l'ARN : modifications post-transcriptionelles

Les modifications post-transcriptionelles des ARN sont relativement nombreuses et permettent la maturation fonctionnelle des ARN en modifiant leurs structures, leurs réactivités, leurs capacités à former des interactions avec d'autres partenaires ou autre.

Ce travail est essentiellement axé autour de la méthylation de l'uridine sur son carbone 5 mais se permettra une légère digression vers la méthylaminométhylation de l'uridine (toujours sur le carbone 5).

1.3.1 Les différentes méthylations post-transcriptionnelles des uridines

Les ribosomes sont des complexes nucléoprotéiques dont les différentes sous-unités sont caractérisées par leurs coefficients de sédimentation (en Svedberg, S). Ils assurent la synthèse protéique en catalysant la liaison peptidique entre deux acides aminés successifs (Fig. 1.17).

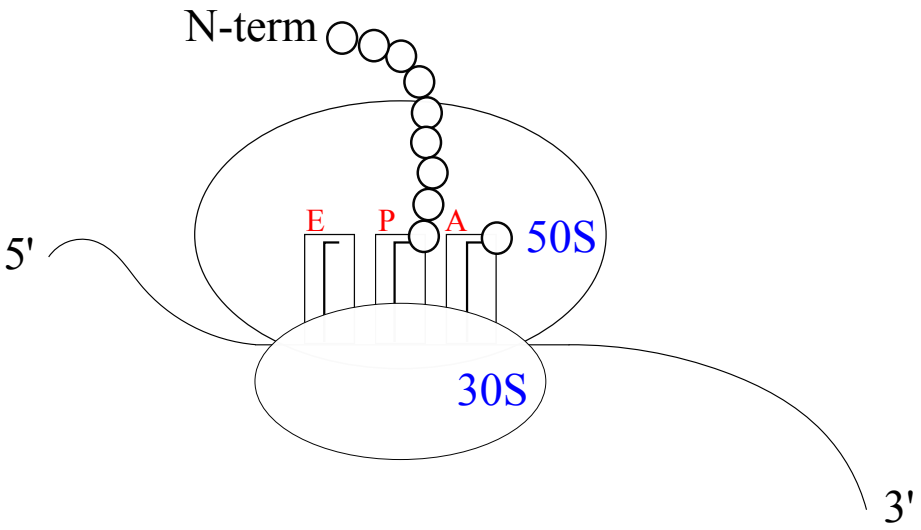

FIGURE 1.17 – **Schéma représentant la synthèse d'une protéine en cours d'élongation.** L'ARNt chargé de l'acide aminé correspondant au codon N+1 entre par le site accepteur (A) pour former une liaison peptidique avec la chaîne en cours d'élongation lié au peptidyl ARNt présent dans le site peptide (P). Enfin, l'ARNt déchargé se retrouve dans le site exit (E) et quittera le ribosome lors de la traduction du codon suivant.

Nous ne nous intéressons ici qu'au ribosome procaryote (70S), le ribosome eucaryote (80S) n'étant *a priori* pas sujet aux méthylations post-transcriptionelles des uridines sur le carbone 5. Le ribosome procaryote est composé de deux sous-unités nucléoprotéiques :

— la petite sous-unité 30S constituée de l'ARNr 16S et de 21 protéines

— la grande sous-unité 50S constituée de l'ARNr 5S, l'ARNr 23S et 34 protéines.

L'ARNr 23S de la grande sous-unité du ribosome procaryote est soumise à deux méthylations post-transcriptionnelles d'uridines numérotés U747 et 1939 chez *E.coli*.

L'uridine 1939 fait partie d'une région conservée (1930-1970) du domaine IV de l'ARNr 23S, impliquée dans une interaction avec l'ARNr 16S selon les données cristallographiques obtenues sur le ribosome 70S [33]. Bien que le rôle de sa méthylation n'ait pas été élucidé, il a été observé que la substitution de cette base augmentait la sensibilité à certains antibiotiques ciblant le ribosome [34]. L'uridine 747 se trouve dans la boucle 35 du domaine II de la grosse sous-unité du ribosome. Le rôle de sa méthylation reste à déterminer.

Les ARN de transfert (ARNt) sont des ARN relativement courts (70 à 100 nucléotides) constituant les intermédiaires clés dans le déchiffrage du code génétique. Chaque ARNt sera chargé d'un acide aminé donné par les aminoacyl-ARNt synthétases ; ainsi fonctionnels, ils pourront tour à tour s'apparier aux différents codons successifs des ARN messager (ARNm) et permettre au ribosome une synthèse peptidique dictée par l'ARNm.

La structure secondaire des ARNt, comparable à un trèfle, présente trois motifs en "tige-boucle" notables : la boucle T-Ψ, la boucle anticodon et la boucle A. Ces motifs se replient pour former une structure tertiaire en forme de "L". On retrouve la boucle anticodon à la pointe du "L" et à sa base, la boucle T-Ψ, le bout du "L" étant constitué des extrémités 5' et 3' (Fig. 1.18).

On trouve, invariablement en position 54 des ARNt, une uridine sujette à une méthylation (très conservée dans le vivant). La méthylation de cette uridine située au sein de la boucle T-Ψ favorise une interaction de type *reverse Hoogsteen* entre la méthyladénosine et cette uridine qui *in fine* renforce la structure en "L" de l'ARNt.

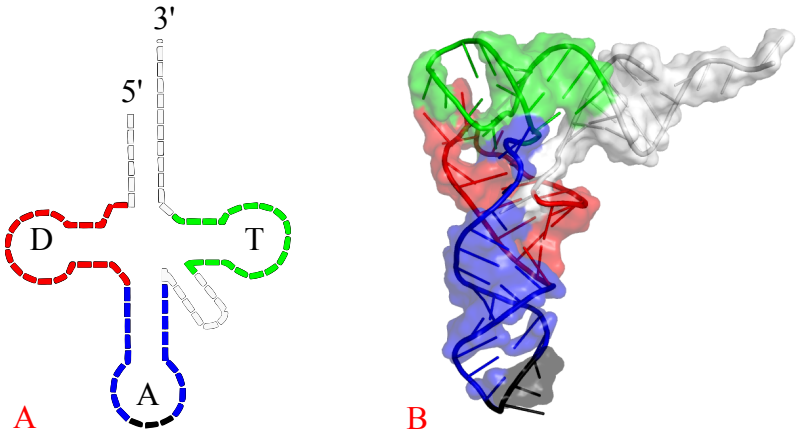

FIGURE 1.18 – **Structures secondaire (A) et tertiaire (B) des ARNt** en l'occurrence, celui de la phénylalanine chez la levure (code PDB : 4TNA). L'interaction entre la ribothymidine 54 et la méthyladénosine 58 au sein de la boucle T (en vert) permet au motif tige-boucle T d'interagir avec la tige-boucle D (en rouge), résultant en une structure des ARNt fonctionnelle en forme de "L" avec, à une extrémité la boucle anticodon (en noir) et à l'autre le 3'OH pouvant fixer l'acide aminé.

1.3.2 Les enzymes de méthylation post-transcriptionelles des uridines

La plupart des enzymes méthylant les uridines des ARN utilisent, comme donneur de méthyle, le S-adénosylméthionine (SAM). L'enzyme la mieux caractérisée utilisant ce mécanisme est TrmA dont la fonction est la méthylation de l'uridine 54 des ARNt chez des organismes appartenant aux trois domaines du vivant. Le mécanisme catalytique de TrmA, une fois résolu [35], a pu être appliqué aux autres enzymes SAM-dépendantes de méthylation post-transcriptionelles des uridines.

Dans un premier temps, l'interaction entre TrmA et son substrat induit un changement conformationnel de l'ARNt permettant la sortie de l'uridine ciblée de la boucle et son insertion dans le site actif de l'enzyme (Fig. 1.19).

S'ensuit une attaque nucléophile de la cystéine numérotée 324 (chez *E.coli*) sur le carbone 6 de l'uridine qui permettra ultérieurement, via un transfert de charges, l'attaque du groupement méthyle du SAM par un des doublets liant le carbone 5 au carbone 6. La réaction se termine par la récupération de l'hydrogène lié au carbone 5, ceci par une base permettant la formation d'une double liaison et la libération du substrat (Fig. 1.20).

Deux autres enzymes de méthylation post-transcriptionelles des ARN sur le carbone 5 d'uridines, et utilisant le SAM comme donneur de méthyle, ont été identifiées ;

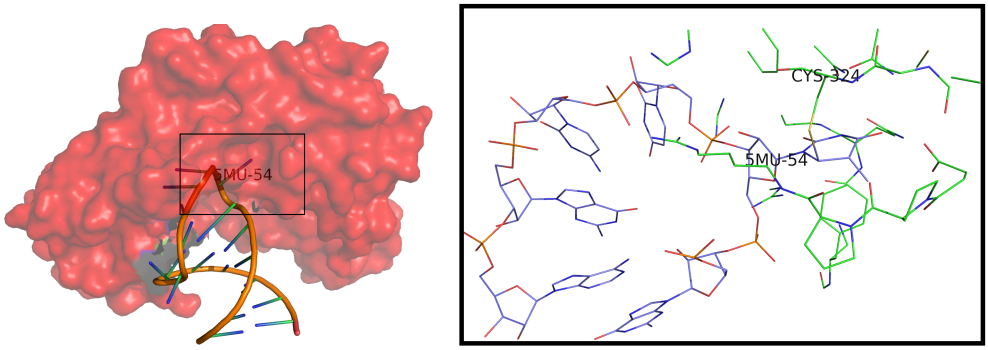

FIGURE 1.19 – **Structure de l'enzyme TrmA d'*E.coli* cocristallisée avec un substrat minimal ribonucléique (code PDB : 3BT7).** L'interaction entre la protéine (en vert dans l'encadré) et son substrat (dont les carbones sont en violet) entraine un changement conformationnel de l'ARN et l'insertion de l'uridine 54 dans le site actif.

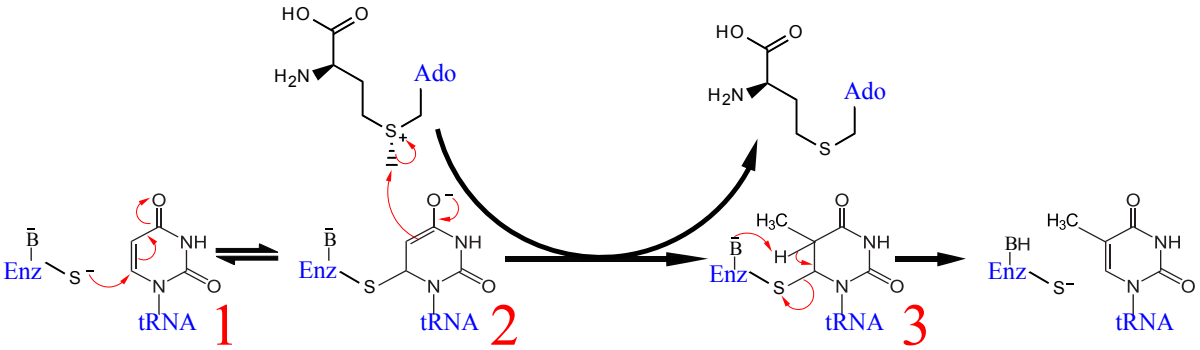

FIGURE 1.20 – **Mécanisme catalytique des méthyltransférases SAM dépendantes.** [36]. 1 : La cystéine 324 (numérotation d'*E.coli*) attaque le carbone 6 de l'uridine. 2 : Attaque d'un doublet sur le groupement méthyle du SAM. 3 : Une base de l'enzyme récupère l'hydrogène lié au carbone 5 de l'uridine permettant la formation d'une double liaison entre les carbones 5 et 6 de l'uridine et la libération du substrat.

les gènes codant pour ces enzymes appartenant à l'ensemble de gènes homologues COG2265 dont fait partie *trmA* :

- RlmC catalyse la méthylation de l'uridine 747 de la sous-unité 23s du ribosome [37].
- RlmD catalyse la méthylation de l'uridine 1939 de la sous-unité 23S du ribosome. La structure de RlmD a été résolue [38] et présente un groupement FeS de rôle inconnu.

On notera en outre que, chez *Bacillus subtilis*, ces deux modifications post-transcriptionnelles sont assurées par une même enzyme judicieusement baptisée RlmCD [39].

Plus tard, une recherche dans la banque de données de gènes orthologues COG (*Cluster of Orthologous Groups of proteins* [40] [41]) a permis d'identifier un ensemble de gènes homologues, baptisé COG1206, ayant avec *trmA* une distribution mutuellement exclusive [42]. De cet ensemble fait partie le gène *trmFO* codant pour une flavoenzyme ayant une activité de méthylation des ARNt [43] [44]. Cette découverte constituait une nouvelle occurrence de méthyltransférase flavine-dépendante similaire à ThyX, utilisant le MTHF comme donneur de méthyle et le NADPH comme agent réducteur, tour en ayant un mécanisme réactionnel légèrement différent.

Malgré l'absence de donnée structurale de TrmFO en présence de son substrat, on peut facilement envisager qu'une activation électrostatique de l'uridine à la manière de ThyX est difficilement concevable pour une macromolécule (remarque *a posteriori*, puisque le mécanisme catalytique de ThyX a été proposé après les considérations mécanistiques de TrmFO détaillées ci-après).

Après identification du gène *trmFO* et caractérisation biochimique de la protéine exprimée par celui-ci, la structure de TrmFO de *Thermus thermophilus* a été résolue par diffraction de rayons X avec une résolution de 1,6 Å [8] (Fig. 1.21), un test d'activité a été développé et a permis d'identifier, avec l'aide de mutagénèses dirigées, l'importance de certains résidus dans la catalyse [45].

Il a été clairement établi que deux cystéines (numérotées 51 et 223 chez *Thermus thermophilus*) avaient une importance cruciale dans la catalyse. Leurs rôles ont été établis ultérieurement chez *Bacillus subtilis* [46]. La cystéine 223 (numérotée 226 chez *Bacillus subtilis*) formant un intermédiaire covalent avec un mini-substrat contenant une 5-fluorouridine à la place de l'uridine cible, est fortement suspecté jouer le rôle de nucléophile activant le substrat.

Un intermédiaire covalent entre le noyau isoalloxazine de la flavine et la cystéine 53 (C51 chez *Thermus thermophilus*) a été identifié dans une préparation de la protéine de *Bacillus subtilis* fraichement purifiée qui s'est avéré capable de méthyler le substrat de l'enzyme [47]. Ces différentes observations ont naturellement mené à la proposition d'un mécanisme catalytique ; lequel est détaillé sur la figure 1.22 [48].

Outre *trmFO*, les protéines codés par deux gènes faisant partie du cluster COG1206 ont été caractérisées. La méthylation post-transcriptionelle de l'uridine 54 des ARNt de la bactérie *Mycoplasma capricolum* est inexistante, malgré la présence sur son

8. Ce manuscrit s'autorise l'usage de l'ångström (Å$= 10^{-10}m$), une unité bien pratique mais ne faisant pas partie du système international.

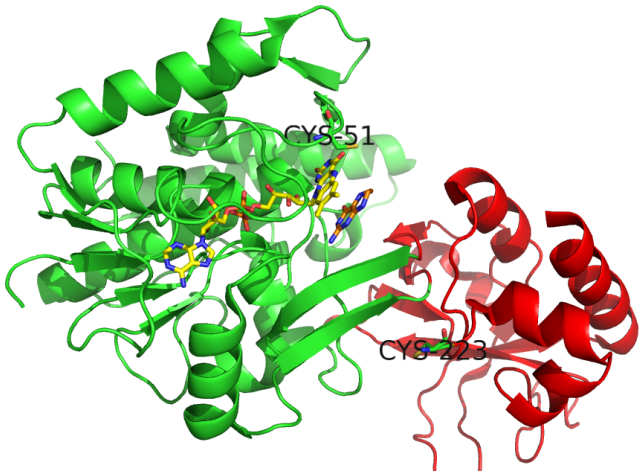

FIGURE 1.21 – **Structure de TrmFO de *Thermus thermophilus*** avec FAD et THF. La protéine s'organise en deux domaines, la cystéine 51 stabilise le groupement méthyle sur le noyau isoalloxazine et la cystéine 223 permet l'attaque nucléophile sur le carbone 5 de l'uridine cible.

chromosome de deux gènes homologues de *trmFO*. La mutation d'un de ces gènes a permis d'identifier la première enzyme flavine-dépendante catalysant la méthylation de l'uridine 1939 de l'ARNr 23S [49] mettant en évidence une certaine versatilité de reconnaissance de substrats dans l'évolution des méthyltransférases flavine-dépendantes. Versatilité non seulement dans la reconnaissance du substrat, mais aussi dans la réaction catalysée.

1.3.3 La méthylaminoéthylation de l'uridine 34 d'ARNt

Le code génétique a la particularité d'être dégénéré, chaque codon code pour un acide aminé spécifique, mais un acide aminé peut être codé par plusieurs codons différents. Différentes modifications de l'anticodon de certains ARNt permettent à un même ARNt de reconnaître différents codons. Une de ces modifications est la méthylaminoéthylation de l'uridine 34 par le complexe MnmEG. Le complexe MnmEG constitue un hétérodimère formé par les protéines MnmE et MnmG, cette dernière codée par un gène du groupe COG1206. Bien que catalysant des réactions différentes, les protéines TrmFO et MnmG partagent des caractéristiques communes. Des portions de alignement des séquences des deux protéines sont indiquées ci-dessous avec en rouge, trois résidus conservés d'une séquence à l'autre discutés dans les quelques paragraphes suivants (Fig. 1.23) :

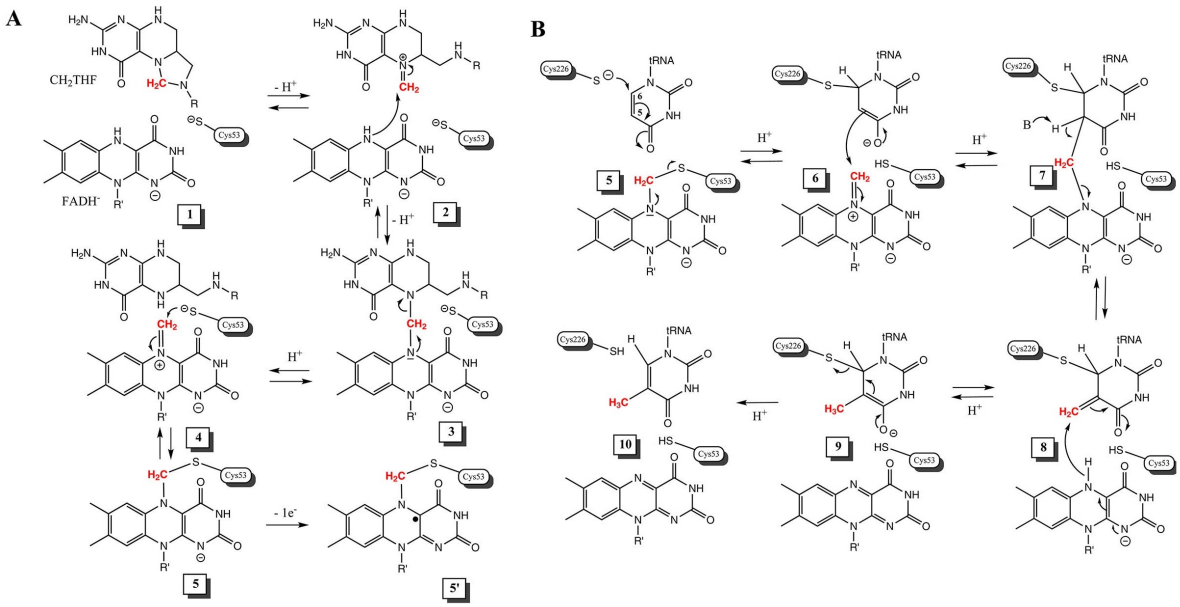

FIGURE 1.22 – **Postulat de mécanisme catalytique pour TrmFO.** La partie A détaille la formation d'un adduit C53-méthylène-FAD passant par la formation d'un adduit FAD-MTHF (3) suivie de l'élimination du tétrahydrofolate dont découle la formation d'un dérivé méthylène-imminium-FAD (4) et se terminant sur l'attaque du C53 sur ce dernier. On peut voir en seconde partie (B) le détail du mécanisme de méthylation de l'ARNt. L'adduit C53-méthylène-FAD est attaqué par le C5 du l'uridine préalablement activée par l'addition de la cystéine 226 sur le carbone 6. Il en résulte une décomposition du composé (5) en un dérivé méthylène imminium FAD permettant la formation d'un intermédiaire covalent entre le FAD et l'uridine de l'ARNt (7). La déprotonation du carbone 5 de l'uridine permettra ensuite la formation d'un groupement méthylène exocyclique sur l'uridine cible (8) qui sera ensuite converti en groupement méthyle par transfert d'hydride du FAD réduit [47].

Au même titre que TrmFO, on trouve dans la séquence de MnmG deux cystéines catalytiques numérotées 48 et 284 chez *Aquifex aeolicus* [50], suggérant des similitudes dans les mécanismes réactionnels de ces deux enzymes malgré le fait qu'elles catalysent des réactions différentes.

Chaque site actif des deux protéines (région à proximité de la flavine) comporte notamment une cystéine proche du noyau isoalloxazine et une tyrosine impliquée dans une liaison hydrogène via son groupement hydroxyle. Dans le cas de TrmFO, cette tyrosine est impliquée dans une interaction de type "$\pi - \pi$ stacking" avec le noyau isoalloxazine (Fig. 1.24). Le rôle de cette tyrosine chez TrmFO de *Thermus thermophilus* sera discuté plus loin.

La localisation de la flavine constitue une différence majeure. Le noyau isoal-

```
mnmG  Aa  10   VVVIGGGHAGIEAALAAARMGAKTAMFVL---------NADTIGQMSCNPAIGG----IA  56
               V V+G G AG EAA     R+G     +F +            D    ++ C+ ++GG     A
TrmFO Tt  4    VNVVGAGLAGSEAAWTLLRLGVPVRLFEMRPKRMTPAHGTDRFAEIVCSNSLGGEGETNA  63

mnmG  Aa  375  AGNFNGTTGYEEAAGQGIVAGINAALRAFGKEPIYLRRDESYIGVMIDDLTT  426
               AG   G  GY E+A  G +AG+NAA +A G  P+    +ES +G ++  L T
TrmFO Tt  334  AGVLAGVEGYLESAATGFLAGLNAARKALGLPPV-APPEESMLGGLVRYLAT  384

mnmG  Aa  270  ALYGGLIK--------GIGPRYC-----PSIEDKIVKFPDKERHQIFLEPEGLDTIEIYP  316
               ++ GGL++        G  P Y      P +E ++ K   KE+ Q      GL+    +
TrmFO Tt  373  SMLGGLVRYLATANPEGFQPMYANWGLVPPVEGRMGK---KEKRQAMYR-RGLEAFSAWL  428

mnmG  Aa  317  NGLSTSLP  324
               +GL+  LP
TrmFO Tt  429  SGLNPPLP  436
```

FIGURE 1.23 – **Alignement partiel de séquences des protéines MnmG de *Aquifex aeolicus* et TrmFO de *Thermus thermophilus*.**

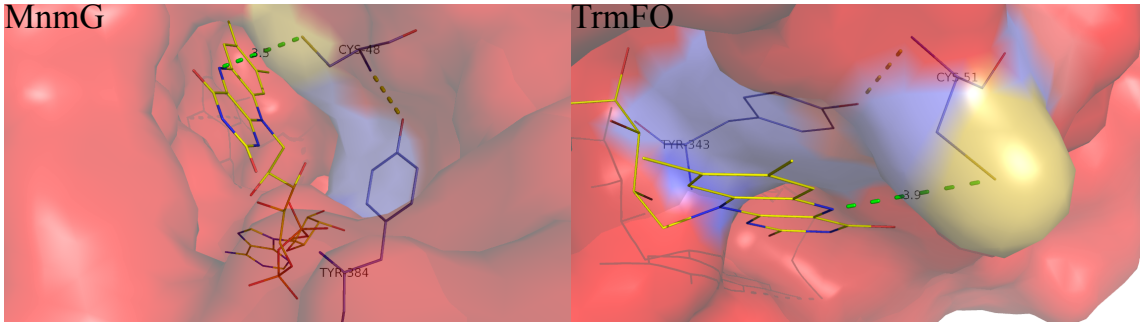

FIGURE 1.24 – **Comparaison des sites actifs de TrmFO de *Thermus thermophilus* (code PDB : 3G5R) et MnmG de *Aquifex aeolicus* (code PDB : 2ZXI).** Chez les deux enzymes, une tyrosine (Y343 pour TrmFO et Y384 pour MnmG)est impliquée dans une liaison hydrogène avec la l'azote d'une cystéine catalytique (C51 pour TrmFO et C48 pour MnmG).

loxazine est assez exposé pour TrmFO tandis qu'il est plutôt enfoui pour MnmG. MnmG étant fonctionnelle sous forme d'hétérodimère, on peut envisager un changement conformationnel de celle-ci lorsqu'elle interagit avec son partenaire, exposant son cofacteur FAD au substrat.

Les méthyltransférases flavine-dépendantes découvertes récemment suscitent un vif intérêt dans la mesure où leur caractérisation élargit le spectre des réactions catalysées par les flavoenzymes.

On notera deux aspects intrigants, leurs apparitions indépendantes sous deux formes différentes au cours de l'évolution en dépit de leurs faibles efficacités catalytiques et la redondance de leurs cibles (le noyau uracile).

1.4 Les flavoprotéines

Les flavoprotéines ont la caractéristique d'être liées à un groupement prosthétique de type flavinique. Ces derniers sont des dérivés de la riboflavine (vitamine B_2) ; il existe deux principaux types de cofacteurs flaviniques, les flavines mononucléotides (FMN) et les flavines adénine dinucléotides (FAD) dont les formules sont représentées en figure 1.25. Le précurseur des flavines, la riboflavine, est composée d'un noyau isoalloxazine lié à un ribitol par son azote 5. La FMN équivaut à une riboflavine phosphorylée sur son carbone 5' et la FAD, à une FMN liée à une adénosine.

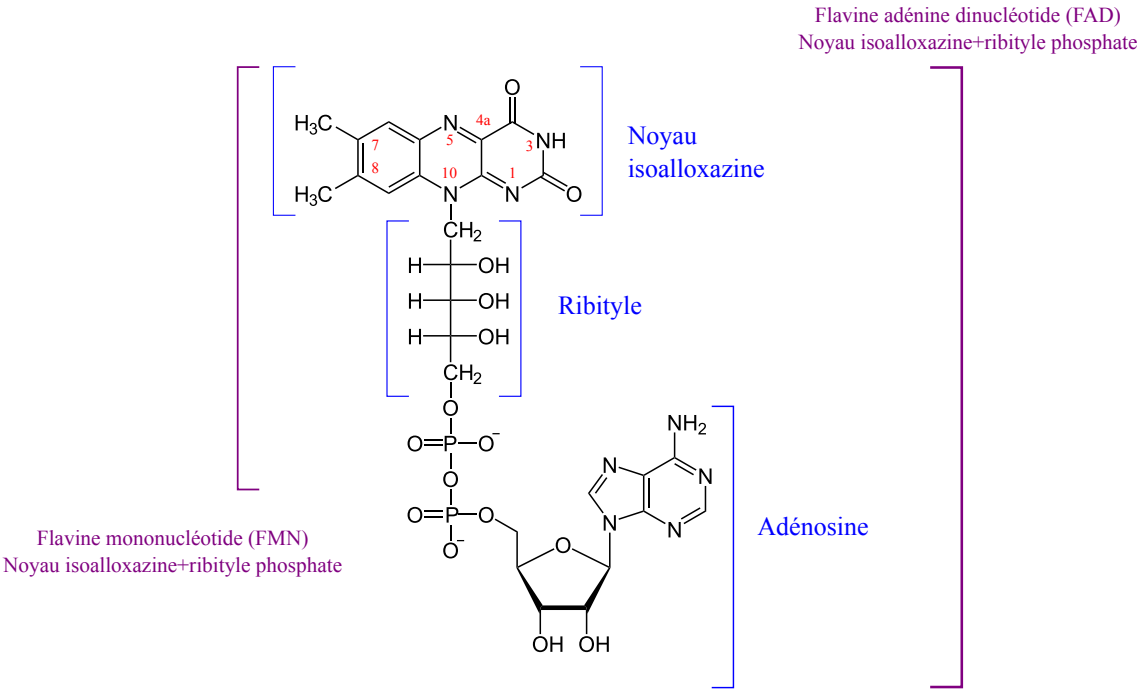

FIGURE 1.25 – **Formule des cofacteurs flaviniques.** La flavine mononucléotide (FMN) se décompose en un noyeau isoalloxazine et une chaîne ribityle phosphate tandis que la flavine adénine dinucléotide (FAD) se décompose en une FMN et une adénosine phosphate.

Les caractéristiques fonctionnelles des flavoprotéines leur sont conférées exclusivement par la partie isoalloxazine du cofacteur, les parties ribityle et adénosine ayant pour rôle principal l'interaction du cofacteur avec la ou les chaînes peptidiques.

1.4.1 Propriétés (photo)chimiques du noyau isoalloxazine

La réduction du noyau isoalloxazine affecte essentiellement la sous fonction "enediamine", aux positions N(1), C(4a) et N(5) avec un potentiel d'environ -200 mV

et par une réaction en une étape à deux électrons ou deux étapes à un électron ($FAD \rightarrow FADH_2$). Le noyau pyrimidique de l'hétérocycle peut être considéré comme un "puits à électron". Au sein du micro-environnement d'une flavoprotéine, des interactions entre ce noyau pyrimidine et des résidus qui tendent à réduire sa densité de charge négative augmenteront son potentiel d'oxydoréduction [51]. Ainsi, dans le vivant, les potentiels redox des flavoprotéines peuvent se situer entre -495 mV [52] et 80 mV [53].

On compte trois différents états d'oxydoréduction des flavines : la forme oxydée (flavoquinone), la forme réduite à un électron (flavosémiquinone) et la forme réduite à deux électrons (flavohydroquinone). Chacun de ces états peut se trouver sous forme cationique, neutre ou anionique donnant lieu à neuf formes de flavines différentes présentées en figure 1.26.

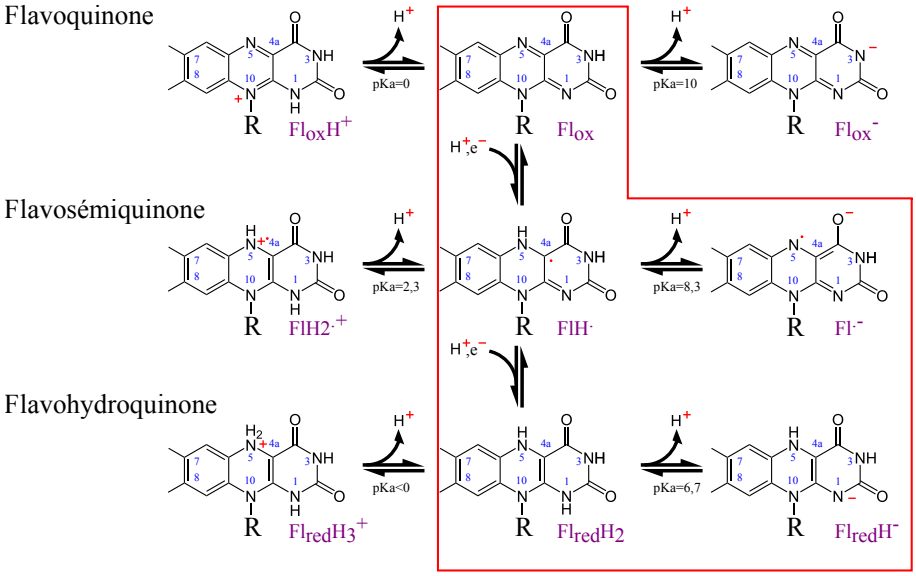

FIGURE 1.26 – **Formules des différentes formes (cationique, neutre et anionique) de flavines pour chacun de leurs états redox** d'après [54]. Les valeurs de pKa ne permettent qu'aux espèces dans l'encadrée rouge d'exister en conditions physiologiques.

En outre, le cycle phényl matérialisé par les carbones C(6) à C(9) confère aux flavines des propriétés de chromophore [55]. La présence de doubles liaisons conjugées au sein de ce cycle donne lieu à un nuage électronique délocalisé pouvant entrer en résonance avec un rayonnement incident et l'absorber.

Toutes ces caractéristiques ont permis d'identifier les flavoprotéines comme biosenseurs et comme catalyseurs de nombreuses réactions d'oxydoréduction, de transfert d'électrons, de déshydrogénation, de monooxygénation...

1.4.2 Caractéristiques spectrales

Les propriétés de chromophores des flavines les rendent particulièrement adaptées aux études spectroscopiques.

La forme oxydée (majoritairement observée dans les préparations de protéines) présente un spectre d'absorption avec deux pics (à 380 nm et à 450 nm). La flavoquinone présente une large bande vers 600 nm sous forme neutre et deux pics à 350 nm et 500 nm sous forme chargée négativement ; enfin, la flavine réduite a un maximum d'absorption vers 350 nm (Fig. 1.27).

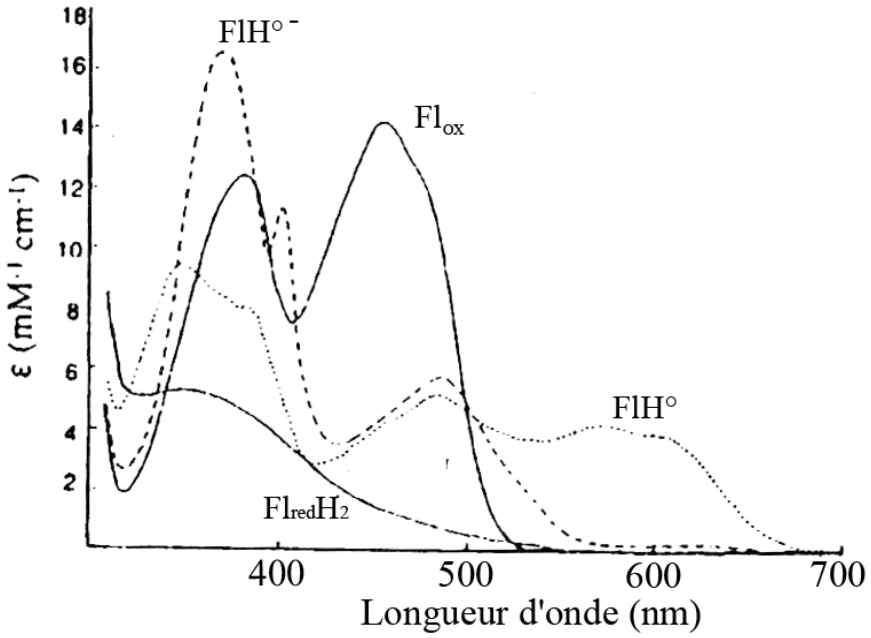

FIGURE 1.27 – **Spectres d'absorption des flavines en fonction de leur différents états d'oxydoreduction.** [56] Fl_{ox} : Flavoquinone neutre, $FlH^{\cdot -}$: Flavosémiquinone anionique, $FlH^{\cdot}$: Flavosémiquinone neutre, $Fl_{red}H_2$: Flavohydroquinone.

En plus de leurs propriétés d'absorption, les flavines sont fluorescentes. Les formes neutres des flavines fluorescent fortement avec un maximum d'émission à 520 nm, tandis que les formes chargées fluorescent peu [57].

Les propriétés de fluorescence et d'absorption des flavines peuvent être fortement modifiées au sein des protéines par la fixation de ligands ou le microenvironnement.

1.4.3 Monooxygénases et déshydrogénases

Les monooxygénases catalysent l'incorporation d'un atome d'oxygène depuis le dioxygène vers un substrat donné selon un mécanisme détaillé en figure 1.28.

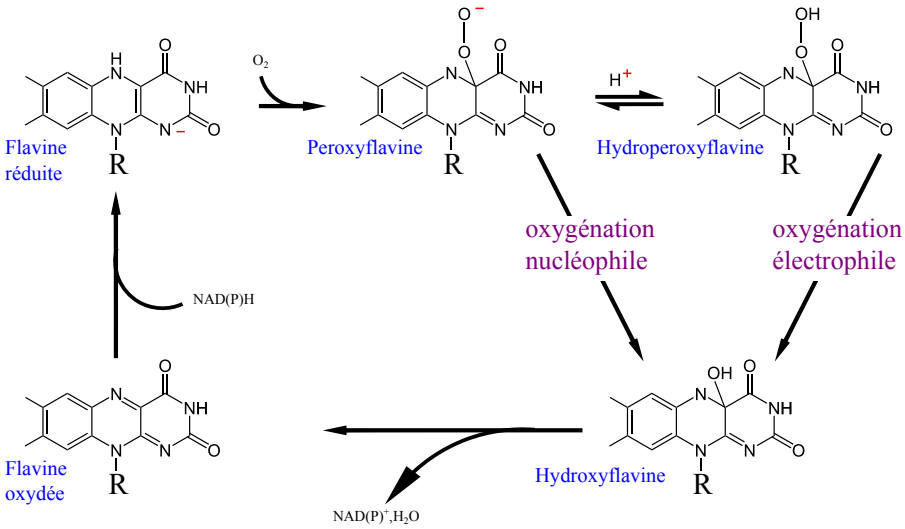

FIGURE 1.28 – **Mécanisme général des monooxygénases flavine dépendantes** d'après [58].

Les monooxygénations catalysées par les flavoprotéines peuvent être de différents types dont les principales sont :

- l'hydroxylation, soit un transfert de groupement OH (para-hydroxybenzoate hydroxylase [59])
- l'époxydation (styrene monooxygénase [60])
- le réarrangement de Baeyer-Villiger [9] [62] [63] (luciférase, cyclohexanone monooxygenase [64])
- l'oxydation d'amine
- l'oxydation de seleniure

Les déshydrogénases : le noyau isoalloxazine ayant deux sites de protonation, ces flavoenzymes sont connues pour catalyser la formation de doubles liaisons carbone-carbone. On les retrouve notamment dans le métabolisme énergétique (par exemple l'acyl-CoA déshydrogénase ou la succinate déshydrogénase), la flavine réduite pourra être ré-oxydée par la chaine respiratoire [65].

9. Les Baeyer-Villiger monooxygénases (BVMO) ont été beaucoup étudiées pour leurs application en synthèse organique [61]

1.4.4 Les biosenseurs flavoprotéiques

Les senseurs de lumière flavoprotéiques se subdivisent en trois principales classes : les protéines à domaines *Light Oxygen Voltage* (LOV), les *Blue Light Sensors Using FAD* (BLUF) et les cryptochromes [66].

Les protéines à domaine LOV se caractérisent par un squelette de cinq feuillets β antiparallèles (Aβ, Bβ, Gβ, Hβ et Iβ) et quatre hélices α les connectant les uns aux autres (Cα, Dα, Eα et Fα). Les protéines à domaine LOV sont présentes chez les plantes, les champignons, les archées et les procaryotes. Elles sont notamment impliquées dans le phototropisme, la relocalisation des chloroplastes et l'ouverture des stomates chez les plantes et dans le rythme circadien des champignons [67]. Caractérisées pour la première fois chez les phototropines des plantes [68] [69], ces protéines absorbent la lumière bleue et UV-A [68] via leur cofacteur flavinique. Après excitation de la flavine, se forme un adduit C4(a) cystéinyl avec une cystéine située sur l'hélice Eα au sein d'une séquence GXNCRFL(Q) conservée [70]. La formation de cet adduit s'effectue sur une durée comprise entre 100 ns et 5ms [71][72] et entraine un *blue shift* du pic d'absorbance vers 390 nm. Le retour à l'état fondamental se fait en 500 ms.

Les protéines à domaine BLUF ont été identifiées pour la première fois de deux manières indépendantes. D'une part, dans la régulation d'expression de gènes de photosynthèse chez la bactérie pourpre *Rhodobacter sphaeroides* [73] [74], d'autre part dans un mécanisme de phototaxie chez le protiste *Euglena gracilis* [75]. Les domaines BLUF se caractérisent pax un module *ferredoxin-like* de deux hélices α parallèles à une feuillet β de cinq brins [76]. Chez les protéines à domaine BLUF, la lumière bleue excite la flavine qui acceptera un électron d'une tyrosine à proximité (conservée chez ces protéines) pour former un bi-radical [77] résultant en un changement conformationnel [78] faisant intervenir une glutamine [79]. Ce changement de conformation induit un *red shift* de 10 nm du spectre d'absorption des BLUF étudiées à ce jour [80][81][82]. Cet état photoinduit peut perdurer jusqu'à plusieurs minutes.

Les cryptochromes sont les éléments clés du rythme circadien des plantes et des animaux. Ces protéines partagent une origine évolutive avec les photolyases, des enzymes de réparation de l'ADN dépendantes de la lumière [83]. Chez les photolyases au même titre que chez les cryptochromes, la flavine est liée à la chaîne peptidique via un pli Rossmann [66]. En absence de lumière, la flavine des cryptochromes est

oxydée *in vivo*. La lumière bleue induit l'accumulation de radicaux $FADH^{\cdot}$ et de flavines réduites $FADH^{-}$. La réoxydation des flavines se fait en plusieurs minutes.

La magnétoception : outre leurs rôles établis dans la perception de la lumière bleue, les cryptochromes ont récemment été proposés comme senseurs de champ magnétique chez certains eukaryotes "superieurs". Chez la drosophile, par exemple, une magnétoception dépendante de la lumière a été établie mettant en cause les cryptochromes [84]. Chez certains oiseaux migrateurs, l'orientation dans le champ magnétique terrestre implique l'activation de structures neuronales visuelles exprimant les cryptochromes [85].

1.5 Problématique

L'anabolisme des acides nucléiques dans le vivant actuel porte les stigmates de l'évolution du support génétique. Dans un premier temps sont synthétisés les ribonucléotides qui, après trois phosphorylations successives, peuvent être polymérisés en ARN aux fonctions diverses [10]. Lors de la réplication, les ribonucléotides seront réduits dans un premier temps (relique du passage de l'ARN à l'ADN) puis, dans un deuxième temps, l'uridine subira une méthylation sur son carbone 5 (passage de l'ADN "U" à l'ADN "T"). Cette méthylation présente un intérêt particulier dans la mesure où tous les organismes actuels intègrent la thymidine dans leur ADN. En outre, on peut observer dans le vivant, une méthylation des uridines sans réduction préalable, dans le cas des modifications post-transcriptionnelles des ARN. Comme détaillé précédemment, la méthylation de l'uridine est apparue au moins quatre fois au cours de l'évolution via quatre mécanismes catalytiques différents dont les principales caractéristiques sont détaillées sur le tableau 1.2.

La question ayant animé le projet dont fait état ce manuscrit est celle des contraintes évolutives ayant mené à l'établissement de cet ordre. Nous nous demanderons pourquoi la méthylation de l'uridine libre est systématiquement postérieure à sa réduction, et pourquoi la ribothymidine n'existe que sous forme d'ARN polymérisés, si une autre voie métabolique est envisageable. Nous chercherons à savoir si les mécanismes catalytiques de méthylation des uridines sont interchangeables, c'est-à-dire si TrmFO peut méthyler l'uridine libre.

10. Nous avons parlé dans cette introduction des ARNt, ARNr et ARNm mais il en existe bien d'autre ayant des rôles de régulation [86], impliqués notamment dans l'épissage des ARNm, dans la maturation des ARNr...

TABLEAU 1.2 – **Principales caractéristiques mécanistiques des différentes enzymes de méthylation des uridines.** La numérotation des cystéines catalytiques est celle de *Escherichia coli* pour ThyA et TrmA, et celle de *Bacillus subtilis* pour TrmFO.

	ThyA	ThyX	TrmA	TrmFO
substrat	dUMP	dUMP	ARNt U54	ARNt U54
donneur de méthyle(ne)	MTHF	MTHF	SAM	MTHF
réducteur	THF/DHF	FAD/NADPH	0	FAD/NADPH
activation de l'uracile	C146	électrostatique	C324	C226
intermédiaire de réaction	U-CH2-S	FAD-CH2-U	U-CH2-S	FAD-CH2-U

La première partie de ce manuscrit décrira une approche métabolique et de biologie synthétique. La complémentation d'un phénotype d'auxotrophie pour la thymidine par de la ribothymidine a été étudiée et la construction d'un double mutant a permis d'identifier la voie permettant cette complémentation. Des expériences d'évolution dirigée par PCR mutagène et synthèse de banque de mutants combinatoire ont été mises en place, visant à transformer TrmFO, une ribothymidylate synthase en thymidylate synthase artificielle.

Dans un second temps, la protéine TrmFO a été étudiée plus en détail. Les rôles de différents résidus dans la fixation de la protéine à un substrat minimal ont pu être déterminés par anisotropie de fluorescence. Une approche spectroscopique d'absorption, de fluorescence à l'état stationnaire ou résolue en temps a pu aboutir à une meilleure compréhension du mécanisme catalytique de TrmFO et des méthyltransférases flavine-dépendantes en général. L'observation et la caractérisation de ce qui pourrait être un intermédiaire réactionnel sera discutée.

Chapitre 2

Matériels et méthodes

2.1 Techniques de biologie moléculaire

2.1.1 Liste des souches d'*E.coli*

TABLEAU 2.1 – **Liste des souches d'*E.coli***

Souche	Génotype	Source
MG1655	F-, λ^-, rph-1	[87]
FE013	MG1655 $\Delta thyA$::*aphA-3*	[22]
STELLAR	$F–, endA1, supE44, thi-1, recA1, relA1, gyrA96,$ $phoA, \Phi F80dlacZ\delta M15, \delta(lacZYA-argF)$ $U169, \delta(mrr-hsdRMS-mcrBC), \delta mcrA, \lambda–$	Clontech
$\Delta deoA$	MG1655 $\Delta deoA$::FLP	[88]
Double mutant	MG1655 $\Delta deoA$::FLP $\Delta thyA$::*aphA-3*	ce travail
XL1	endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[: :Tn10 proAB$^+$ lacI^q Δ(lacZ)M15] hsdR17(r$_K^-$ m$_K^+$)	Agilent Technologies
BL21 DE3	B F$^-$ *ompT gal dcm lon* $hsdS_B(r_B^-m_B^-)$ $\lambda(DE3[lacI\ lacUV5\text{-}T7p07\ ind1\ sam7\ nin5])$ $[malB^+]_{K-12}(\lambda^S)$	[89]

MG1655 est la souche sauvage d'*E.coli* de référence, la première à avoir été séquencée [87]. La souche **FE013** est auxotrophe pour la thymidine, elle a été utilisée pour les tentatives de sélection de gènes complémentant ce phénotype. L'ADN génomique de FE013 a aussi servi de matrice à l'amplification de la cassette d'inac-

tivation du gène *thyA* chez le simple mutant $\Delta deoA$.

La souche **STELLAR®** (Clontech) est optimisée pour le clonage et a été utilisée pour une première transformation de produits de mutagénèse aléatoire après clonage.

Le simple mutant $\Delta deoA$ a été utilisé pour la construction du double mutant, afin de procéder au test de complémentation du phénotype $\Delta thyA$ par des analogues nucléosidiques. La souche a été fournie par Olivier Tenaillon du laboratoire *Infection, Antimicrobials, Modelling, Evolution.*

XL1 (Agilent Technologies) a servi à la transformation des produits de mutagénèse dirigée.

La souche **BL21** est mutée pour certaines protéases, et compte le gène codant pour l'ARN-polymérase du bactériophage T7 [1], sous contrôle du promoteur de l'opéron lactose, la rendant optimale pour l'expression de protéines recombinantes [89].

2.1.2 Composition des milieux de culture

Le milieu LB (*lysogeny broth*)[90] est un milieu riche composé d'extrait de levure (5 g/L), de NaCl (10 g/L) et de lysat de caséine (tryptone 10 g/L). L'extrait de levure contient la plupart des composés nécessaires à la croissance des souches d'*E.coli* délétés pour les gènes non essentiels

Le milieu SOC [91] est est généralement utilisé pour augmenter les efficacités de transformation des plasmides. Il est composé de 2% de tryptone, 0,5% d'extrait de levure, 8,56 mM de NaCl, 2,5 mM de KCl, 10 mM de $MgCl_2$ et 10 mM de $MgSO_4$, 20 mM glucose.

Le milieu minimal M9 est composé des éléments chimiques nécessaires et suffisants à la croissance d'*E.coli* de type sauvage. Dans le cas des travaux présentés ici, il a permis d'identifier des seuls évènements de complémentation de phénotypes d'auxotrophie (incapacité à synthétiser un composé essentiel). On y trouve, pour cinq litres d'eau déminéralisée, 33.9 g de Na_2HPO_4, 15 g de KH_2PO_4, 5g de NH_4Cl et 2.5 g/ de NaCl. Après stérilisation de la solution par autoclave, 20 mL de glucose 20% et 2 mL de $MgSO_4$ 1 M (stérilisés par filtration à 2 μm) ont été ajoutés à 1 L de solution pour obtenir des concentrations finales de 0,4 % et 2mM respectivement.

Si nécessaire, différents antibiotiques, l'ampicilline (Amp), la kanamycine (Kan)

1. Cette caractéristique n'a pas été exploitée ici, le plasmide d'expression utilisé ayant un promoteur T5, reconnu par l'ARN polymérase d'*E.coli*

ou le chloramphénicol (Cm) ont été ajoutés aux milieux de cultures à des concentration finales de 100 μg/mL, 40 μg/mL et 20 μg/mL respectivement.

2.1.3 Cartes des plasmides utilisés

Le plasmide pKOBEG [92] compte une origine de réplication ori101 et le gène repA101^{ts} permettant la répression de cette dernière à haute température, le gène *cat* de résistance au chloramphénicol, le gène *araC* codant pour le répresseur de l'opéron arabinose dont le répresseur contrôle l'expression des protéines Bet, Exo et Gam (Fig. 2.1). Ces gènes permettent la recombinaison chez *E.coli* d'ADN linéaires et ainsi l'inactivation de gènes par insertion de cassette.

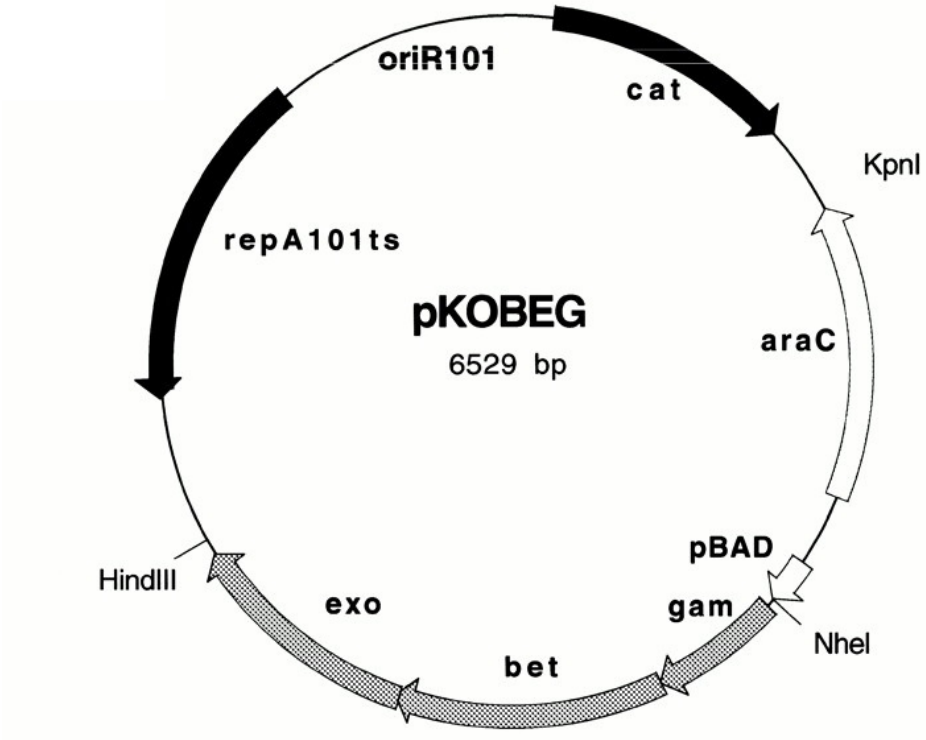

FIGURE 2.1 – **Carte du plasmide pKOBEG** utilisé pour l'inactivation de gènes chez *E.coli* par insertion de cassette.

Le plasmide pCP-20 [88] permet l'excision des gènes flanqués par des régions FRT par l'expression de la recombinase FLP (spécifique de ces régions). pCP-20 compte une origine de réplication thermosensible (au même titre que pKOBEG), un gène de résistance à l'ampicilline (*bla* qui code pour la β-lactamase) et un gène *cat* de résistance au chloramphénicol (Fig. 2.2).

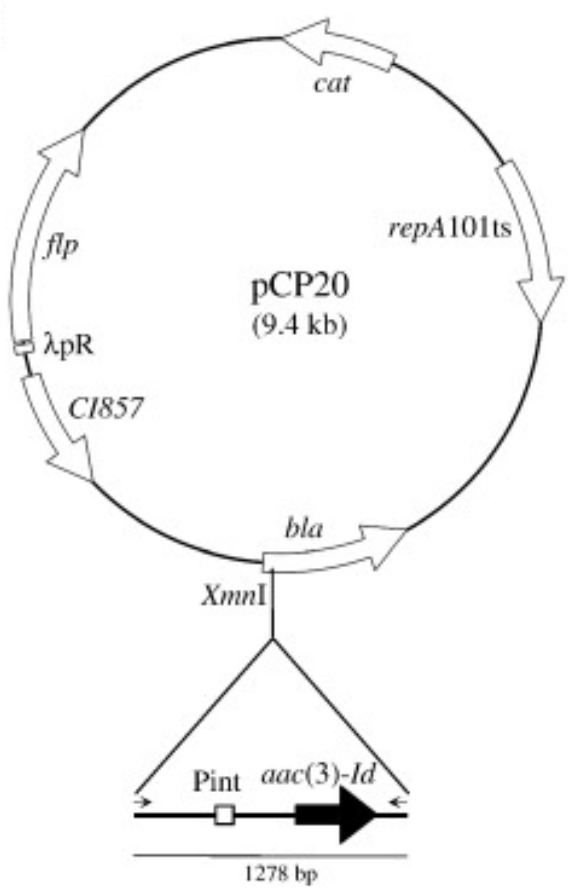

FIGURE 2.2 – **Carte du plasmide pCP-20** codant pour la flipase FLP qui catalyse l'excision de cassettes d'inactivation flanquées de régions FRT.

Le plasmide d'expréssion pQE-80L (QUIAGEN) comprend une origine de réplication chez *E.coli* (ColE1), un gène de résistance à l'ampicilline permettant de sélectionner sa transformation dans les cellules compétentes, le gène $laqI^q$ codant pour le répresseur catabolique de l'opéron lactose. Dans le cas de ce plasmide, le répresseur LacI peut se fixer sur le promoteur T5 [2] situé en amont du site multiple de clonage (MSC) grace aux séquences opératrices lacO. Le système permet l'induction par l'IPTG de l'expression du gène cloné dans le MCS (Fig. 2.3).

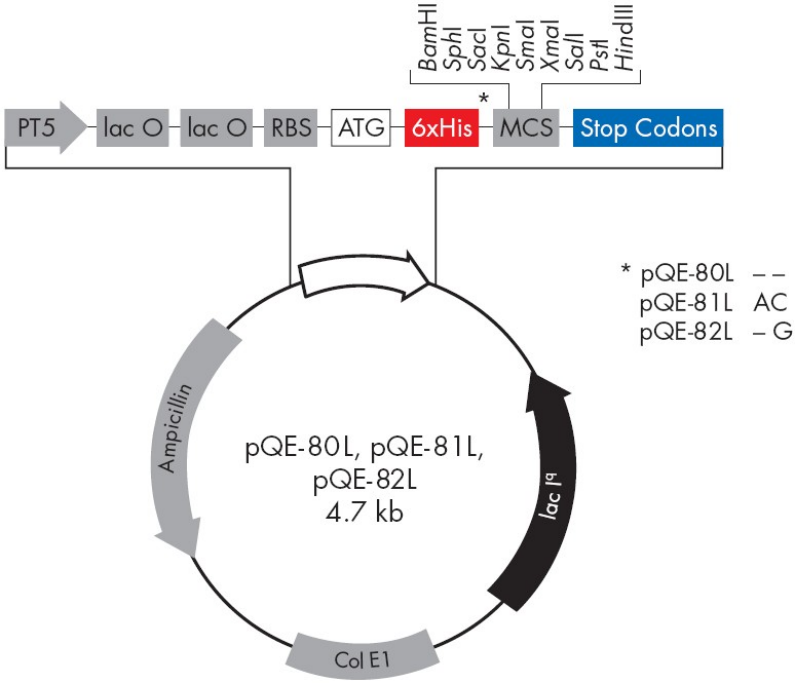

FIGURE 2.3 – **Carte du plasmide d'expression pQE-80L.**

2.1.4 Transformation bacterienne

La transformation bacterienne permet d'introduire dans une bactérie de l'ADN étranger. Deux techniques de transformation ont été utilisées au cours de ce travail.

2. le promoteur T5 est reconnu par l'ARN polymérase native d'*E.coli*, ainsi, ce plasmide peut être utilisé pour l'expression de protéine dans n'importe quelle souche.

Transformation par choc thermique : Jusqu'à 5 μL d'une solution de plasmide à concentration variable (entre 10 et 200 ng/μL) ont été ajoutés à 50 μL de bactéries chimio-compétentes dans un tube Eppendorf de 1,5 mL, le mélange a été incubé 30 min sur glace pour ensuite être soumis à un choc thermique d'exactement 1 min à 42°C suivi d'une incubation de 2 min sur glace. Après quoi, 450 μL de milieu *Super Optimal broth with Catabolite repression* (SOC) [91] ou de LB ont été ajoutés aux cellules et incubées 1 h à 37°C sous agitation à 300 rpm dans un bain sec. Après transformation, les suspensions de cellules ont été étalées sur milieu sélectif supplémenté en antibiotique ad hoc.

Transformation par électroporation : Jusqu'à 5 μL d'une solution d'ADN à concentration variable (entre 10 et 200 ng/μL) ont été ajoutés à 50 μL de bactéries électrocompétentes dans une cuve d'électroporation pour être soumises à un choc électrique de 2500 V, le contenu a ensuite été resuspendu dans 450 μL de milieu LB pour une incubation d'1 h à 37°C sous agitation à 300 rpm dans un bain sec. Dans le cas des sélections de gène complémentant le phénotype d'auxotrophie pour la thymidine, de l'Isopropyl β-D-1-thiogalactopyranoside (IPTG) a été ajouté à une concentration finale de 1 mM pour permettre l'expression des éventuels gènes de complémentation avant sélection. Enfin, la transformation a été étalée sur milieu sélectif.

2.1.5 Préparation de cellules d'*Escherichia coli* électrocompétentes

Les bactéries électrocompétentes sont destinées à être transformées avec de l'ADN (plasmidique ou autre) par un choc électrique de 2500 V. Pour 1 L de culture, 1 L de milieu *lysogeny broth* (LB) [90] a été inoculé par 25 mL d'une préculture et laissé incubé à 37°C sous agitation à 150 rpm jusqu'à une $DO_{600} = 0,6 - 0,75$. Dans le cas de FE013 (Tab. 2.1), de la thymidine a été ajoutée au milieu à une concentration finale de 40 μg/mL, la souche étant auxotrophe pour la thymidine [3]. La culture ainsi en phase exponentielle a été refroidie sur glace pendant 10 min. Pour la suite du protocole, les cellules ont été conservées à 4°C tout au long de la suite du protocole afin de les figer en phase exponentielle.

3. Dans le cas de FE013, le manque de thymidine limite la division cellulaire. On arrive alors relativement vite à une phase stationnaire bien que la DO_{600} continue d'augementer, du fait d'un allongement des cellules et l'efficacité de transformation s'en trouve largement diminuée.

Les cellules ont ensuite été lavées trois fois par 10 min de centrifugation à 5000 rpm et resuspension par 200 mL, 150 mL, 100 mL d'une solution de 10% de glycérol stérile maintenu à 4°C tout au long du processus. Cette étape permet de débarasser la culture des sels conducteurs contenus dans le milieu afin d'éviter un arc électrique lors de l'étape de transformation, et la mort des cellules. Enfin après une dernière centrifugation, les cellules ont été resuspendues dans un volume final de 2 mL de cette même solution de glycérol, aliquotées par 100 μL et stockées à -80°C jusqu'à utilisation. On peut noter qu'une culture de 1 L mène à 2 mL de bactéries compétentes, et que la culture est concentrée 500 fois.

2.1.6 Inactivation du gène *thyA* chez la souche Δ*deoA*

La cassette d'inactivation du gène *thyA* par le gène *aphA-3* de résistance à la kanamycine [22] a été amplifiée par réaction de polymérisation en chaîne (PCR) a partir de 10 ng d'ADN génomique de la souche FE013 extrait à l'aide du kit NucleoSpin® Tissue de Macherey-Nagel par lyse alcaline et adsorption sur membrane de silice. La *Pfu* DNA polymerase de Promega® a été utilisée à cette fin, avec les paramètres de thermocycleur suivants :

94°C 2 min	
94°C 30 sec	30X
55°C 20 sec	
68°C 2 min	
68°C 5 min	

Le produit de cette amplification a été purifié par affinité à la silice à l'aide du kit NucleoSpin® Gel and PCR Clean-up de Macherey-Nagel pour ensuite être transformé dans une souche Δ*deoA* électrocompétente et recombinante préparée comme il suit.

100 mL de LB+Cm ont été inoculés par 5 mL de préculture de la souche Δ*deoA* (gène codant pour la thymidine phosphorylase) portant le plasmide pKOBEG [92]. L'expression des gènes *bet*, *exo* et *gam* organisés en opéron sur le plasmide pKOBEG sous contrôle du promoteur pBAD a été induite par 0,2% d'arabinose à $DO_{600} = 0,15$. La culture a ensuite été refroidie 10 minutes sur glace une fois la DO_{600}=0,87 atteinte. La culture a ensuite été soumise à trois lavages par 50 mL glycérol à 10% et resuspendue dans 300 μL de glycérol 10%. 60 μL de cellules compétentes ont été transformées par 160 ng d'ADN *aphA-3* amplifié précédemment, étalées sur boîte LB+Kan (40 μg/mL) et incubées à 37°C pendant une nuit.

2.1.7 Vérification des génotypes Δ*deoA* et *thyA*::*aphA-3*

L'excision de la cassette *kan* de la souche Δ*deoA*::*kan* a été vérifiée par PCR sur colonie pour trois colonies. La One*Taq*® DNA polymerase de chez New England Biolabs a été utilisée à cette fin, avec comme paramètres de thermocycleur :

94°C 2 min
94°C 30 sec
55°C 20 sec — 30X
68°C 2 min
68°C 5 min

La vérification de l'inactivation de *thyA* par la cassette *aphA-3* a été faite dans des conditions similaires à deux différences près : un temps d'élongation de 1 min 30 et un temps d'hybridation de 40 sec.

2.1.8 Évolution dirigée de TrmFO

L'évolution dirigée permet entre autre la sélection de nouvelles activités catalytiques via un processus similaire à l'évolution par sélection naturelle. Une expérience d'évolution dirigée consiste en une étape de génération de diversité et une étape de sélection.

Ce travail décrit les tentatives d'obtenir une thymidylate synthase artificielle à partir de l'enzyme TrmFO. Afin d'y parvenir, deux techniques de génération de diversité ont été mises en application : la synthèse d'une banque de mutant combinatoire pour six lysines d'une part, et la PCR mutagène d'autre part.

La banque de mutants combinatoire a été synthétisée par la société DNA 2.0, les séquences ayant été optimisées pour l'expression de la protéine TrmFO de *Bacillus subtilis* chez *E.coli*.

Les différentes PCR *error prone* ont été réalisées à l'aide du kit GeneMorph II Random Mutagenesis d'Agilent Technologies ; un kit de PCR dont la polymérase a un faible biais de mutagénèse. La matrice pour la PCR a été le plasmide pQE-80L dans lequel était inséré le gène *trmFO* de *Bacillus subtilis*, de *Thermus thermophilus* ou la banque de mutants combinatoire.

Différentes quantités d'ADN matrice (entre 10 et 200 ng) ont été utilisées pour faire varier le taux de mutations avec les paramètres de thermocycleur suivant :

94°C 2 min
94°C 30 sec | 30X
55°C 20 sec |
72°C 2 min |
72°C 5 min

Le produit de PCR mutagène a été soumis à une électrophorèse sur gel préparatif d'agarose 1%+BET, 30 min à 100V. La portion du gel contenant le produit de PCR a pu être identifiée par une révélation aux UV et isolée du reste du gel pour ensuite être purifiée par chromatographie d'affinité sur membrane de silice.

Le produit de PCR purifié a ensuite été cloné dans le vecteur pQE-80L par une double digestion *Bam*HI/*Hind*III (du vecteur et d'une part et du produit de PCR d'autre part) de 15 min à 37°C, suivi d'une inactivation de 10 min à 85°C (enzymes *Fast digest* de Thermo Fisher). Le vecteur digéré a ensuite été déphosphorylé (protocole *Antarctic phosphatase* NEB). L'insert et le vecteur déphosphorylé ont été mélangés à masse égale (pour un rapport molaire d'environ 1/4) évaluée au Nano drop, et la ligature a été effectuée 15 min à température ambiante. En parallèle, la même quantité de vecteur déphosphorylé a subi le même traitement à titre de témoin. Les différents produits de ligatures ont été transformé dans la souche STELLAR et étalés sur LB+Amp. Après incubation d'une nuit à 37°C, les colonies ont été soumises à une extraction plasmidique, les plasmides ont été transformés dans la souche FE013 pour effectuer la sélection.

2.1.9 Liste des amorces PCR

Les amorces suivantes ont été utilisées pour vérifier l'excision de la cassette *kan* inactivant le gène *deoA*, la vérification du phénotype $\Delta thyA$ ou la mutagénèse aléatoire 2.4. L'amorce H50 s'hybride sur la cassette *apha-3* de résistance à la kanamycine, pQE FW et pQE REV s'hybrident de part et d'autre du site de clonage du plasmide pQE-80L.

SK15 :	5’ GGCTCCCGTACAGAAGATATT 3’
SK16 :	5’ CCCTGATAACGTTGATACTGT 3’
SK72 :	5’ CGGGATCCATGAAACAGTATTTAGAACTG 3’
SK73 :	5’ CCCAAGCTTTTAGATAGCCACCGGCGCTTT 3’
H50 :	5’ CCGGTGATATTCTCATTTTAGCC 3’
deoA FW :	5’ CATGATGGAAGTGATCCGTG 3’
deoA REV :	5’ GGCCTTTACAACCGTTATCA 3’
pQE FW :	5’ CCCGAAAAGTGCCACCTG 3’
pQE REV :	5’ CTTCTGAGGTCATTACTGG 3’

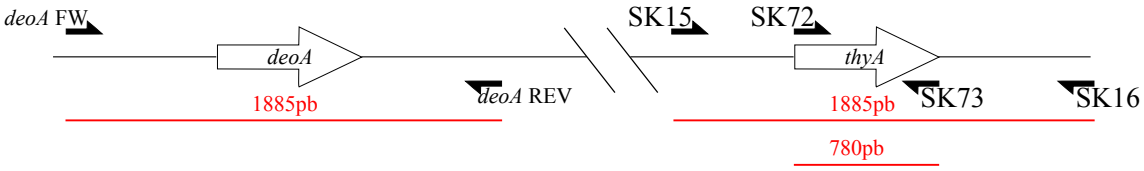

FIGURE 2.4 – **Hybridation des différentes amorces sur le chromosome d’*E.coli* de type sauvage.**

2.1.10 Mutagénèse dirigée

Les différentes mutagénèses dirigées ont été effectuées en amplifiant par PCR le plasmide entier pQE-80L-*trmFO Tt* à l’aide d’oligonucléotides portant les mutations d’intérêt. Les différentes PCR ont été réalisées a l’aide du kit QuikChange II Site-Directed Mutagenesis de chez Agilent Technologies et les amorces ont été élaborés à l’aide du site (www.genomics.agilent.com/primerDesignProgram.jsp) avant d’être synthétisés par la société EUROGENTEC.

Chaque réaction d’amplification a été réalisée dans un volume final de 50 μL, contenant 50 ng d’ADN matrice, 125 ng de chaque oligonucléotide, 1 μL d’une solution de dNTP fournie par le kit et 1μL d’ADN polymérase *Pfu* Ultra HF. Les paramètres utilisés pour le thermocycleur ont été :

95°C 2 min
95°C 30 sec
58°C 1 min
68°C 7 min
} 16X

Un traitement par l’endonucléase de restriction *Dpn*I après amplification a permis d’éliminer spécifiquement l’ADN matrice. La présence d’amplicons a été vérifiée par migration de 4 μL de chaque produit de PCR avec 1 μL de tampon de charge 5X

dans un gel d'agarose 1% contenant du bromure d'éthidium (BET) à 100V pendant 30 min. Les produits d'amplification ont ensuite été transformés dans des cellules chimio-compétentes de la souche XL1-Blue par choc thermique et les cellules ont été étalées sur LB+Amp. La souche XL1-Blue permet la recircularisation du produit d'amplification en un plasmide circulaire fonctionnel. Les transformants ont ainsi pu être soumis à une extraction d'ADN plasmidque et envoyés à séquencer. Les plasmides vérifiés par séquencage ont ensuite pu être transformés dans la souche d'expréssion BL21 en vue d'une expression et purification des protéines.

2.1.11 Liste des amorces utiliséés pour la mutagénèse dirigée

Y343F for :	5'-CGGGTGTTGAAGGCTGGCTGGAATCTGCGGCG-3'
Y343F rev :	5'-CGCCGCAGATTCCAGCCAGCCTTCAACACCCG-3'
C51A for :	5'-CCCAGGCTGTTGCTGGCGACGATCTCAGCGAA-3'
C51A rev :	5'-TTCGCTGAGATCGTCGCCAGCAACAGCCTGGG-3'
H308G for :	5'-CAGGTAGGTGTTGCGACCCATAACGCCGTAGCGA-3'
H308G rev :	5'-TCGCTACGGCGTTATGGGTCGCAACACCTACCTG-3'
C51S for :	5'-CAGGCTGTTGCTGCTGACGATCTCAGCGA-3'
C51S rev :	5'-TCGCTGAGATCGTCAGCAGCAACAGCCTG-3'
C51V for :	5'-GCCACCCAGGCTGTTGCTAACGACGATCTCAGCGAAACG-3'
C51V rev :	5'-CGTTTCGCTGAGATCGTCGTTAGCAAGAGCCTGGGTGGC-3'
C223A for :	5'-CCAGTTCTTCCACCGGCACCGCCGCTTCGAAAAATTCCAGT-3'
C223A rev :	5'-ACTGGAATTTTTCGAAGCGGCGGTGCCGGTGGAAGAACTGG-3'
W214F for :	5'-TCGAAAAATTCCAGTTTTTCAAAATCGTGTGGGGTGTGACGC-3'
W214F rev :	5'-GCGTCACACCCCACACGATTTTGAAAAACTGGAATTTTTCGA-3'
K282A for :	5'-CGTTTCTGTTCCGGCCACGCCAGACCCGTCTGGAAGC-3'
K282A rev :	5'-GCTTCCAGACGGGTCTGGCGTGGCCGGAACAGAAACG-3'
K287A for :	5'-GTCTGAAATGGCCGGAACAGGCGCGCCTGATCCAGATGATTCC-3'
K287A rev :	5'-GGAATCATCTGGATCAGGCGCGCCTGTTCCGGCCATTTCAGAC-3'
K409A for :	5'-GGTAGAAGGCCGTATGGGCGCGAAAGAAAAACGTCAAGCG -3'
K409A rev :	5'-CGCTTGACGTTTTTCTTTCGCGCCCATACGGCCTTCTACC -3'
K410A for :	5'-ATACATCGCTTGACGTTTTTCCGCCTTGCCCATACGGCCTTCTAC -3'
K410A rev :	5'-GTAGAAGGCCGTATGGGCAAGGCGGAAAAACGTCAAGCGATGTAT -3'

2.1.12 Séquence du gène de synthèse codant pour TrmFO de *Thermus thermophilus*

L'encodage et la synthèse du gène ont été réalisés par la société Genecust. Les sites de restriction *BamH*I et *Sma*I ayant permis le sous clonage dans le vecteur d'expression pQE-80L sont représentés en bleu, les codons initiateur et STOP le sont en rouge (Fig. 2.5).

GGATCCATGAATCAGCAAACTGTAAACGTTATTGGTGCTGGTCTGGCGGGTAGCGAAGCGGCTTGGCAGCTGGCGAAACG
TGGTATCCAGGTAAAACTGTACGAGATGCGTCCGGTTAAGCAGACTCCGGCACACCATACGGATAAATTCGCGGAACTGG
TATGTAGCAACTCTCTGCGTTCCAACACTCTGGCGAACGCGGTGGGCGTCCTGAAAGAAGAAATGCGTGCCCTGGACTCT
GCCATCATCGCAGCAGCGGACGAATGCTCCGTACCGGCGGGTGGTGCGCTGGCTGTAGATCGTCATGAGTTCGCTGCGAG
CGTTACTAACCGTGTGAAAAATCACCCAAACGTAACTGTTATTAACGAAGAAGTTACCGAAATCCCGGAGGGTCCGACGA
TCATCGCTACCGGCCCACTGACCAGCGAATCTCTGAGCGCACAGCTGAAGGAACTGACCGGCGAGGACTATCTGTACTTT
TACGACGCAGCGGCGCCGATCGTAGAAAAAGACTCCCTGGACATGGATAAAGTCTACCTGAAATCTCGCTATGACAAAGG
CGAGGCAGCCTACCTGAACTGTCCTATGACCGAAGAAGAGTTCGATCGTTTCCACGAAGCACTGACGTCTGCCGAAACTG
TCCCGCTGAAAGAATTTGAAAAGGAAATCTTCTTCGAAGGCTGCATGCCGATCGAAGTTATGGCAAAACGTGGCAAAAAA
ACGATGCTGTTCGGCCCGATGAAACCTGTGGGCCTGGAACATCCGGTTACTGGCAAACGTCCGTACGCTGTGGTCCAGCT
GCGCCAAGATGACGCAGCCGGTACCCTGTATAACATCGTTGGTTTTCAGACCCACCTGAAATGGGGTGATCAGAAAGAAG
TGCTGAAACTGATTCCGGGTCTGGAGAACGTTGAAATCGTGCGCTATGGTGTTATGCACCGTAACACCTTTATTAACTCT
CCGTCCCTGCTGAAGCCGACTTACCAGTTCAAAAACCGTAGCGATCTGTTCTTTGCTGGCCAGATGACCGGTGTGGAAGG
CTATGTGGAATCCGCGGCTAGCGGTCTGGTTGCTGGTATTAACGCTGCTAAACTGGTTCTGGGTGAAGAGCTGGTTATTT
TCCCGCAGGAAACCGCTATCGGCTCTATGGCCCACTACATCACCACCACCAACCAGAAAAACTTCCAGCCAATGAACGCC
AATTTCGGTCTGCTGAAAGAGCTGCCGGTGAAAATCAAGAATAAAAAGGAGCGCAACGAACAGTACGCGAACCGCGCAAT
TGAAACCATCCAAACCATTTCCAAGACCATCTAACCCGGG

FIGURE 2.5 – **Séquence codant pour TrmFO de *Thermus thermophilus*.** les sites de restriction *Bam*HI et *Sma*I sont en bleu et les codons initiateur et stop sont représentés en rouge.

2.2 Techniques d'études biochimiques

2.2.1 Expression et purification de TrmFO et de ses mutants

Expression : 500 mL de LB+Amp (100 μg/mL) ont été inoculés par 10 mL de préculture en phase stationnaire et laissé incubés à 37°C sous agitation. Une fois la phase exponentielle atteinte ($D0_{600} = 0,6 - 0,8$) l'expression protéique a été induite par 1 mM d'IPTG en concentration finale pendant 3h ; après quoi la culture a été culotée par une centrifugation de 20 min à 5000 rpm pour être resuspendue dans 5 mL de tampon *lysis equilibrium wash* (LEW : 50 mM NaH_2PO_4, 300 mM NaCl, pH 8,0) et stockée à -80°C jusqu'à l'étape de purification.

Les suspensions bactériennes ont été traitées 30 min sur glace à la DNase et au lysozyme (une pointe de spatule de chaque pour 5 ml de suspension bacterienne) pour ensuite subir 5 cycles de 20 secondes de sonications à 70% et 20 secondes sur glace afin de compléter la lyse.

Afin d'optimiser le ratio holoenzyme/apoenzyme, le lysat a été supplémenté en FAD 100 μM final et laissé 10 min sur glace afin de favoriser la fixation du FAD à la protéine. Une centrifugation d'une heure à 13200 rpm a ensuite permis de ne récupérer que la fraction soluble des protéines du lysat, contenant la protéine d'intérêt.

Purification : La fraction soluble (environ 4,5 mL) a été déposé sur une colonne *Protino*® Ni-TED Resin 2000 (Macherey Nagel) pré équilibrée par 5 mL de tampon LEW puis récupéré, La colonne a ensuite été lavée par 25 mL de tampon LEW et la protéine a été éluée dans 3 mL de tampon d'élution (LEW + imidazole 250 mM).

Cette dernière étape a été réitérée après lavage de la colonne par 5 mL de tampon LEW avec la fraction de protéines solubles récupérée préalablement après premier passage. La protéine éluée a été dessalé par chromatographie d'exclusion sur colonne de polyacrylamide Econopac 10DG (bio-rad) et récupérée dans 4 mL de tampon HEPES 50 mM NaCl 300 mM pH8 avant d'avoir été concentré sur filtre AMICON® à 10 kDa d'exclusion de masse moléculaire.

Afin d'évaluer la pureté des différents mutants, les purifications ont été soumises à une migration de 30 min à 100V sur gel d'acrylamide 10% en conditions dénaturantes (SDS-PAGE) suivi d'une révélation à l'instantBlue (Expedeon).

2.3 Approches spectroscopiques

2.3.1 Anisotropie de fluorescence

L'anisotropie de fluorescence permet de suivre la fixation d'un fluorophore à une protéine via le changement de la vitesse rotationnelle de ce dernier. Ce suivi peut être effectué par le calcul du rapport d'intensités d'émission de fluorescence polarisée verticalement et horizontalement.

Les expériences d'anisotropie de fluorescence ont été réalisées dans des cuves en quartz de 1x1 cm contenant 1 mL d'une solution du substrat minimal ribonucléique à 10 nM dans du tampon HEPES 20 mM, NaCl 100 mM. Le mini-substrat (correspondant au motif $T-\Psi$ de l'ARNt^{Phe} de *Thermus thermophilus*) portant en 5' le

fluorophore Alexa 532 a été excité par une lumière de 530 nm de longueur d'onde avec une fente de 10 nm et l'émission a été enregistrée entre 540 et 600 nm avec une valeur de photo-multiplicateur de 850V et un échantillonage de 1 nm.

L'ARN non structuré portant le fluorophore CY5 a été excité par une lumière de 650 nm de longueur d'onde avec une largeur de 10 nm et l'émission a été enregistrée entre 600 et 700 nm avec une valeur de photo-multiplicateur de 850V et un échantillonage de 1 nm. Sauf indication contraire, les mesures ont été effectuées à 45°C et les valeurs d'anisotropie ont été calculées en appliquant l'équation à l'intégrale des courbes de spectre d'émission de fluorescence. Pour le tracé des courbes d'anisotropie en fonction de la concentration en protéine, les préparations de TrmFO ont été dosés par la méthode de Bradford [93].

2.3.2 Spectroscopie d'absorption

Les mesures de spectroscopie d'absorption à l'équilibre ont été réalisées à l'aide d'un spectromètre Shimadzu dans des cuves de 1 mm de chemin optique.

2.3.3 Photoconversion de la protéine TrmFO et de ses différents mutants

Les expériences d'illumination ont été réalisées dans des cuvettes de quartz de 1 mm de chemin optique. Une lampe de microscope de 700 mW couplée à une fibre optique a servi de source de lumière, réglée à puissance réduite (au tiers environ) sur une durée totale de 2 à 3 heures, en prenant une mesure du spectre d'absorption toutes les 15 minutes.

2.3.4 Chromatographie en phase liquide couplée à un spectromètre de masses

Une préparation de TrmFO de type sauvage a été dialysée deux fois une heure contre une solution de NaCl à 100 mM afin d'éliminer le tampon et de faire baisser la concentration en sel, concentrée puis diluée deux fois pour obtenir une concentration finale en NaCl de 50 mM. La protéine a ensuite été précipitée par trois cycles de chauffage à 99°C 15 min et centrifugation de 20 min à 13000 rpm et élimination du culot. Le surnageant a ensuite été analysé par temps de vol en électrospray positif par

Sophie BOURCIER du Laboratoire de Chimie Moléculaire de l'École Polytechnique.

2.3.5 Fluorescence résolue en temps

La fluorescence résolue en temps a été étudiée chez TrmFO et différents mutants en utilisant un montage de spectroscopie ultra-rapide de type *Kerr gate* (Fig. 2.6). Les différents milieux de Kerr utilisés ont été le benzène, le suprasil ou le CS_2. Le décalage spatial du LASER est directement lié à l'espacement temporel des mesures à une échelle de la centaine de femto secondes. Au moins trente mesures ont été effectuées pour une dynamique donnée avec un milieu de Kerr donné.

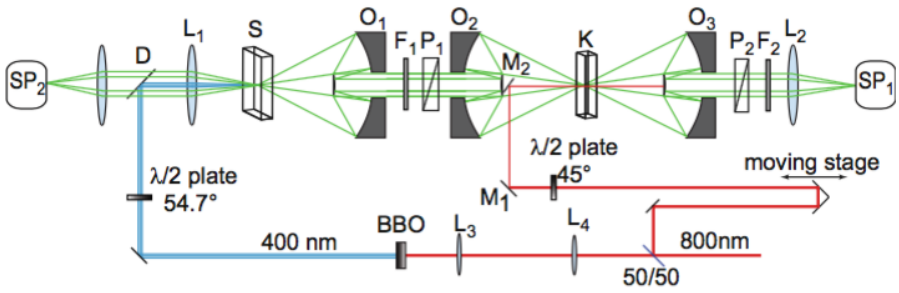

FIGURE 2.6 – **Schéma du montage de spectroscopie ultrarapide.** Le milieu de Kerr sert d'obturateur. Le BBO permet le changement de la longueur d'onde de 800 nm à 400 nm afin d'exciter l'échantillon. Les temps d'obturation et de réponse sont définis par les milieux de Kerr utilisés et la partie amovible permet de moduler le moment à des intervals de l'ordre de la centaine de fs.

2.4 Informatique

L'utilitaire d'optimisation d'encodage des gènes pour la mutagénèse aléatoire a été écrit en Java Script par Simon Dadoun sous licence GPL ; il est accessible sur le site (www.lob.cnrs.fr/mutagenesis/sms2/mutagenesis_bias_fraction.html) mais n'est pas encore tout-à-fait fonctionnel.

Les figures de structures protéiques de cet écrit ont été réalisées à l'aide du logiciel PyMOL, les différentes courbes ont été tracées à l'aide du logiciel OriginPro8 et les figures et schémas ont été créées grâce au logiciel Inkscape.

Chapitre 3

Métabolisme des pyrimidines chez *E.coli*

Chez *E.coli*, le métabolisme des pyrimidines (Fig. 3.1) reflète assez bien l'évolution du support génétique décrit en introduction. Les différentes voies de biosynthèse donnent lieu, dans un premier temps aux ribonucléotides ; ceux-ci peuvent être intégrés dans l'ARN après trois phosphorylations successives. Pour la réplication du programme génétique, ces ribonucléotides seront réduits en 2'désoxyribonucléotides par les ribonucléotide réductases. C'est uniquement après cette étape de réduction qu'intervient enfin la méthylation du dUMP en dTMP.

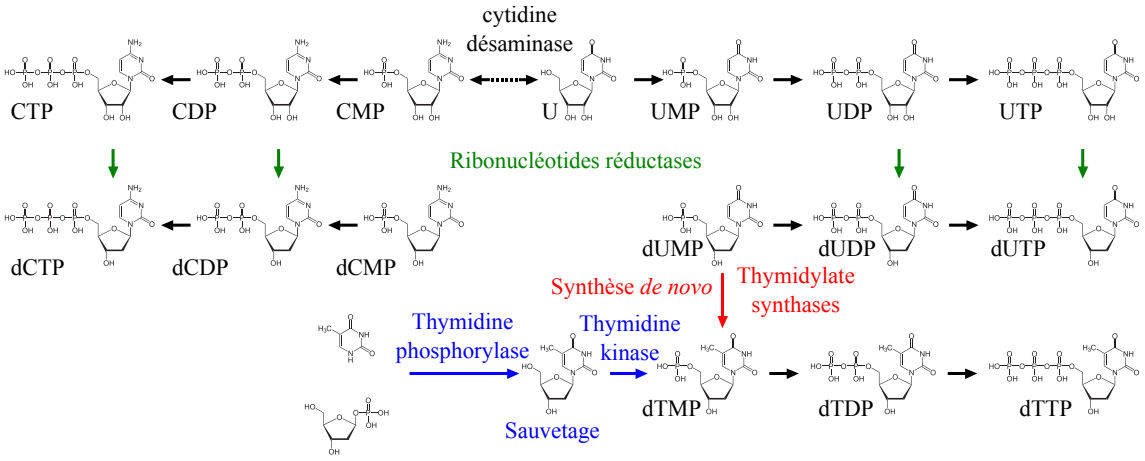

FIGURE 3.1 – **Représentation schématique du métabolisme des pyrimidines.** Les flèches vertes matérialisent l'action des ribonucléotides réductases, les flèches bleues, la voie de sauvetage du phénotype auxotrophe pour la thymidine et la flèche rouge, la synthèse *de novo* du thymidylate.

Dans ce chapitre, je propose une possibilité alternative de méthylation de l'uridine antérieure à sa réduction qui donnerait lieu à de la ribothymidine 5'monophosphate ou 5-méthyluridine (5mU), un composé ne faisant *a priori* pas partie du métabolisme du vivant. Mes réflexions à ce sujet s'articuleront autour de deux axes : la complémentation du phénotype d'auxotrophie pour la thymidine par des analogues de nucléotides dans un premier temps, et la tentative de création d'une thymidylate synthase artificielle par évolution dirigée dans un second temps.

3.1 Voie de complémentation du phénotype Δ*thyA* par des analogues nucléosidiques

3.1.1 Complémentation du phénotype Δ*thyA* par la ribothymidine (5mU) et la méthylcytidine (5mC)

La méthylation sur le carbone 5 de la 2'désoxyribo-uridine 5'monophosphate produisant la 2'désoxyribo-thymidine 5'monophosphate est systématiquement postérieure à la réduction du ribose par les ribonucléotide réductases, empêchant la formation de ribothymidine (5mU) dans la cellule. Cependant, j'ai pu mettre en évidence une promiscuité *in vivo* d'enzymes du métabolisme des pyrimidines d'*Escherichia coli* permettant la complémentation du phénotype Δ*thyA* d'une souche nommée FE013 [22], par des analogues nucléosidiques (Fig. 3.2) comme la 5mU et la 5mC (Tab. 3.1).

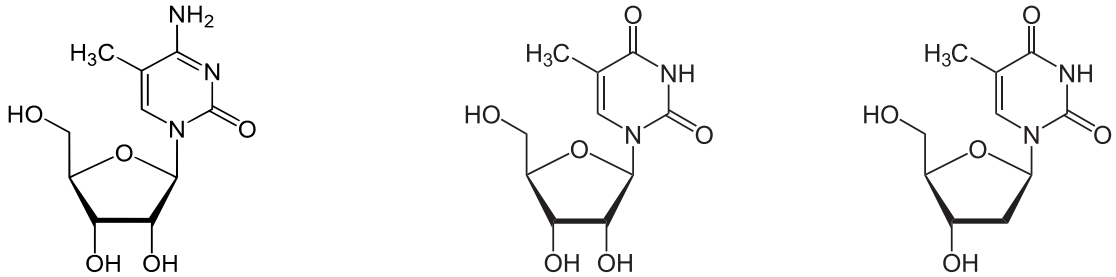

FIGURE 3.2 – **Structures des analogues nucléosidiques** et comparaison avec la thymidine classique (2'désoxyribo-thymidine).

Se soulève alors la question des enzymes permettant cette complémentation. L'hypothèse testée et validée en section suivante est l'implication de la thymidine

TABLEAU 3.1 – **Complémentation de la souche FE013 par des analogues nucléosidiques.** Sur milieu minimal supplémenté en thymidine, la souche auxotrophe pousse, au même titre que sur les analogues nucléosidiques 5mU et 5mC.

	M9	M9+dT	M9+5mC	M9+5mU
FE013	0	++	+	+
MG1655	+++	+++	+++	+++

phosphorylase (E.C.2.4.2.4) codée par le gène *deoA*, une enzyme catalysant, chez *E.coli*, la phosphorolyse de la thymidine en désoxyribose-1-phosphate et en thymine (Fig. 3.3). Il a été préalablement établi qu'un double mutant Δ*thyA* Δ*deoA* peut se développer en présence de thymidine, mais pas en présence de thymine [94].

Les analogues nucléosidiques pourraient subir chez *E.coli* une phosphorolyse dans un premier temps, produisant du ribose-1-phosphate et de la thymine. Cette thymine serait dans un deuxième temps fixée sur un désoxy-ribose-1-phosphate par la thymidine phosphorylase produisant *in fine* de la 2'-désoxythymidine.

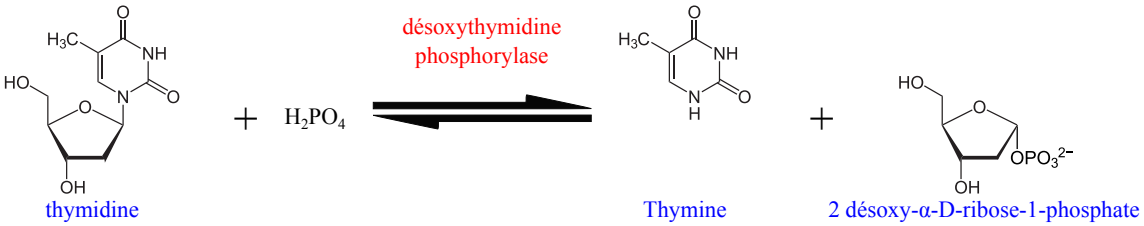

FIGURE 3.3 – **Réaction catalysée par la désoxythymidine phosphorylase.** Sur milieu minimal seul, la souche auxotrophe pour la thymidine ne pousse pas (par définition), tandis que la souche sauvage pousse.

Pour tester cette hypothèse, j'ai construit un double mutant pour les gènes *thyA* et *deoA* comme décrit ci-après.

3.1.2 Construction du double mutant Δ*deoA*Δ*thyA*

Le double mutant a été construit à partir d'une souche Δ*deoA*::*kan* issue de la banque KEIO, une collection de 3985 mutants d'*E.coli* K12 pour des gènes non essentiels, remplacés par une cassette de résistance à la kanamycine [88]. Cette construction a été faite en deux étapes (Fig. 3.4) :

— L'excision de la cassette *kan* de résistance à la kanamycine (Kan) via le plasmide pCP-20 [95]

— L'inactivation du gène *thyA* chez *E.coli* par insertion d'une autre cassette de résistance à la kanamycine *apha-3*

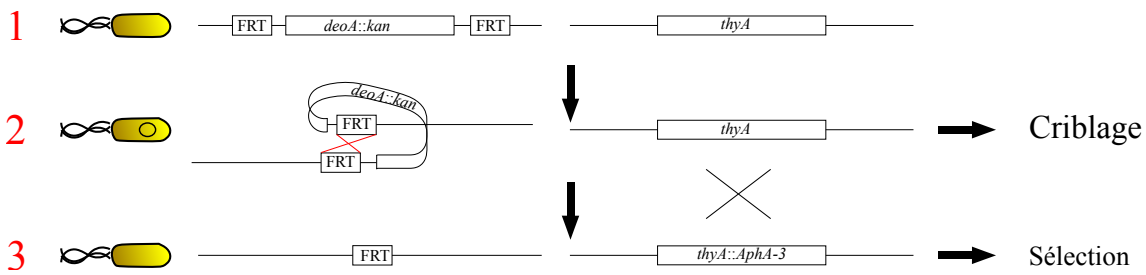

FIGURE 3.4 – **Stratégie de construction du double mutant Δ*deoA*Δ*thyA*.** La souche Δ*deoA*::*kan* issue de la banque KEIO (1) a été transformée par le plasmide pCP-20 codant pour la flipase FLP qui permet l'excision de la cassette d'inactivation *kan* (2). Enfin, la souche alors sensible à la kanamycine peut être mutée pour le gène *thyA* par la cassette aphA-3 (3).

La cassette d'inactivation du gène *deoA* (1323 pb) est flanquée par des régions FRT, ciblées par la recombinase FLP codée par le plasmide pCP-20, une construction comportant une origine de réplication thermosensible et des gènes de résistance au chloramphénicol (Cm) et à l'ampicilline (Amp). En principe, le plasmide se réplique à 30°C et ne se réplique pas à 42°C, seulement, le taux de mutations spontanées chez *E.coli*, de concert avec le nombre important de copies de plasmides dans une culture, mène à un faible nombre de plasmides thermorésistants. Ces plasmides seront néanmoins sélectionnés si la souche qui les porte est mise en culture à de plus hautes températures. C'est pourquoi il faut prendre soin de systématiquement mettre en culture à 30°C et non 37°C les souches d'*E.coli* qui portent un plasmide thermosensible avant de s'en débarrasser à 42°C.

Le plasmide pCP-20, a été transformé dans la souche d'*E.coli* STELLAR^{TM} (Clonetech) [1], laquelle a été étalée sur LB+Amp+Cm et incubée à 30°C deux nuits. Afin de s'assurer de la thermosensibilité du plasmide, quatre colonies ont été remises en culture liquide sans antibiotique à 30°C d'une part et à 42°C d'autre part. Les cultures ont été striées le lendemain sur LB+Amp+Cm et comparées après une nuit à 37°C. Un tapis de cellules a pu être observé sur chaque strie issues des cultures à 30°C, tandis que quelques colonies isolées ont été distinguées sur les stries issues des cultures à 42°C. Les résistances aux antibiotiques ont donc été majoritairement perdues après culture à 42°C, indiquant un non maintient du plasmide dans la souche

1. Une souche *recA1* optimisée pour les transformations et le clonage.

et sa bonne thermosensibilité. Les cultures liquides a 30°C ont été récupérées pour une extraction d'ADN plasmidique.

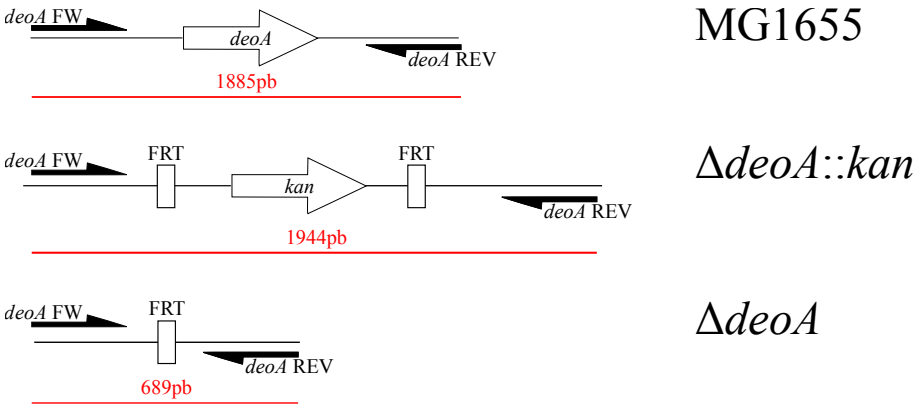

FIGURE 3.5 – **Stratégie de vérification de l'excision de la cassette *kan* par PCR.** Les amorces s'hybrident à quelques centaines de nucléotides en amont et en aval du gène, la taille du produit d'amplification permet de discriminer entre la présence et l'excision de la cassette *kan*.

Le plasmide pCP-20 a ensuite été transformé dans la souche Δ*deoA*::*kan*, laquelle a été étalée sur LB+Amp+Cm et incubée deux nuits à 30°C. Huit colonies ont été repiquées sur LB+Kan pour cribler la perte de résistance à la kanamycine. Trois clones ayant passé le criblage avec succès ont ensuite servi de matrice pour une PCR sur colonie afin de s'assurer de l'excision de la cassette (Fig. 3.5). Une colonie n'ayant pas été transformée par le plasmide pCP-20 a été utilisée comme témoin et les produits de PCR ont migré sur gel d'agarose pour en comparer les tailles (Fig. 3.6).

On observe chez le témoin, une bande correspondant à 1885 pb, taille attendue pour l'amplification de la cassette chez la souche d'origine. En revanche, pour chaque colonie testée, on peut observer une bande correspondant à la migration d'un fragment de 689 pb, taille attendue pour la cassette excisée (Fig. 3.5). On peut conclure que l'excision a bien été efficace.

La souche Δ*deoA* sensible à la kanamycine ainsi obtenue, il est devenu possible d'en inactiver le gène *thyA* par insertion de la cassette de résistance à la kanamycine *aphA-3*. Pour cela, le plasmide pKOBEG [96] a été transformé dans cette souche afin d'en faire des cellules compétentes ayant exprimé les protéines de recombinaison du phage gamma *bet*, *exo* et *gam*, permettant l'inactivation d'un gène par recombinaison homologue. D'autre part, la cassette *aphA-3* a été amplifiée à partir de l'ADN génomique de FE013.

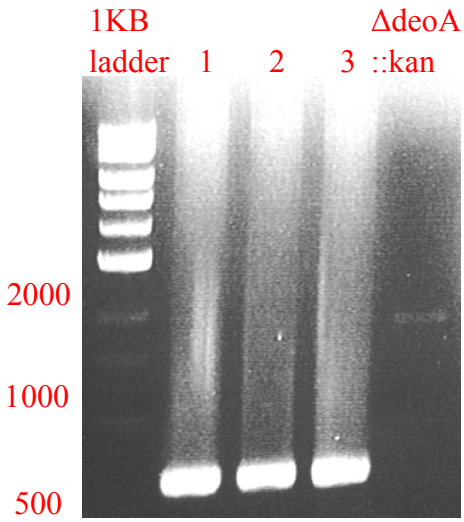

FIGURE 3.6 – **Vérification de l'excision de la cassette *kan* par migration sur gel d'agarose des produits de PCR.** 1kb ladder : marqueur de poids moléculaire, Δ*deoA*::*kan* : colonie non transformée par pCP-20, 1, 2 et 3 : colonies testées. Une bande est présente pour le témoin positif aux alentours de 2000 pb, attendue pour l'amplification de la cassette. Pour chaque colonie testée, on distingue une bande de 689 pb attendue pour l'amplification de la région après excision de la cassette.

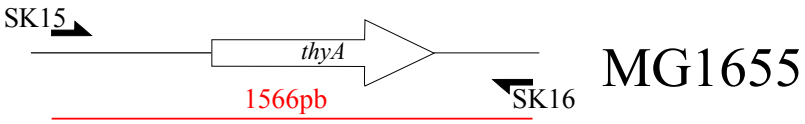

FIGURE 3.7 – **Amplification de la cassette *aphA-3* à partir de l'ADN génomique de la souche FE013 flanquée par les régions en aval et en amont du gène *thyA*.**

Le produit de cette amplification a enfin été transformé dans la souche Δ*deoA*, portant le plasmide pKOBEG qui code pour les protéines de recombinaison du phage lambda Bet, Exo et Gam. La protéine Gam empêche la dégradation des ADN linéaires double brin par les nucléases RecBCD et SbcCD d'*E.coli*. L'exonucléase Exo dégrade l'ADN double brin de 5' vers 3' en ADN simple brin ; enfin, la protéine Beta se fixe sur l'ADN simple brin issu de l'action d'Exo, et favorise l'appariement homologue de la cassette d'inactivation avec la région ciblée du chromosome [97] [98]. La souche ainsi transformée a été étalée sur LB+Kan. Deux colonies obtenues après incubation une nuit à 37°C ont été testées par PCR pour l'inactivation du gène *thyA* par insertion de la cassette.

La PCR sur une colonie de la souche K12 MG1655 ne montre aucun signal pour le couple d'amorces SK15/H50, comme attendu, la souche K12 n'ayant pas de cassette

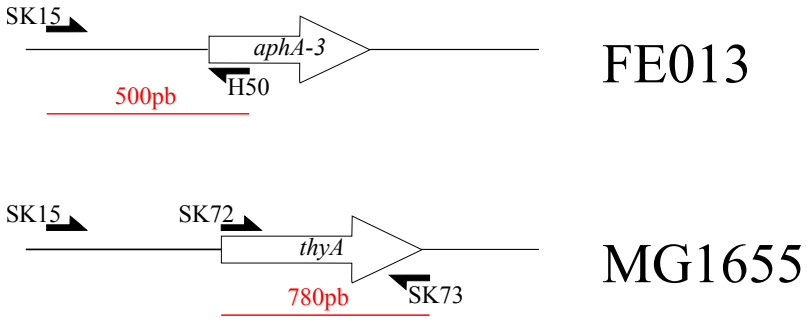

FIGURE 3.8 – **Stratégie de vérification de l'insertion de la cassette *aphA-3* dans le gène *thyA*.** L'amorce SK15 s'hybride en amont du gène *thyA* ou *aphA-3*, suivant la construction, l'amorce H50 s'hybride sur le gène *aphA-3* et l'amorce SK73 à la fin du gène *thyA*.

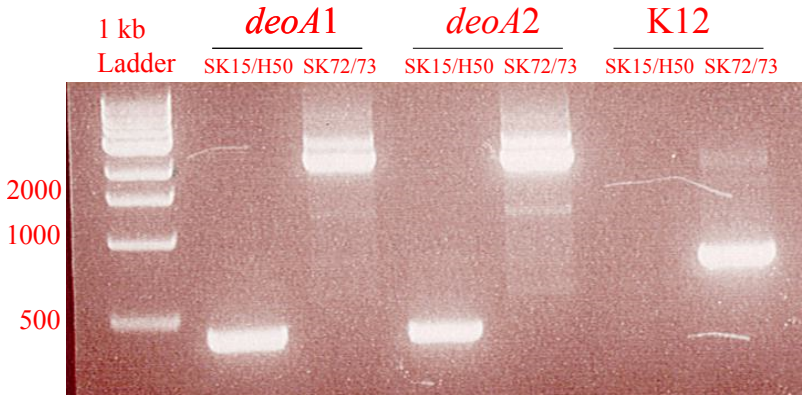

FIGURE 3.9 – **Vérification de l'inactivation du gène Δ*thyA* par la cassette *aphA-3*** par migration sur gel d'agarose des produits de PCR sur colonie. 1 kb ladder : marqueur de poids moléculaire, *deoA*1 et 2 : colonies testées, K12 : contrôle (souche sauvage). Les amorces SK15, SK72, SK73 et H50 sont représentées en figure 3.8.

aphA-3. Pour le couple SK72/SK73, on observe une bande à 780 pb correspondant à la région ciblée, ainsi qu'une bande d'intensité plus faible aux alentours de 3000 pb, que l'on attribue à une amplification non spécifique d'une région non ciblée du chromosome. Pour les deux colonies testées, on retrouve cette même bande avec une intensité plus importante, mais pas de signal spécifique de la région ciblée. Pour le couple SK15/H50, on observe une bande de 500 pb, correspondant à l'amplification d'une partie de la cassette *aphA-3* et à la région en aval. Cela montre que la cassette est bien présente sur le chromosome à l'endroit ciblé et que l'inactivation du gène *thyA* a bien été efficace (Fig. 3.9).

3.1.3 Tests de complémentation des différents mutants par des analogues nucléosidiques

Le double mutant Δ*deoA*Δ*thyA*::*aphA-3* ainsi validé a été strié sur différentes boîtes de milieu minimal M9 supplémenté ou non en analogues nucléosidiques (Fig. 3.10). Le tableau 3.2 résume le résultat de la complémentation.

TABLEAU 3.2 – **Récapitulatif des résultats de complémentation du phénotype d'auxotrophie pour la thymidine** des souches MG1655, Δ*deoA*, FE013 et Δ*deoA*Δ*thyA*::*aphA-3* par la thymidine (dT) la thymine, la ribothymidine (5mU) et la 5-méthylcytidine (5mC).

	MG1655	Δ*deoA*	FE013	Δ*deoA*Δ*thyA*::*aphA-3*
M9	+	+	0	0
M9+dT	+	+	+	+
M9+thymine	+	+	+	0
M9+5mU	+	+	+	0
M9+5mC	+	+	+	0

La souche sauvage MG1655 se développe sur toutes les boîtes au même titre que la souche Δ*deoA*, ce qui montre qu'aucun des composés n'est toxique et que la seule délétion de *deoA* n'empêche pas le développement des bactéries (comme attendu, le gène n'étant pas essentiel). En revanche, FE013 ne pousse pas sur du milieu M9 seul, comme attendu, mais se développe en présence de chaque composé. Les analogues nucléosidiques 5mU et 5mC peuvent pallier l'absence de thymidylate synthase au même titre que la thymidine et la thymine. Le double mutant Δ*deoA*Δ*thyA*::*aphA-3* peut se développer sur M9+thymidine, mais pas sur M9+thymine comme observé précédemment [94], il ne peut pas non plus se développer sur M9+5mU ou M9+5mC.

On peut en déduire que la thymidine phosphorylase est nécessaire au sauvetage du phénotype d'auxotrophie pour la thymidine par les analogues nucléosidiques 5mU et 5mC ; et que la méthylation de l'uridine antérieure à sa réduction en 2'désoxyribouridine ne semble pas favorable.

Toujours dans l'optique d'envisager des voies alternatives de biosynthèse du thymidylate, la versatilité de l'enzyme TrmFO a été explorée dans le but de créer une thymidylate synthase artificielle. TrmFO catalyse la méthylation de l'uridine 54 des ARNt chez certains micro-organismes [44]. Cette réaction étant assez similaire à celle catalysée par les thymidylate synthases, on peut y voir une forme de convergence

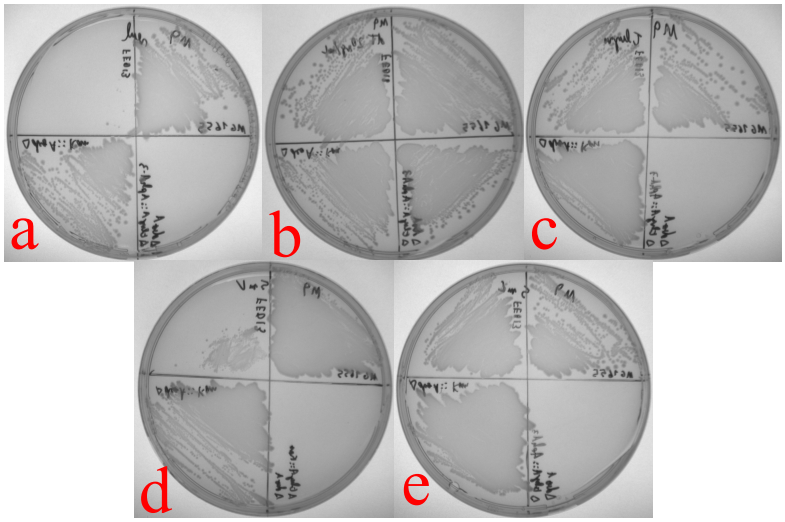

FIGURE 3.10 – **Complémentation des différentes souches d'*Escherichia coli* par des analogues nucléosidiques.** Boîtes de Pétri M9 supplémentées ou non (a) en thymidine (b), thymine (c), 5mU (d) ou 5mC (e) après stries des souches MG1655 (type sauvage, coin supérieur droit de chaque boîte), FE013 (Δ*thyA*, coin supérieur gauche) Δ*deoA* (coin inférieur gauche) et Δ*deoA*Δ*thyA*::*aphA-3* (coin inférieur droit). Le gène *deoA* est nécessaire à la complémentation du phénotype d'auxotrophie pour la thymidine par les analogues nucléosidiques au même titre qu'il l'est pour la thymine.

évolutive. La partie suivante vise à pousser jusqu'au bout cette pseudo-convergence en accélérant *in vitro* l'évolution de TrmFO.

3.2 Évolution dirigée de *trmFO*

Le thymidylate étant un précurseur essentiel de la synthèse d'ADN, les organismes en ont besoin pour survivre et se répliquer. Il existe deux principales voies métaboliques permettant la synthèse du thymidylate :

— La synthèse *de novo* consistant en la méthylation du dUMP par la thymidylate synthase.

— La voie de récupération consistant en la phosphorylation sur le carbone 5' de la 2'désoxythymidine, issue de milieu de culture, par la thymidine kinase.

Un organisme auxotrophe pour la thymidine n'a donc comme seul moyen de survie la récupération de thymidine exogène et ne peut pas pousser sur milieu minimal non supplémenté en thymidine.

3.2.1 Présentation de la méthode

L'évolution dirigée englobe l'ensemble des techniques visant à l'ingénierie d'un gène ou d'un organisme, par une approche stochastique consistant à reproduire *in vitro* le processus d'évolution par sélection naturelle [99]. Les applications de cette méthode sont nombreuses, pour n'en citer que quelques-unes ; elle a permis notamment d'augmenter la thermostabilité de protéines, de créer des enzymes catalysant de nouvelles réactions ou encore d'améliorer la spécificité d'une protéine pour un ligand. En général, une expérience d'évolution dirigée consiste en un processus itératif de deux étapes : la génération d'une diversité génétique et le criblage ou la sélection des mutants d'intérêt. Différentes techniques peuvent être employées à la génération de diversité génétique dont voici les principales :

— La recombinaison de fragments d'ADN homologues ou *DNA shuffling* [100] consiste à digérer partiellement différentes séquences homologues codant pour une protéine d'une fonction donnée pour ensuite reconstituer une séquence "mosaique" complète (Fig. 3.11)

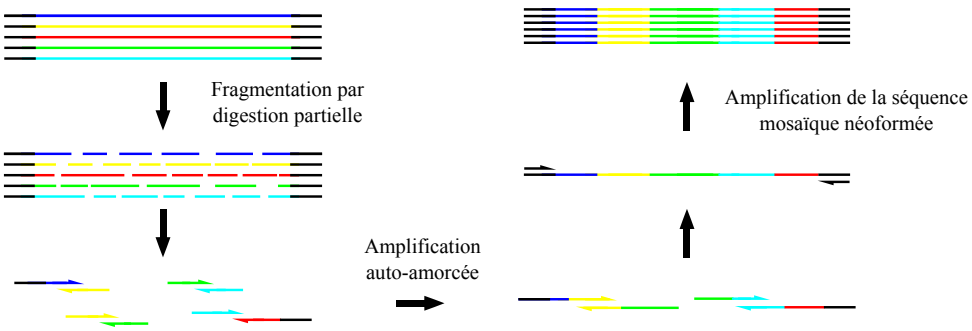

FIGURE 3.11 – **Schéma d'une procédure de recombinaison de fragments par homologie de séquences.** Les fragments homologues sont d'abord soumis à une digestion partielle, donnant lieu à des amorces d'environ 50 kb qui s'hybrideront entre elles pour reconstituer des séquences mosaïques via une amplification auto-amorcée. Les séquences mosaïques seront ensuite amplifiées pour obtenir une banque de mutants.

— La PCR mutagène ou *error prone PCR* est une technique d'amplification exponentielle d'une séquence d'ADN dont la caractéristique est sa faible fidélité. Les produits d'amplification ainsi générés contiennent des erreurs et constituent une banque de mutants.
Différents paramètres d'une PCR peuvent être ajustés pour promouvoir les erreurs comme entre autres la concentration en magnésium, le remplacement du magnésium par du manganèse [101], un déséquilibre de concentration en

dNTP, une polymérase peu fidèle [102].

— La mutagénèse à saturation permet de tester, pour une position donnée, la substitution par tous les acides aminés protéinogènes, via une amplification du gène tout entier par des amorces portant des codons dégénérés. Cette technique est généralement utilisée dans des cas où les connaissances structure/fonction de la protéine cible sont assez avancées.

Le criblage ou la sélection d'une banque de mutants peut prendre plusieurs formes et dépend du but recherché.

3.2.2 Système de sélection

Pour sélectionner une thymidylate synthase artificielle, je disposais d'un souche FE013 [22], auxotrophe pour la thymidine, de génotype Δ*thyA*::*aphA-3*, pour laquelle le gène *thyA* avait été inactivé par l'insertion d'une cassette de résistance à la kanamycine. Nous partons du principe qu'un gène capable de complémenter le phénotype Δ*thyA* coderait pour une thymidylate synthase. Le système de sélection consiste en une simple transformation dans la souche FE013, des banques de gènes issues de mutagénèses (décrites dans les sections suivantes) clonées dans un vecteur d'expression pQE-80L.

3.2.3 Banque de mutants combinatoire du gène *trmFO* de *Bacillus subtilis*, approche semi-rationnelle

La seule structure de TrmFO disponible, à ce jour, est celle venant de la bactérie thermophile (s'il est besoin de préciser) *Thermus thermophilus*. Cette structure avait été résolue avec 1,6 Å de résolution, à l'occasion d'une étude antérieure à ce projet [45]. L'observation de la structure de TrmFO permet de distinguer un sillon de charges positives en aval du site actif (Fig.3.12). Ce sillon pourrait très probablement être responsable de l'interaction entre la protéine et son substrat naturel, l'ARNt.

On suppose ici qu'en empêchant la fixation de l'ARNt à la protéine, le dUMP pourrait diffuser dans le site actif de TrmFO (celui-ci étant spécifique d'une uridine), et se trouver méthylé par cette dernière. Une banque de mutants combinatoire du gène *trmFO* de la bactérie *Bacillus subtilis* a été synthétisée pour tester cette hypothèse, ciblant des résidus lysines susceptible d'être impliqués dans la fixation de l'ARNt.

Les résidus K175, K246, K285, K290, K413 et K414 dont les équivalents chez *Thermus thermophilus* sont respectivement K173, K243, K282, K287, K409 et K410 (Fig. 3.13) ont été mutées en aspartate (D) ou en alanines (A). Une mutation en alanine neutralise la charge positive, tandis qu'une mutation en aspartate la remplace par une charge négative. Les lysines 282, 287, 409 et 410 de l'enzyme de *Thermus thermophilus* sont constitutifs du sillon de charges positives évoqué précédemment et sont donc de bon candidats (Fig. 3.12).

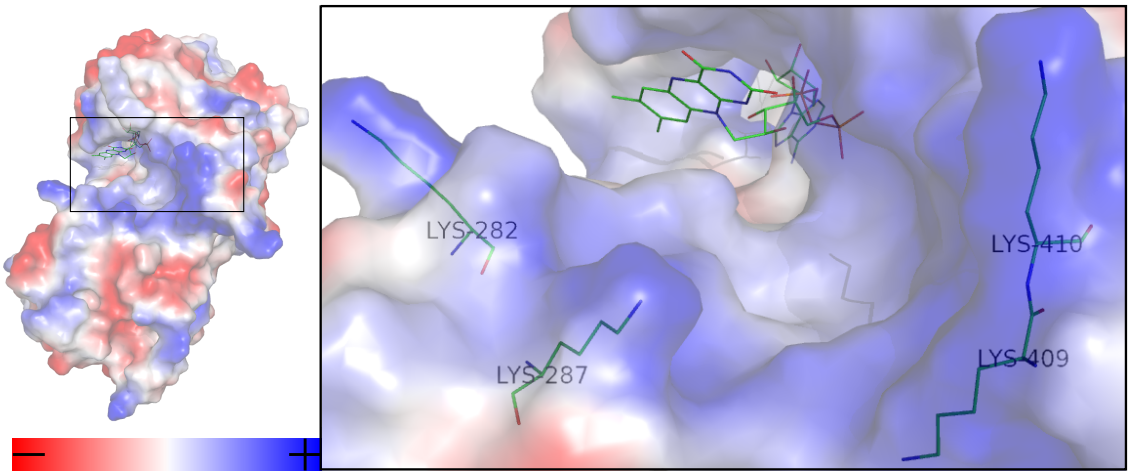

FIGURE 3.12 – **Structure de TrmFO de *Thermus thermophilus* (PDB : 3G5R)** coloré en fonction des charges de surface (les charges négatives en rouge et les positives en bleu). Les résidus K282, K287, K409 et K410 sont mis en évidence pour leur participation dans un sillon de charges positives en aval du site actif (materialisé par le noyau isoalloxazine de la flavine) susceptibles d'interagir avec le substrat naturel de l'enzyme, l'ARNt.

TABLEAU 3.3 – **Résidus substitués dans la banque de mutants combinatoire de TrmFO de *Bacillus subtilis*** (B.s). Les résidus sont listés avec leurs équivalents chez *Thermus thermophilus* (T.t), leurs scores de conservation déterminés par un alignement multiple ClustalW [103] (les plus conservés ont un score négatif) et l'activité catalytique de leurs mutants en alanine chez *Thermus thermophilus* selon [45].

résidus chez *T.t*	résidus chez *B.s*	score de conservation	activité chez *T.t*
K173	K175	-0,477	NI
K243	K246	-0,718	NI
K282	K285	-0,686	6 %
K287	K290	-0,397	5 %
K409	K413	-0,717	65 %
K410	K414	0,646	30 %

```
B.s   124  TEIPEGPTIIATGPLTSESLSAQLKELTGEDYLYFYDAAAPIVEKDSLDMDKVYLKSRYD  183
           EIP G T++ATGPLTSE+L+  LK   G+ +L +YDAA+PIV  +S+D+ K +   RY
T.t   122  REIPPGITVLATGPLTSEALAEALKRRFGDHFLAYYDAASPIVLYESIDLTKCFRAGRYG  181

B.s   184  KGEAAYLNCPMTEEEFDRFHEALTSAETVPLKEFEKEIFFEGCMPIEVMAKRGKKTMLFG  243
           +  A YLNCPMTEEE+ RFH+AL  A+     ++EK  FFE C+P+E +A+RG +T+LFG
T.t   182  Q-SADYLNCPMTEEEYRRFHQALLEAQRHTPHDWEKLEFFEACVPVEELARRGYQTLLFG  240

B.s   244  PMKPVGLEHPVTGKRPYAVVQLRQDDAAGTLYNIVGFQTHLKWGDQKEVLKLIPGLENVE  303
           PMKPVGL  P TGK P+AVVQLRQ+D AG ++++VGFQT LKW +QK ++++IPGLEN E
T.t   241  PMKPVGLVDPRTGKEPFAVVQLRQEDKAGRMWSLVGFQTGLKWPEQKRLIQMIPGLENAE  300

B.s   304  IVRYGVMHRNTFINSPSLLKPTYQFKNRSDLFFAGQMTGVEGYVESAASGLVAGINAAKL  363
           IVRYGVMHRNT++N+P LL  T +F+    L+ AG + GVEGY+ESAA+G +AG+NAA+
T.t   301  IVRYGVMHRNTYLNAPRLLGETLEFREAEGLYAAGVLAGVEGYLESAATGFLAGLNAARK  360

B.s   364  VLGEELVIFPQETAIGSMAHYITTTNQKNFQPMNANFGLLKELPVKIKNKKERNEQYANR  423
            LG   V  P+E+ +G +  Y+ T N + FQPM AN+GL+  +  ++  K++R   Y  R
T.t   361  ALGLPPVAPPEESMLGGLVRYLATANPEGFQPMYANWGLVPPVEGRMGKKEKRQAMY-RR  419
```

FIGURE 3.13 – **Alignement partiel des séquences protéiques de TrmFO de *Bacillus subtilis* et *Thermus thermophilus*.** Obtenu a l'aide de l'algorythme BLAST, l'alignement (50,7% de similitude) met en valeur (en rouge) les lysines ciblées par la librairie de mutants combinatoires.

Les activités catalytiques des mutants K282A K287A K409A et K410A (pour la protéine TrmFO de *Thermus thermophilus*) ont été testés [45] et donnent respectivement 6 %, 5 %, 65 % et 30 % (Tab. 3.3). Ces résidus étant relativement éloignés du site actif, on peut supposer qu'ils ne sont pas impliqués dans le mécanisme catalytique, et que la baisse d'activité s'explique par une baisse d'affinité pour le substrat. Trois possibilités pour chacune des six positions ciblées donnent lieu à $3^6 = 729$ combinaisons différentes. On part du principe que le test des différentes combinaisons de mutants équivaut à un tirage aléatoire avec remise. La probabilité d'avoir criblé la totalité de la banque en fonction du nombre de tirages est :

$$\sum_{i=0}^{k} (-1)^{k-i} \begin{pmatrix} k \\ i \end{pmatrix} \left(\tfrac{i}{k}\right)^n$$

avec k, le nombre de variants (729) et n le nombre de tirages (nombre de transformants). Le tracé de la probabilité d'avoir la banque de mutants entièrement représentée en fonction du nombre de transformants donne la figure 3.14.

On peut ainsi déterminer qu'avec 9000 transformants, la totalité de la banque est représentée avec quasi certitude. Pour estimer le nombre de transformants, on

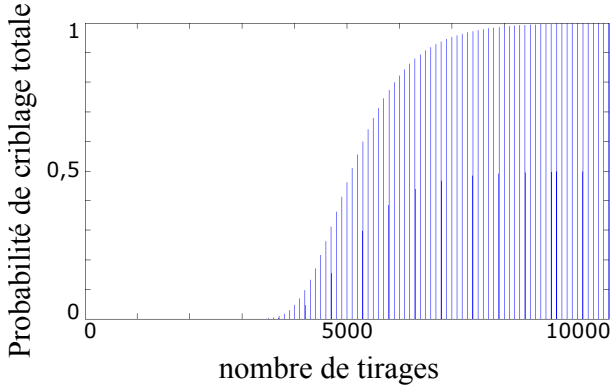

FIGURE 3.14 – **Probabilité d'avoir la banque de mutants entièrement représentée en fonction du nombre de transformants.** Modélisation Matlab d'un tirage avec remise de 729 événements équiprobables.

réalise un contrôle positif avec le plasmide pHM460 qui correspond au gène *thyX* de *Chlamydia trachomatis* cloné dans le vecteur d'expression pQE-80L.

Deux transformations de la banque ont été réalisées pour lesquelles on a pu estimer à 5000 et 6000 le nombre d'évènements de transformation, soit un total de 11000 transformants, parmi lesquels aucun n'a pu croître sur milieu sélectif. Ainsi, aucune des 729 combinaisons de mutations ne peut complémenter le phénotype $\Delta thyA$ sous nos conditions expérimentales.

3.2.4 Mutagénèse aléatoire par *error prone* PCR

Le gène *trmFO* de *Bacillus subtilis* cloné dans le vecteur pQE-80L a été amplifié par PCR mutagène sous différentes conditions afin d'obtenir des taux d'erreurs plus ou moins importants. Le produit d'amplification a été recloné dans le même plasmide d'expression pQE-80L avant d'être transformé dans une première souche d'*E.coli* optimisée pour l'efficacité de transformation (STELLAR) et sélectionnée sur LB+Amp. La totalité des colonies transformées par ce plasmide recombinant ont été grattées de leur boîte pour en extraire les plasmides dans une même préparation (Fig. 3.15).

Le produit de cette extraction d'ADN plasmidique constituant notre banque de mutants a été transformé dans 50 μL d'une suspension de cellules électrocompétentes de la souche FE013, laquelle a été étalée sur milieu sélectif.

Après une première expérience de mutagénèse aléatoire de *trmFO* de *Bacillus subtilis*, on a pu observer environ 2750 colonies contre 115 colonies sur un

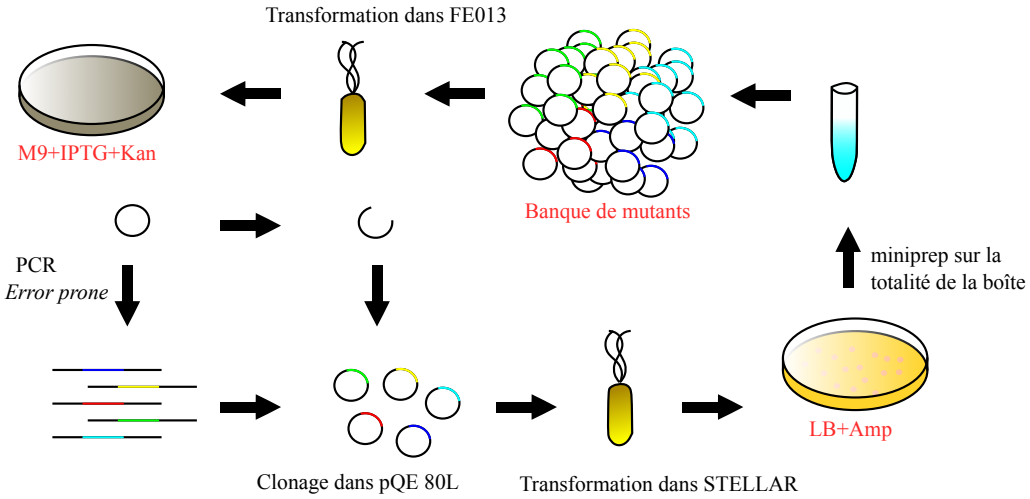

FIGURE 3.15 – **Protocole d'évolution dirigée par PCR mutagène de *trmFO* en enzyme complémentant le phénotype Δ*thyA***. Le plasmide est d'abord amplifié, cloné dans pQE-80L, le produit de clonage est transformé dans la souche STELLAR qui sera étalée sur milieu sélectif. Les colonies formées sur boîte sont ensuite grattées pour en extraire les plasmides qui constituent la banque de mutants ; celle-ci sera enfin transformée dans FE013 pour sélectionner un ou plusieurs gènes complémentant le phénotype Δ*thyA*.

témoin de recircularisation, suggérant un crible de 2635 séquences potentiellement différentes. Une seconde expérience pour laquelle différents mutants ponctuels ainsi que la banque de mutants combinatoire ont été utilisés comme matrice a donné lieu à 1200 mutants différents pour cinq recircularisations. Aucun de ces variants n'a pu complémenter le phénotype d'auxotrophie pour la thymidine. On notera que le nombre de mutants obtenus au cours de la deuxième expérience est particulièrement faible par rapport au nombre minimal requis pour cribler ne serait-ce que la banque de mutants combinatoire donné en figure 3.14. Le nombre de possibilités explorées est donc relativement faible, et le système de criblage pourrait être optimisé par une technique de clonage plus efficace. En outre, la génération de diversité génétique par PCR mutagène peut, elle aussi, être améliorée comme décrit dans la section suivante.

3.2.5 Développement d'un utilitaire d'optimisation de l'encodage des gènes pour la mutagénèse aléatoire

Il existe différentes limites à la génération d'une diversité génétique par PCR mutagène dont l'une est intrinsèquement liée à l'usage du code. On peut considérer,

vu les taux de mutation d'une expérience classique de PCR mutagène, qu'un codon ne sera jamais muté plus d'une fois. L'encodage d'un acide aminé donné peut donc, d'une certaine manière, déterminer son potentiel devenir à l'issue d'une PCR *error prone* ; l'exemple un peu extrême de la sérine, illustré par la figure 3.16, montre qu'il est plus intéressant d'encoder une sérine par AGC que par UCA.

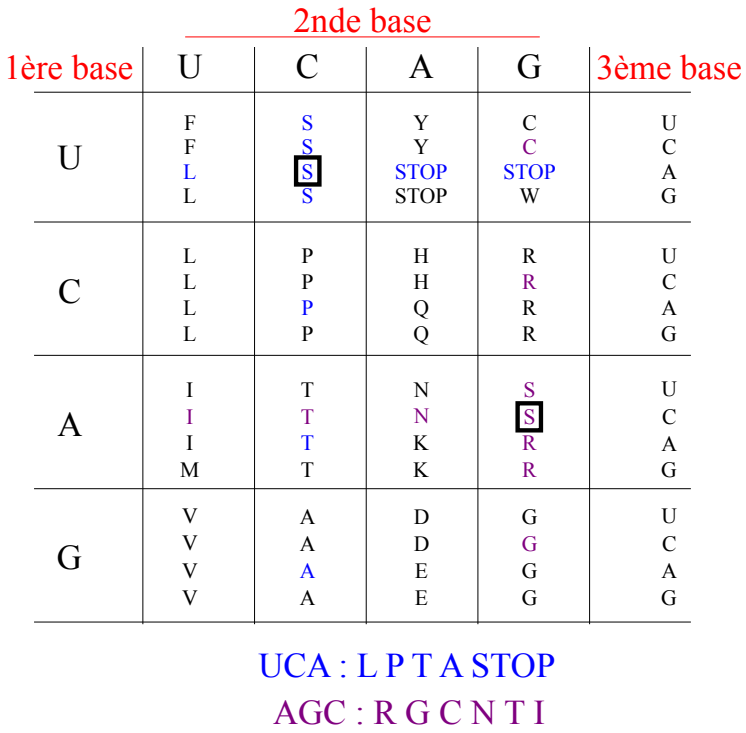

FIGURE 3.16 – **Représentation du code génétique.** L'exemple des sérines illustre l'importance de l'encodage d'un gène pour une expérience de PCR mutagène. La sérine codée par le codon UCA et ses potentielles mutations sont en bleu, la sérine codée par AGC et ses potentielles mutations sont en mauve.

Lors d'une expérience de mutagénèse aléatoire, certains événements peuvent être particulièrement indésirables, comme l'apparition d'un codon "STOP" au milieu d'une séquence qui mènerait à une protéine tronquée ou l'apparition d'une proline susceptible de perturber la structure de la protéine. En revanche, il est souhaitable d'augmenter la diversité des mutations des différents résidus. Un programme de prédiction prenant en entrée la séquence et proposant en sortie la probabilité d'apparition des différents résidus a déjà été développé [104].

Nous proposons ici une démarche inverse, un programme (toujours en cours de développement) d'encodage prenant en entrée un score attribué à chaque critère (vers une proline, vers un codon "STOP", optimisation de la diversité de mutations pos-

sible) et proposant en sortie une table d'usage du code, utilisable par un programme de traduction inverse. Le code source de cet utilitaire est présenté en annexe. D'autre part, les erreurs que commettent les polymérases lors d'une expérience d'évolution dirigée ne sont pas distribuées équitablement, c'est-à-dire que les trois bases issues de la mutation d'une base donnée n'ont pas la même probabilité d'apparaître (Tab. 3.4).

TABLEAU 3.4 – **Biais de mutagénèse de différentes polymérases utilisées pour la PCR mutagène.** Les différentes valeurs ont été récupérées à partir du manuel du kit GeneMorph II Random Mutagenesis d'Agilent Technologies.

Type de mutation	MutazymeII	MutazymeI	*Taq*
A⇒N, T⇒N	50,7%	25,6%	75,9%
G⇒N, C⇒N	43,8%	72,5%	19,6%
A⇒G, T⇒C	17,5%	10,3%	27,6%
G⇒A, C⇒T	25,5%	43,7%	13,6%
A⇒T, T⇒A	28,5%	11,1%	40,9%
A⇒C, T⇒G	4,7%	4,2%	7,3%
G⇒C, C⇒G	4,1%	8,8%	1,4%
G⇒T, C⇒A	14,1%	20,0%	4,5%

Il en résulte un biais de mutagénèse propre à chaque technique de mutagénèse aléatoire. Une autre fonctionnalité du programme proposée est la prise en compte de ce biais de mutagénèse en pondérant le score de chaque mutation par sa probabilité d'apparaître.

3.3 Conclusions et ouvertures

Au cours de ce chapitre, j'ai pu me pencher sur la séquentialité du métabolisme des pyrimidines et, en particulier, sur le fait que les réactions chimiques des différents métabolites s'effectuent dans le même ordre que celui de leur apparition au cours de l'évolution, à savoir : dans un premier temps une réduction des nucléotides en 2'désoxyribonucléotides, (ARN vers ADN) et dans un second temps, la méthylation de l'uridine (ADN U vers ADN T).

S'est alors posée la question de la versatilité de ces voies métaboliques, jusqu'à quel point le système est-il figé ? Peut on envisager une forme de vie artificielle avec un anabolisme du thymidylate différent, une thymidylate synthase artificielle dont

le mécanisme catalytique est celui d'une enzyme de méthylation des ARNt, une voie métabolique différente et peut-être plus avantageuse de biosynthèse *de novo* du thymidylate ?

Ce travail a permis de mettre en évidence qu'une souche d'*E.coli* auxotrophe pour la thymidine peut pousser sur milieu contenant de la ribothymidine au même titre que la thymidine ou de la thymine ; ce sauvetage n'est possible que par l'action de la thymidine phosphorylase. Cette nouvelle observation suggère que les enzymes de la voie de sauvetage, la thymidine kinase et/ou la thymidylate kinase n'ont pas d'activité, suffisante du moins, sur la ribothymidine pour permettre son incorporation dans l'ADN.

L'idée soutenue par ce chapitre est que la méthylation de l'uridine postérieure a sa réduction constitue une erreur figée ou *frozen mistake* (comme définie par F.Crick [16]) et que la méthylation de l'uridine antérieure à sa réduction permettrait de réduire considérablement la concentration intracellulaire en dUTP, réduisant ainsi son incorporation dans l'ADN, elle même réduisant le besoin en enzymes de réparation du type Uracil-N-glycosylases (UNG) relativement couteuses en énergie pour la cellule. Seulement, la méthylation de l'uridine étant apparue après la réduction des nucléotides au cours de l'évolution, il semble que les différentes enzymes du métabolisme des pyrimidines sont incapables de prendre en charge cette voie potentielle.

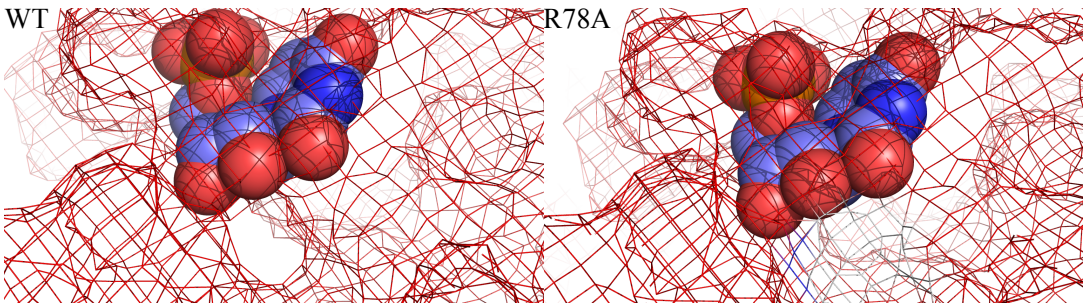

FIGURE 3.17 – **Modélisation de l'UMP dans le site actif de ThyX de *Thermotoga maritima*** à partir de la structure 1O25. WT : l'UMP est en représentation "spheres" et l'enzyme en rouge et en représentation "mesh", permet de voir un clash stérique entre le groupement 2'OH ajouté manuellement et l'enzyme. R78A : Il n'y a plus de clash après simulation de l'effet sur la structure d'une mutation de l'arginine 78 en alanine.

De cette éventuelle voie alternative de synthèse *de novo* du thymidylate, la première étape serait une ribothymidylate synthase. Les thymidylates synthases semblent spécifiques du dUMP, bien que certaines données de diffraction de rayons

X aient suggéré la présence de ribothymidylate co-cristallisé avec une thymidylate synthase ThyX (H.Myllykallio, com. verb.). Une simple modélisation de l'UMP dans le site actif de la thymidylate synthase ThyX de *Thermotoga maritima* permet de mettre en évidence un clash stérique du groupement hydroxyle en 2' avec le résidu R78 du motif RHR très conservé chez les ThyX (Fig. 3.17), qui pourrait être la cause de leur manque de spécificité envers l'UMP.

Ceci dit, du TMP a été très récemment observé *in vivo* chez *Streptomyces viridochromogenes* comme intermédiaire dans la chaîne de biosynthèse de polyoxynes [105] (une famille d'antibiotiques). L'enzyme responsable de cette réaction, codée par un gène homologue de *thyX*, catalyse indifféremment la méthylation du dUMP ou de l'UMP ; cette enzyme est capable de complémenter un phénotype d'auxotrophie pour la thymidine. Ce qui montre que le TMP compte parmi les métabolites du vivant après tout. Néanmoins, le TMP semble dans l'incapacité de court-circuiter la voie métabolique traditionnelle de biosynthèse du thymidylate chez *Escherichia coli*.

L'explication pourrait être l'éventuelle production de TTP après phosphorylations successives du TMP par les nucléotide kinases. Sans discrimination entre TTP et UTP par les ARN polymérases, le ribothymidylate pourrait être incorporé dans l'ARN à la place de l'uridylate et engendrer d'importantes perturbations structurales et fonctionnelles (Fig. 3.18). Cette limite pourrait expliquer pourquoi la méthylation de l'uridine postérieure à sa réduction a été conservée dans cet ordre par le vivant.

Il apparaît donc assez clairement que cette voie métabolique alternative est très peu probable. L'élaboration d'un nouveau mécanisme catalytique de méthylation du dUMP reste, quand à elle, une possibilité intéressante.

Au cours de ce travail, une approche de mutagénèse aléatoire a été développé en vue de créer une (ribo)thymidylate synthase artificielle à partir de TrmFO, une enzyme catalysant une réaction similaire à celle des thymidylate synthases, sur la boucle T-Ψ d'un ARN de transfert.

Les différents essais ont pu de mettre en évidence certaines limites de l'expérience, à savoir l'étape de clonage des produits de mutagénèse aléatoire ainsi que l'importance de l'usage du code dans les expériences de PCR mutagènes en général. Le fait de s'être penché sur ce problème a permis de proposer un utilitaire ayant le potentiel d'améliorer l'efficacité des expériences de PCR mutagènes.

Le fait qu'aucun gène complémentant le phénotype d'auxotrophie pour la thymidine n'ait été obtenu à partir de *trmFO* ne veut pas dire que c'est impossible. Le

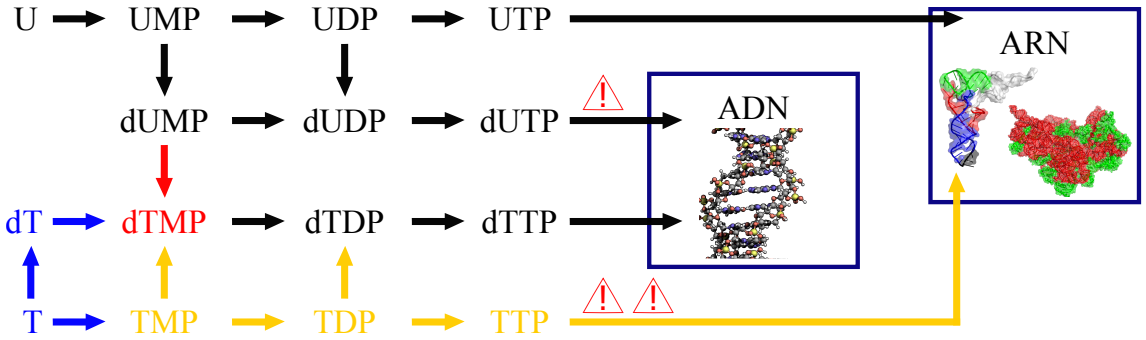

FIGURE 3.18 – **Schéma récapitulatif des voies classique et hypothétique de biosynthèse du thymidylate.** La voie de sauvetage est en bleu, en noir, les voies existantes et en orange, la voie envisagée. L'insertion d'uridine dans l'ADN est délétère dans la mesure ou elle nécessite l'activité des UNG. Cette insertion d'uridines dans l'ADN pourrait être fortement diminuée en substituant à la méthylation du dUMP (voie de synthèse *de novo* en rouge), celle de l'UMP (voie en orange). L'inconvénient de cette voie pourrait être l'insertion de thymidine dans l'ARN probablement beaucoup plus délétère, qui expliquerait pourquoi cette voie n'a pas été sélectionnée au cours de l'évolution.

nombre de mutants possibles pour une séquence peptidique donnée étant considérable (20^n avec n=nombre de résidus), on peut être certain de ne pas avoir exploré toutes les possibilités de mutants pour TrmFO. Gardant à l'esprit que le système de sélection est particulièrement astringent (tout ou rien), on pourrait envisager un système plus graduel, avec par exemple, de faibles concentration en thymidine. La direction prise par mes travaux a plutôt consisté en une approche rationnelle, seulement, les données sur TrmFO n'étant pas suffisantes, une meilleure caractérisation de cette enzyme est nécessaire. Tel est l'objet du chapitre suivant.

Chapitre 4

Caractérisation biochimique de la flavoenzyme TrmFO de *Thermus thermophilus*

Les flavines sont des molécules dotées d'une forte versatilité chimique, faisant d'elles des groupements prosthétiques d'enzymes catalysant de multiples réactions [106]. Les flavoprotéines sont impliquées dans des processus divers (Fig. 4.1) de transferts d'électrons (notamment dans la chaîne respiratoire), de photoréception (comme le rythme circadien ou le phototropisme chez les plantes [107]), de bioluminescence [108] comme de déshydrogénation à deux électrons (intervenant dans le métabolisme énergétique ou le métabolisme des xénobiotiques [109]) etc...

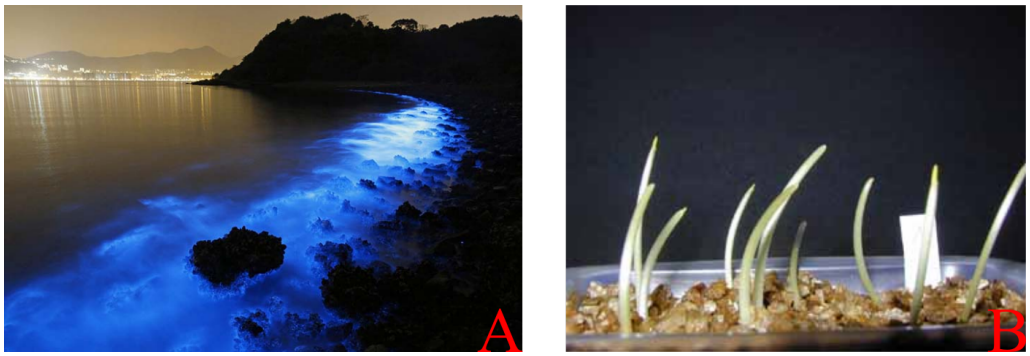

FIGURE 4.1 – **Différents processus biologiques impliquant des flavoprotéines.** A : Bioluminescence issue d'une floraison de *Noctiluca scintillans* dans la baie de Hong Kong (www.theatlantic.com/photo/2015/01/a-bioluminescent-bloom-in-hong-kong/384759/) B : Phototropisme de germes de blé (www.snv.jussieu.fr/bmedia/mouvements/trop-photo.htm).

Les cofacteurs flaviniques se caractérisent par leur noyau isoalloxazine qui leur confère une couleur jaune. On en compte trois principaux types (Fig. 4.2) : les riboflavines composées du noyau isoalloxazine et d'un ribitol, les flavines mono-nucléotides (FMN) dont le ribitol est phosphorylé et les flavines adénine dinucléotides (FAD), pour lesquelles une adénosine diphosphate est liée au ribitol.

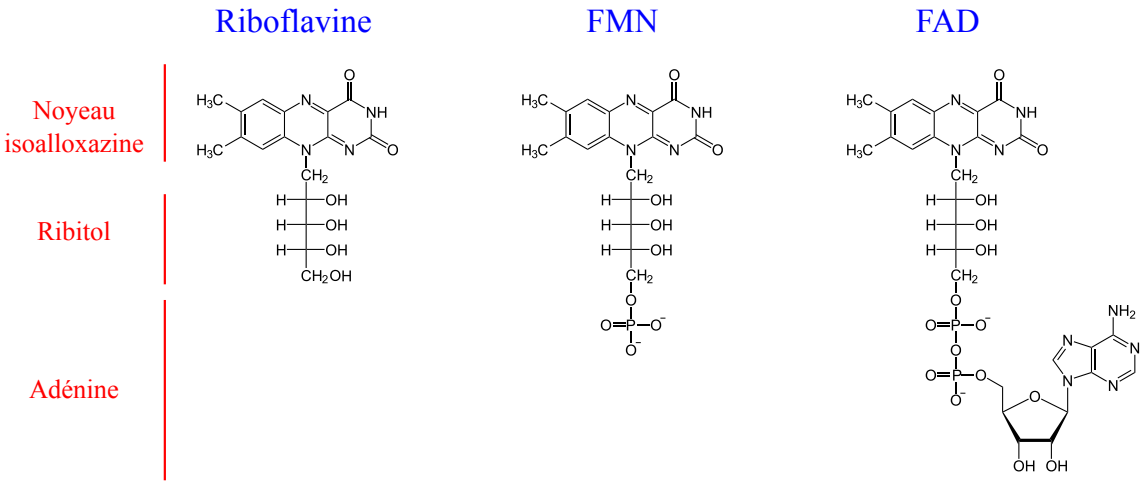

FIGURE 4.2 – **Les trois principaux types de cofacteurs flaviniques.** La riboflavine est composée d'un noyau isoalloxazine et d'un ribitol. La flavine mononucléotide (FMN) consiste en une riboflavine phosphorylée. La flavine adénine dinucléotide est composée d'une FMN à laquelle est lié un groupement adénosine monophosphate.

Les caractéristiques spectrales des flavines font des flavoprotéines des sujets d'étude de choix en spectroscopie d'absorption et de fluorescence. En effet, les spectres d'absorption des flavoprotéines peuvent changer drastiquement en fonction de leurs états d'oxydo-réduction, comme présenté en figure 4.3.

Les flavines ont un pouvoir fluorescent important avec, pour la forme oxydée, un maximum d'excitation à 450 nm et un pic d'émission à 520 nm. Cette fluorescence des flavines est très versatile et peut être éteinte par leur micro-environnement au sein des flavoprotéines ou par une interaction entre leur noyau isoalloxazine et un substrat. Certaines modifications chimiques peuvent aussi engendrer des décalages de longueur d'onde d'excitation ou d'émission.

Les flavoprotéines, de par ces caractéristiques spectrales, se prêtent bien aux études de dynamique du site actif par spectroscopie résolue en temps [111].

L'activité de transfert de carbone à partir du MTHF catalysée par des flavoprotéines est un phénomène récemment découvert. On compte aujourd'hui parmi ces protéines, outre TrmFO, la thymidylate synthase ThyX et l'enzyme RlmFO de

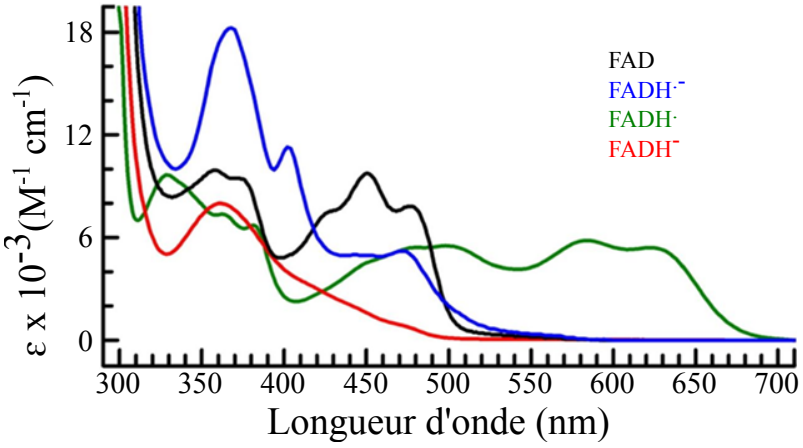

FIGURE 4.3 – **Les différents spectres d'absorption des flavines en fonction de leurs états d'oxydation ou de réduction.** D'après [110]

méthylation d'ARN ribosomal. TrmFO et RlmFO ont une origine évolutive commune que ThyX ne partage pas ; néanmoins, la compréhension de certains aspects du mécanisme réactionnel d'une enzyme a pu être extrapolée aux autres, permettant différents angles d'attaque dans l'étude d'un phénomène plus global qu'est le mécanisme réactionnel des flavoenzymes méthyle-transférases.

La méthylation de l'uridine 54 sur son carbone 5 constitue une modification post-transcriptionelle des ARNt très conservée parmi les organismes vivants, elle favorise une interaction de type *reverse Hoogsteen* avec la méthyladénosine 58 au sein de la boucle T-Ψ (Fig. 4.4), stabilisant la structure en "L" de l'ARNt.

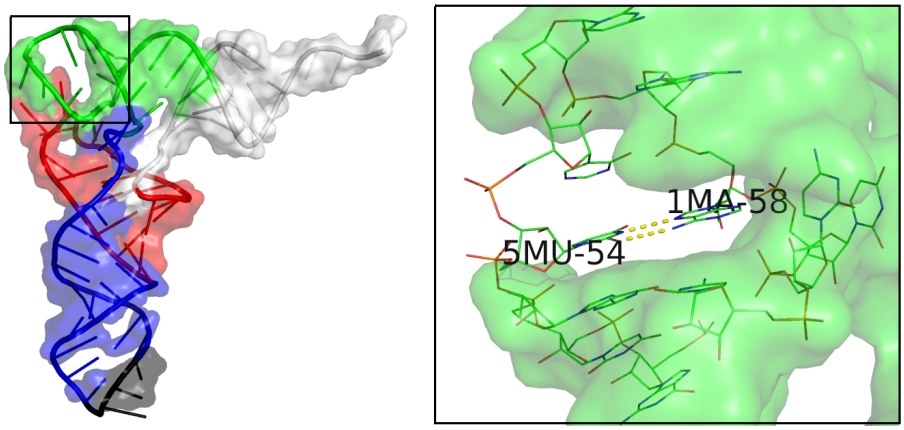

FIGURE 4.4 – **Structure de la boucle T-Ψ de l'ARNt^{Phe} de levure (PDB : 4TNA)** (PDB : 4TNA). La 5 méthyluridine et la 1 méthyladénosine 58 forment une interaction de type *reverse Hoogsteen* qui stabilise la structure en "L" de l'ARNt.

Deux enzymes catalysant cette réaction sont connues, TrmA et TrmFO ; elles sont codées par des gènes hétérologues et sont apparues indépendamment. Ainsi, la méthylation de l'uridine 54 constitue une convergence évolutive. TrmA utilise le S-adénosyl-L-méthionine (SAM) comme donneur de méthyle tandis que TrmFO est une flavoenzyme qui utilise le MTHF comme donneur de méthyle.

Ce chapitre résume des travaux effectués sur l'enzyme TrmFO de *Thermus thermophilus*. Au commencement de ce projet, l'enzyme de *Bacillus subtilis* avait été étudiée sous ses aspects mécanistiques principalement [46] [112], avec pour appui les données structurales et fonctionnelles collectées chez celle de *Thermus thermophilus* [45]. D'autres études menées sur l'enzyme de *Thermus thermophilus* avaient permis de déterminer le substrat minimal, fournissant une base intéressante pour l'étude des résidus de la protéine impliqués dans la fixation du substrat ribonucléique [113]. Sur cette base, j'ai effectué des études de structure/fonction de TrmFO, via une approche de mutagénèse dirigée, d'anisotropie de fluorescence et de spectroscopie, j'ai pu déterminer l'implication de différents résidus dans des phénomènes potentiellement pertinents pour la compréhension de l'activité de cette enzyme.

4.1 Fixation du substrat

Au cours du chapitre précédent, la construction d'une banque de mutants combinatoire TrmFO a été décrite, basée sur un postulat selon lequel certains résidus seraient impliqués dans la fixation de l'ARN sans être catalytiques. Au cours d'une étude précédant mes travaux, quatre de ces résidus, les lysines 282, 287, 409 et 410 (Fig. 4.5), avaient été mutés en alanines chez *Thermus thermophilus* et les activités catalytiques de ces mutants, déterminée [45]. Les mutants K409A et K410A s'étaient avérés avoir une activité catalytique réduite de moitié environ par rapport à l'enzyme de type sauvage tandis que les mutants K287A et K282A étaient catalytiquement inactifs. Ces quatre lysines sont relativement éloignées du site actif et ne semblent pas impliquées dans le mécanisme réactionnel proposé préalablement [47]. Les lysines étant des résidus chargés positivement, et les phosphates des acides nucléiques étant chargés négativement, ces résidus sont de bons candidats au rôle de fixation du substrat. Pour tester cette hypothèse, un test de fixation d'un mini-substrat ayant un fluorophore Alexa 532 lié en son extrémité 5' (Fig. 4.6) a été mis en place (le substrat minimum ayant été préalablement identifié [113]).

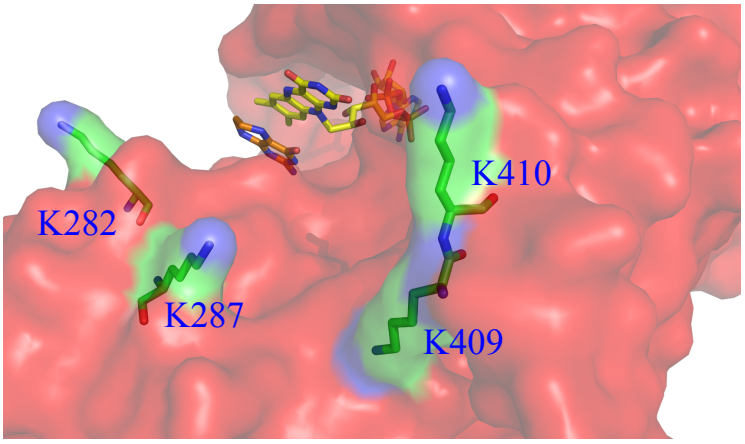

FIGURE 4.5 – **Emplacement des lysines suspectées d'être impliquées dans la fixation du substrat par TrmFO de *Thermus thermophilus* (code PDB : 3G5R).** Les Lysines 282, 287, 409 et 410 ayant un rôle dans l'activité enzymatique sont relativement éloignées du site actif et pourraient être impliquées dans la fixation du substrat ribonucléique.

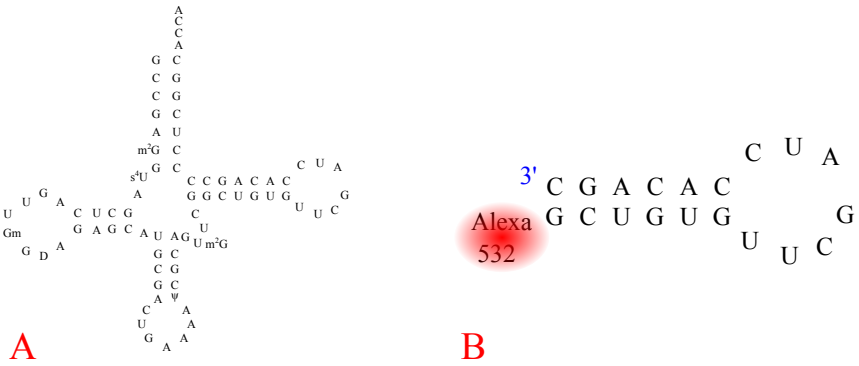

FIGURE 4.6 – **Structures secondaires de l'ARNt^{Phe} de *Thermus thermophilus* (A) et du substrat minimal utilisé en anisotropie de fluorescence (B).**

Afin d'étudier la protéine TrmFO, la synthèse du gène *trmFO* de *Thermus thermophilus* ré-encodé pour une expression optimisée chez *E.coli* a été commandée. Le gène une fois reçu a été sous-cloné dans le vecteur d'expression pQE-80L, lequel a été transformé dans une souche d'*E.coli* BL21 pour exprimer et purifier la protéine. La protéine une fois purifiée a été soumise à une SDS-PAGE afin d'en vérifier la taille et d'en estimer la pureté (Fig. 4.7) [1]. D'autre part, le même plasmide recombinant a pu être utilisé comme matrice de mutagénèse dirigée pour étudier l'implication des quatre lysines dans la fixation du substrat.

1. L'intérêt pour cette étude des différents mutants présentés sur cette figure sera discuté plus loin.

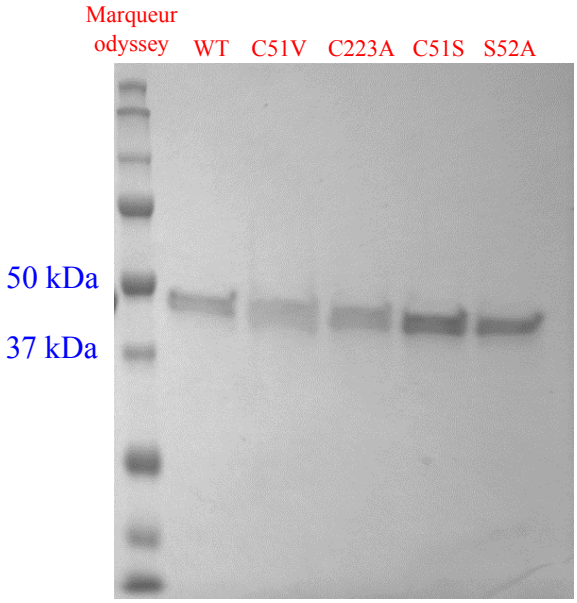

FIGURE 4.7 – **Analyse SDS-PAGE de TrmFO de type sauvage des mutants C51V, C223A, C51S et S52A.** Les signaux des différentes protéines correspondent à la masse moléculaire attendue pour TrmFO (49 kDa) après 30 minutes de migration à 100 V dans un gel de 12 % d'acrylamide.

4.1.1 Principe de l'anisotropie de fluorescence

ANISOTROPE *Se dit d'un corps dont les propriétés physiques dépendent de la direction suivant laquelle on les évalue : les corps cristallisés sont généralement anisotropes.* [114]

L'anisotropie de fluorescence permet d'observer la fixation d'une protéine à une molécule fluorescente par un changement de dépolarisation de la lumière émise par rapport à la lumière excitatrice [115]. L'anisotropie de fluorescence (r) d'une solution est directement liée au degré de liberté rotationnelle des fluorophores qui la composent, à l'échelle temporelle de leur durée de vie de fluorescence. En solution, le ratio d'émission de lumière polarisée parallèlement et perpendiculairement a une lumière excitatrice sera lié directement aux mouvements browniens du fluorophore. Lorsque une macromolécule forme un complexe avec un fluorophore, elle contraint les mouvements browniens de celui-ci et on peut observer une augmentation du rapport I_{VV}/I_{VH} (intensité de lumière émise polarisée verticalement sur l'intensité de lumière émise polarisée horizontalement après excitation par une lumière polarisée verticalement). Ce rapport intégré dans l'équation suivante permet une relation linéaire entre l'anisotropie du milieu et la dépolarisation de la lumière émise [116] :

$$r = \frac{I_{VV} - I_{VH}}{I_{VV} + 2I_{VH}}$$

avec en indices, la disposition des polarisateurs de la lumière excitatrice et de la lumière captée respectivement.

Ceci dit, l'appareil de mesure polarise la lumière (Fig. 4.8), il convient donc d'intégrer à cette équation un facteur de correction. L'intensité d'émission de lumière polarisée perpendiculairement à la lumière excitatrice, pour une solution donnée, doit être constante. Gardant à l'esprit que la composante horizontale de la lumière émise est perpendiculaire à la composante horizontale de la lumière captée (en raison de la disposition du capteur par rapport à la source dans un fluorimètre), le coefficient de correction sera donc :

$$G = \frac{I_{HV}}{I_{HH}}$$

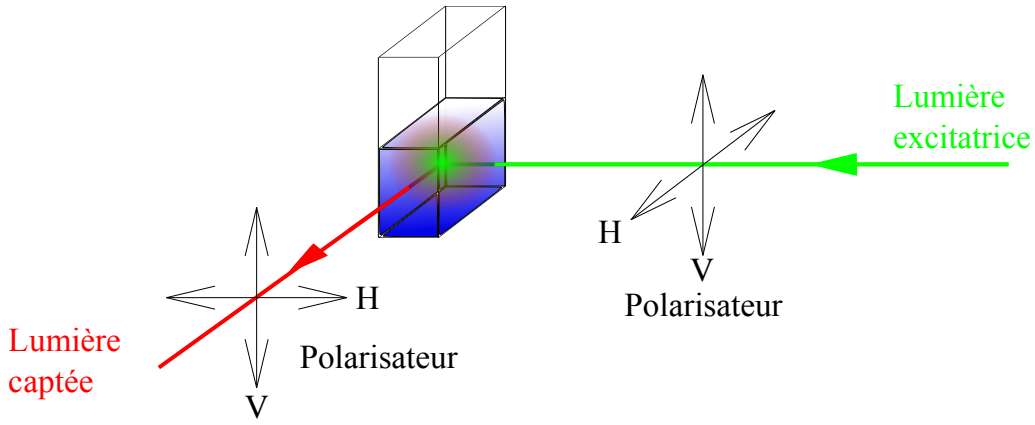

FIGURE 4.8 – **Schéma du montage permettant l'anisotropie de fluorescence.** Un polarisateur se situe après la source de lumière et un autre avant le capteur. On peut noter qu'une seule des quatre configurations de polarisateurs (VV) permet de capter une lumière polarisée parallèlement à la lumière excitatrice.

Ce coefficient, une fois intégré à l'équation qui donne l'anisotropie de fluorescence, on obtient :

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}$$

4.1.2 Résidus impliqués dans la fixation de TrmFO à son substrat

Ainsi ont pu être testées les interactions entre le mini-substrat et différentes protéines dont les résultats sont présentés sur la figure 4.9. On constate tout d'abord que l'augmentation de la concentration en albumine sérique bovine (BSA) n'a pas d'influence sur l'anisotropie contrairement à l'augmentation de la concentration en protéine TrmFO de type sauvage. On peut en déduire une interaction spécifique entre le substrat minimal fluorescent et TrmFO avec une constante de dissociation sub-micromolaire. D'autre part, les courbes de fixation des mutants K282A et K287A se superposent plus ou moins à celle de la protéine de type sauvage, laissant supposer que les lysines 282 et 287 ne jouent pas de rôle important dans l'interaction entre la protéine et son substrat, en revanche, les mutants K409A et K410A semblent fixer le substrat significativement moins efficacement.

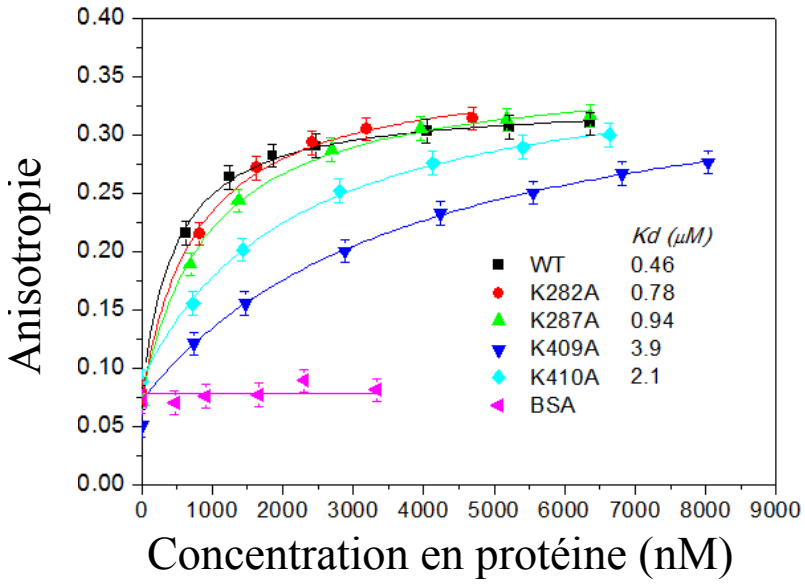

FIGURE 4.9 – **Courbes de fixation de TrmFO et différents mutants à son substrat minimal** Les constantes de dissociation ont été calculées en adaptant les différentes courbes à l'équation $y = a/(1 + b/x) + c$ pour une concentration en substrat de 10 nM. La BSA n'a pas d'influence sur l'anisotropie de fluorescence, tandis que la protéine TrmFO l'augmente, montrant que l'interaction entre le mini-substrat et TrmFO est bien spécifique. Les lysines 282 et 287 ne semblent pas impliquées dans la fixation du substrat, contrairement aux lysines 409 et 410.

Ayant établi que l'interaction avec le substrat est spécifique de la protéine, il reste à s'assurer de la réciproque (spécificité de la protéine pour le substrat). Une

expérience similaire a été effectuée avec un ARN non structuré de même taille lié à un fluorophore Cy5. Les résultats présentés en figure 4.10 montrent que la concentration en BSA n'a pas d'influence sur l'anisotropie de fluorescence de ce nouvel ARN, contrairement à la concentration de TrmFO de type sauvage ; ceci indique une fixation de la protéine à cet ARN.

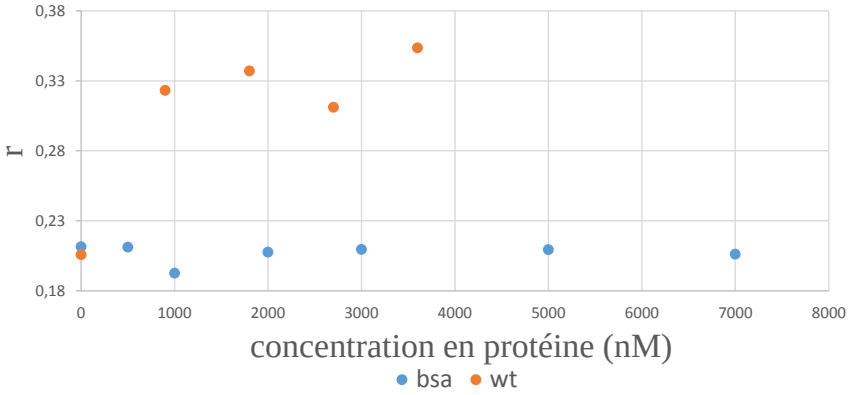

FIGURE 4.10 – **Fixation de TrmFO de type sauvage à un ARN non structuré.** La BSA, faisant office de contrôle négatif, n'a aucune influence sur l'anisotropie de fluorescence de cet ARN, tandis que TrmFO l'augmente significativement, traduisant une interaction avec celui-ci.

Les expériences décrites précédemment ont permis de mettre en évidence la fixation spécifique de TrmFO à son mini-substrat au même titre qu'à un ARN non structuré. On peut en déduire que la structure de l'ARN n'est pas déterminante pour son interaction avec la protéine TrmFO. En outre, ces dernières expériences ont pu mettre en évidence l'implication des résidus K409 et K410 dans la fixation du substrat. Il est intéressant de constater que la mutation des résidus K282 et K287 n'a pas (ou que peu) d'influence sur la fixation du mini-substrat tout en annihilant l'activité enzymatique (d'après [45]). Ces résidus n'ont d'ailleurs pas de rôle apparent dans le mécanisme réactionnel proposé pour TrmFO de *Bacillus subtilis* [46] [112]. On peut alors proposer un mécanisme catalytique en trois temps :

— D'abord, l'enzyme fixe un ARN (sans discrimination de structure) via les lysines K409 et K410 pour ne citer qu'elles.

— Ensuite a lieu la reconnaissance de la structure et l'insertion de l'uracile cible dans le site actif de l'enzyme grâce aux résidus K282 et K287 à la manière de l'enzyme TrmA (qui effectue une réaction similaire via un mécanisme SAM dépendant).

— Enfin peut avoir lieu la réaction de méthylation de l'uridine sur son carbone 5.

Selon des études préalables sur l'enzyme ThyX, l'interaction du dUMP avec le noyau isoalloxazine induit un décalage du pic d'absorption de la flavine [111]. De même manière, les aspects spectroscopiques de TrmFO ont été étudiés dans l'espoir de visualiser un changement de spectre d'absorption induit par le mini-substrat ribonucléique.

4.2 Caractérisation spectrale de TrmFO de *Thermus thermophilus*

Le spectre d'absorption de TrmFO fraîchement purifiée a été mesuré dans le but d'observer une différence attribuable à l'ajout du mini-substrat (Fig. 4.11). Bien que n'ayant pu déceler de changement du spectre induit par le substrat minimal ARN, il a été constaté la protéine TrmFO présentait un spectre d'absorption particulier avec une contribution inhabituelle aux alentours de 400 nm.

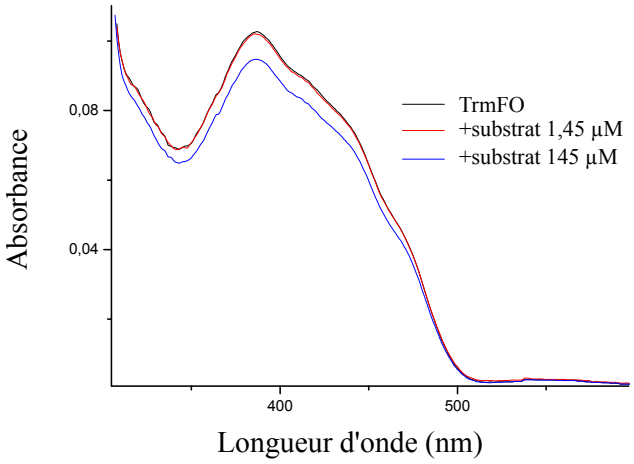

FIGURE 4.11 – **Étude de l'effet du substrat minimal ribonucléique sur le spectre d'absorption de TrmFO.** L'absorbance à 450 nm est d'environ 0,1 et la concentration en flavine peut être estimée à 100 μM. À température ambiante, l'ajout d'une concentration équivalente de substrat n'a pas d'influence sur le spectre global, la baisse d'absorption pouvant être attribuée à un effet de dilution.

4.2.1 Spectroscopie d'absorption UV-visible

Ayant constaté que cette contribution spectrale à 400 nm diminuait irréversiblement lorsque la protéine était exposée à la lumière (tout en restant constante lorsque la protéine était conservée à l'obscurité), j'ai pu affirmer que la protéine fraîchement purifiée comportait une espèce chimique de flavine photosensible (Fig. 4.12A). Après soustraction d'un spectre de flavoprotéine oxydée (en l'occurrence, le mutant Y343F de TrmFO) multiplié par un coefficient correcteur aux différents spectres d'absorption de TrmFO au cours de l'illumination, on obtient un pic à 400 nm qui diminue pour finalement donner un spectre semblable à celui d'une forme $FADH^-$ (Fig. 4.12B). On peut supposer qu'une préparation de TrmFO fraîchement purifiée contient, pour la partie holoenzymatique, un mélange de flavine oxydée et de cette espèce particulière absorbant à 400 nm.

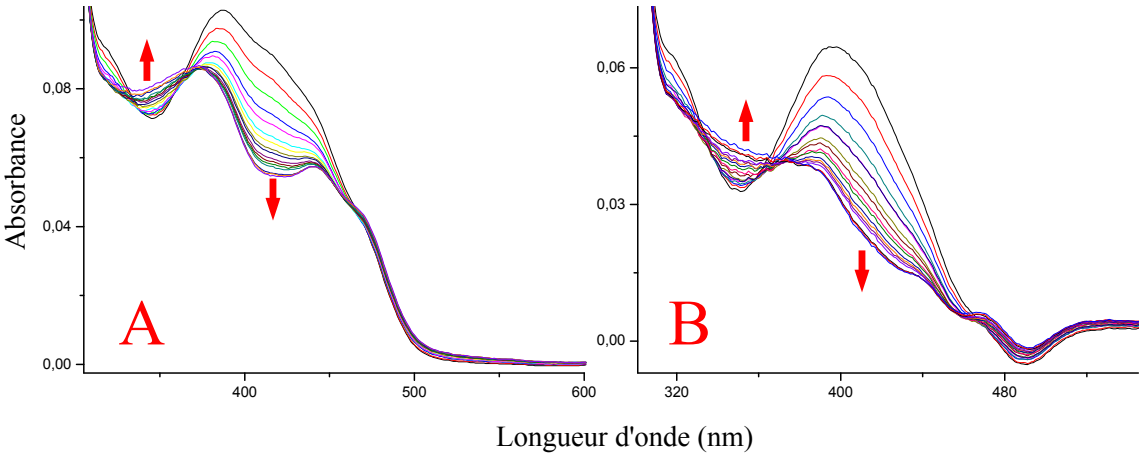

FIGURE 4.12 – **Évolution du spectre d'absorption de TrmFO au fil de son illumination.** A : spectre de la protéine fraîchement purifiée, B : spectre de la protéine duquel est soustrait le spectre du mutant Y343F qui n'a pas de composante à 400 nm. Au cours de l'illumination de TrmFO l'absorption à 400 nm diminue de pair avec une augmentation d'absorption vers 350.

Les données cristallographiques obtenues sur l'enzyme TrmFO de *Thermus thermophilus* [45] ont pu mettre en évidence certains résidus notables aux alentours du noyau isoalloxazine de la flavine (Fig. 4.13). On note ici la présence d'une tyrosine en position 343 à 3,8 Å de la flavine, dont le noyau aromatique est impliqué, avec

cette dernière, dans une interaction de type $\pi - \pi$ et dont le groupement hydroxyle permet une liaison hydrogène avec l'azote d'une cystéine en position 51. Il a été démontré préalablement que cette cystéine 51 est impliquée dans la catalyse [45] au même titre que son équivalent en position 54 de l'enzyme de *Bacillus subtilis* [46]. Afin de déterminer l'implication de ces différents résidus dans la formation de cet espèce (potentiellement révélatrice d'un intermédiaire réactionnel), différentes mutagénèses dirigées ont été réalisées.

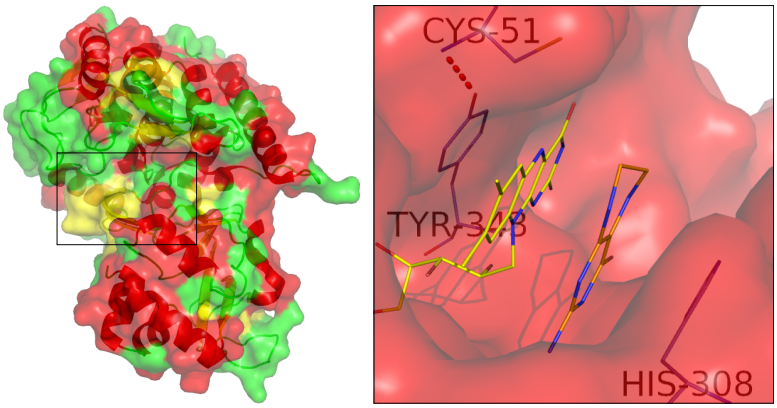

FIGURE 4.13 – **Représentation du site actif de TrmFO de *Thermus thermophilus* (PDB : 3G5R)** co-cristallisée avec FAD (en jaune) et tétrahydrofolate (en orange). On observe des interactions d'empilement $\pi - \pi$ entre la Tyrosine 343, le noyau isoalloxazine de la flavine, le tétrahydrofolate et l'histidine 308. En outre, le groupement hydroxyle de la tyrosine 343 est impliqué dans une liaison hydrogène avec l'azote de la cystéine catalytique C51.

Les plasmides codant pour les mutants C51A, Y343F, et H308G ont été synthétisés par PCR pour être transformés dans la souche d'expression BL21 et *in fine*, exprimer et purifier les protéines pour lesquelles ils codent.

La mutation d'une cystéine en alanine équivaut à une excision du groupement thiol réactif ; celle de la tyrosine en phénylalanine revient à ôter la fonction hydroxyle portée par le noyau φ ; enfin, la mutation d'une histidine en glycine équivaut à dépléter le résidu de sa chaîne latérale.

Les différents mutants purifiés présentaient des spectres de flavine oxydée classiques avec deux pics à 450 nm et 380 nm, indiquant une implication de ces trois résidus dans la formation ou dans la stabilisation de cette espèce absorbant à 400 nm. D'autre part, j'ai pu remarquer que, dans le cas du mutant Y343F, la flavine était photo-réductible, tandis qu'elle ne l'était pas dans le cas des mutants H308G et C51A (Fig. 4.14).

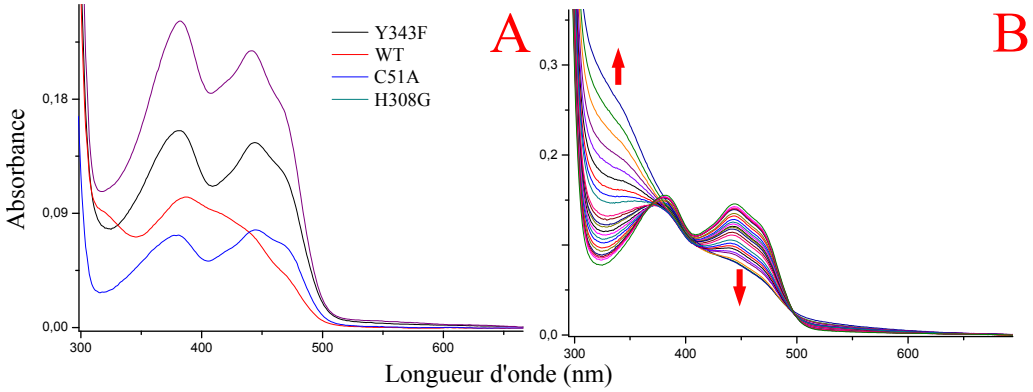

FIGURE 4.14 – **Spectres d'absorption des mutants Y343F et C51A (A) et photoréduction du mutant Y343F (B).** Le spectre d'absorption de la protéine TrmFO fraîchement purifiée présente une composante inhabituelle à 400 nm. Les spectres d'absorption des mutants Y343F, H308G et C51A n'ont pas cette caractéristique. Les résidus Y343, H308 et C51 sont donc probablement impliqués dans la formation ou la stabilisation d'une espèce de flavine absorbant à 400 nm. En outre le mutant Y343F est photo-réductible.

En outre, le rôle d'une seconde cystéine (C223) catalytique a été examiné et le mutant C223A a été conçu, exprimé et purifié. Son spectre d'absorption s'est avéré similaire à celui de la protéine de type sauvage au même titre que son comportement vis-à-vis de la lumière (Fig. 4.15). À ce stade, ce pic d'absorption à 400 nm observé pour l'enzyme de type sauvage pouvait refléter un intermédiaire réactionnel stabilisé dans la protéine fraîchement purifiée, dont la formation nécessitait l'action du résidu C51 mais pas celle du résidu C223.

4.2.2 Identification de l'espèce chimique absorbant à 400 nm

Quelques rares occurrences de flavoprotéines dont le spectre d'absorption présente une contribution inhabituelle à 400 nm apparaissent dans la littérature. On compte parmi elles, le cas d'une flavoprotéine de transfert d'électron (ETF) qui, après ajout de flavine libre, présente un spectre d'absorption similaire [117]. Dans le cas de TrmFO, l'ajout de FAD dans l'extrait cellulaire n'a pas d'effet sur le spectre d'absorption et la contribution à 400 nm est toujours observable. Dans le cas de la chito-oligosaccharide oxydase de *Fusarium graminearum*, la flavine est covalament liée par son carbone 5 à une cystéine et une histidine par son groupement méthyle porté sur le carbone 8. Le spectre d'absorption de cette enzyme dont l'histidine impliquée a été mutée en alanine compte, lui aussi, une contribution à 400 nm [118], probablement caractéristique d'un simple adduit cystéinyl.

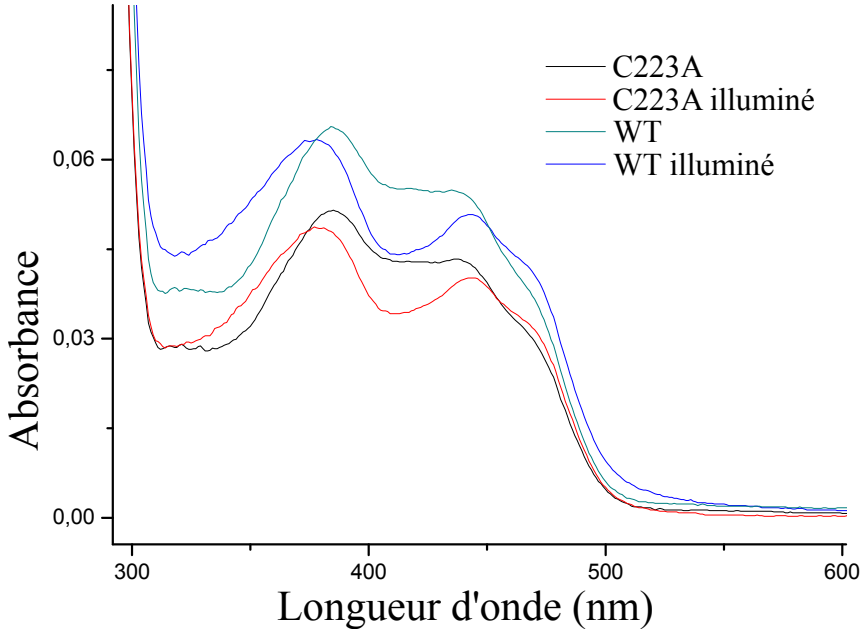

FIGURE 4.15 – **Spectres d'absorption des protéines de type sauvage (WT) et mutante (C223A) avant et après illumination.** Les protéines WT et C223A fraîchement purifiées ont des spectres d'absorption similaires avec une composante à 400 nm et ont le même comportement vis-à-vis de la lumière. Les mesures ont été réalisées dans des cuves de quartz d'1 mm et les concentration en flavine du type sauvage et du mutant peuvent être estimée à 40 μM et 55 μM respectivement

Enfin, des expériences de *stopped flow* sur la sarcosine oxydase ont révélé un intermédiaire réactionnel absorbant à 400 nm [119] [120]. L'intermédiaire en question s'avère être un ion iminium (un groupement methylène sur l'azote 5 du noyeau isoalloxazine). Cette dernière hypothèse sera discutée plus en détails pour sa pertinence mécanistique. En effet, tous les mécanismes catalytiques proposés jusqu'ici pour les méthyle-transférases flavine dépendantes impliquent un transfert de méthylène vers la flavine dans un premier temps, puis vers l'uridine dans un second.

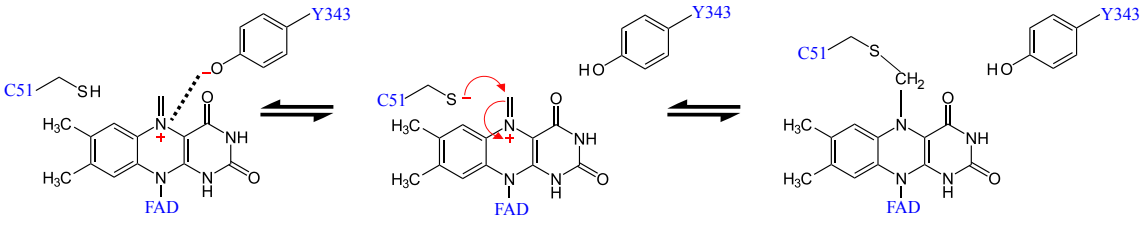

FIGURE 4.16 – **Stabilisation de l'intermédiaire iminium par les résidus C51 et Y343.** La tyrosine 343 permettrait une stabilisation électrostatique de l'iminium ; et la cystéine 51, pourrait le stabiliser via une liaison covalente réversible.

En outre, ce travail a pu affirmer que les résidus H308, Y343 et C51 sont nécessaires à la formation ou la stabilisation de l'espèce chimique absorbant à 400 nm. L'histidine 308 pourrait être responsable du positionnement du MTHF donneur de méthylène en face de la flavine. L'intermédiaire iminium étant probablement instable, on pourrait envisager un rôle stabilisateur pour les résidus C51 et Y343. La figure 4.16 représente un mécanisme postulé de stabilisation de l'iminium par ces derniers.

Il est à noter que le spectre d'absorption de la protéine TrmFO de *Bacillus subtilis* fraîchement purifiée présente des similitudes avec celui de TrmFO de *Thermus thermophilus* après illumination (et après soustraction du spectre d'une flavine oxydée) [112]. Chez *Bacillus subtilis*, ce spectre a été attribué à une flavine alkylée (pont méthylène entre la cystéine 54 et l'azote 5 du noyau isoalloxazine). Dans le cas de TrmFO, cette forme alkylée pourrait résulter d'un effet irréversible de la lumière sur la forme de droite de la figure 4.16. Le mécanisme réactionnel de la lumière sur cette forme reste à élucider. En outre, la forme alkylée chez *Bacillus subtilis* s'oxyde, tandis qu'elle est stable chez *Thermus thermophilus*.

L'instabilité de l'intermédiaire iminium rend son étude difficile, le voici désormais à disposition, stabilisé par le micro-environnement de TrmFO ; ceci devrait permettre de mieux le caractériser, notamment en fluorescence.

4.2.3 Spectroscopie d'émission de fluorescence à l'état stationnaire

Afin de mieux caractériser ce phénomène, des mesures de fluorescence ont été effectuées. Dans le cas de la protéine de type sauvage excitée à une longueur d'onde de 400 nm, on observe un spectre d'émission de fluorescence ayant un maximum aux environs de 540 nm. Dans le cas du mutant C51A excité par une lumière de 450 nm, l'émission de fluorescence est quasi nulle. L'observation des données structurales de TrmFO montrant la tyrosine 343 à proximité du noyau isoalloxazine de la flavine, laissait supposer que cette tyrosine était un *quencher* de fluorescence, à juste titre, dans la mesure où l'on peut constater que la fluorescence est restaurée dans le cas du double mutant C51A Y343F, seulement, avec un pic d'émission à 520 nm après excitation à 450 nm.

Il en a été conclu que la protéine TrmFO de *Thermus thermophilus* fraîchement purifiée contient, pour la partie holoenzymatique, un mélange de protéines dont

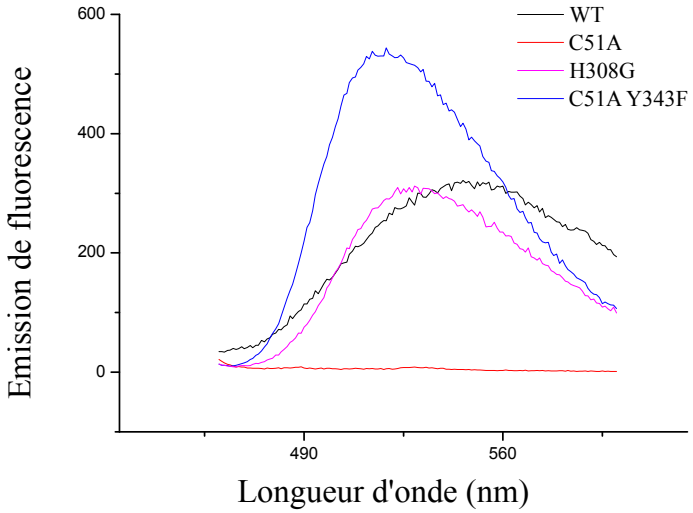

FIGURE 4.17 – **Spectres d'émission de fluorescence de différents mutants de la protéine TrmFO.** La composante absorbant à 400 nm de la partie holoenzymatique de la protéine fraîchement purifiée présente un pic d'émission de fluorescence à 520 nm. La composante "flavine oxydée" (omniprésente chez le mutant C51A) trouve sa fluorescence éteinte par la tyrosine 343, le double mutant C51A Y343F trouvant sa fluorescence restaurée avec un pic d'émission classique à 500 nm.

une partie contient une flavine oxydée et une autre partie contient une espèce de flavine non identifiée. La partie contenant une flavine oxydée trouve sa fluorescence éteinte par la tyrosine 343 ; l'autre partie dont la formation nécessite une cystéine en position 51, absorbe à 400 nm et fluoresce avec un maximum d'émission à 540 nm en plus de s'avérer photoconvertible en une potentielle flavine alkylée.

Pour s'assurer que l'extinction de fluorescence observée chez le mutant C51A n'était pas du fait de la présence d'une alanine à cette position, les plasmides codant pour des mutants C51V et C51S ont été synthétisés, et les protéines correspondantes exprimées et purifiées afin d'être caractériseés en spectrofluorimétrie [2].

Dans les cas de ces trois mutants, l'intensité d'émission de fluorescence s'est trouvée considérablement diminuée (en comparaison avec le spectre d'émission d'une flavine oxydée classique). Néanmoins, la comparaison entre ces trois différents mutants a permis de constater des différences significatives. Tout d'abord, l'émission de fluorescence de C51V est plus importante que celle de C51A (alors que sa concentration en cofacteur donné par l'absorbance à 380 nm est plus faible). En outre, l'intensité d'émission de fluorescence du mutant C51S semble environ sept fois moins

2. Ces dernières expériences ont été réalisées par Maeva SABBAN, à l'occasion de son stage de licence effectué sous ma bienveillante supervision.

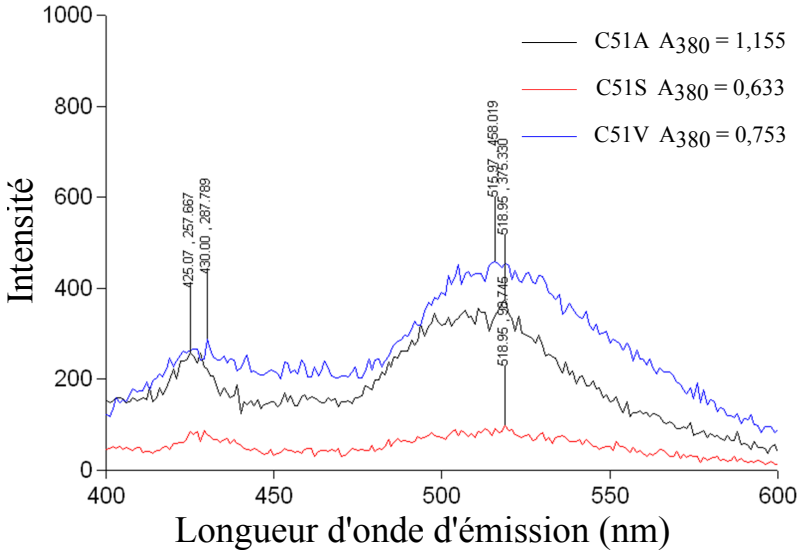

FIGURE 4.18 – **Spectres d'émission de fluorescence de trois différents mutants pour la cystéine 51.** Le photo-multiplicateur a été réglé à 1000 V, et la longueur d'onde d'excitation à 380 nm. L'absorbance à 380 nm pour chaque préparation est indiquée et permet de rapporter les intensités de fluorescence à la concentration en flavine. C51S fluoresce trois fois moins que C51A qui lui même fluoresce deux fois moins que C51V.

importante que celle de C51A pour une différence de concentration d'un facteur deux. La nature du résidu en position 51 a donc une influence notable sur l'extinction de fluorescence de la flavine par la tyrosine 343.

Afin d'étudier plus en détail ces différences observables en état stationnaire, les différentes protéines ont été étudiées en fluorimétrie résolue en temps.

4.3 Spectroscopie ultra-rapide

4.3.1 Principe

La spectroscopie ultra-rapide permet, entre autres, d'étudier des phénomènes de fluorescence avec une résolution de l'ordre de la centaine de femto-seconde. Le principe repose sur un obturateur Kerr. Le milieu de Kerr a la particularité de tourner la polarisation de la lumière qui le traverse sous l'effet d'un pulse à 800 nm. Lorsqu'il est placé entre deux polarisateurs perpendiculaires, il joue le rôle d'obturateur (à la manière de celui d'un appareil photo). Sans pulse à 800 nm, la lumière passe par le premier polarisateur, puis le milieu de Kerr qui n'a pas

d'influence sur sa polarisation, et sera donc arrêtée par le second polarisateur orienté perpendiculairement au premier. Après un pulse de lumière à 800 nm, le changement de polarisation du milieu de Kerr permettera le passage de lumière (une émission de fluorescence en ce qui concerne ce travail). Le temps d'obturation sera défini par le milieu de Kerr utilisé, le CS_2 a un temps de réponse d'une pico-seconde tandis que le suprasil et le benzène en ont un d'environ 200 et 500 femto-secondes respectivement [121].

Afin de décrire un phénomène dans sa globalité, il est nécessaire de l'échantilloner par une fréquence d'acquisition largement supérieure à la durée du phénomène lui même. Par exemple, pour photographier le vol d'un pélican pendant un cinquième de seconde (Fig. 4.19), avec une résolution temporelle satisfaisante, il faudrait prendre 25 photographies par seconde.

FIGURE 4.19 – **Chronophotographie du vol du pélican.** Prise à 25 images par seconde par Etienne-Jules Marey [122].

Pour obtenir un espacement temporel des acquisitions, le laser d'obturation (qui envoie des pulses de lumière à 800 nm sur le milieu de Kerr) s'éloigne du milieu de Kerr pour chaque point. Le décalage spatial induit un décalage temporel du moment d'obturation (le temps parcouru par la lumière étant plus long) permettant l'échantillonnage de l'émission de fluorescence à des intervales de l'ordre de la centaine de femto-seconde.

4.3.2 Identification d'un second *quencher* de fluorescence

On peut constater que la décroissance de fluorescence de la protéine de type sauvage est multiphasique (Fig. 4.20) [3]. Une première phase rapide ($\approx$1 ps) avec un

3. Ces expériences ont été réalisées en collaboration avec Lipsa Nag et Marten Vos.

maximum d'émission vers 520 nm correspond probablement à la fluorescence d'une flavine oxydée et une autre phase plus lente (≈2 ns) avec un maximum d'émission vers 540 nm correspondant à l'espèce absorbant à 400 nm (Fig. 4.21). Le fait que la fluorescence de l'iminium supposé est 2000 fois plus longue que celle de la flavine oxydée explique pourquoi son spectre est largement majoritaire en fluorescence à l'état stationnaire. La fluorescence du mutant C51A décroit considérablement plus rapidement, du fait d'un transfert de charge vers la tyrosine 343. Dans le cas du double mutant C51A Y343F, on constate aussi une phase de décroissance rapide, suggérant que la tyrosine 343 n'est pas le seul résidu à pouvoir éteindre la fluorescence de la flavine.

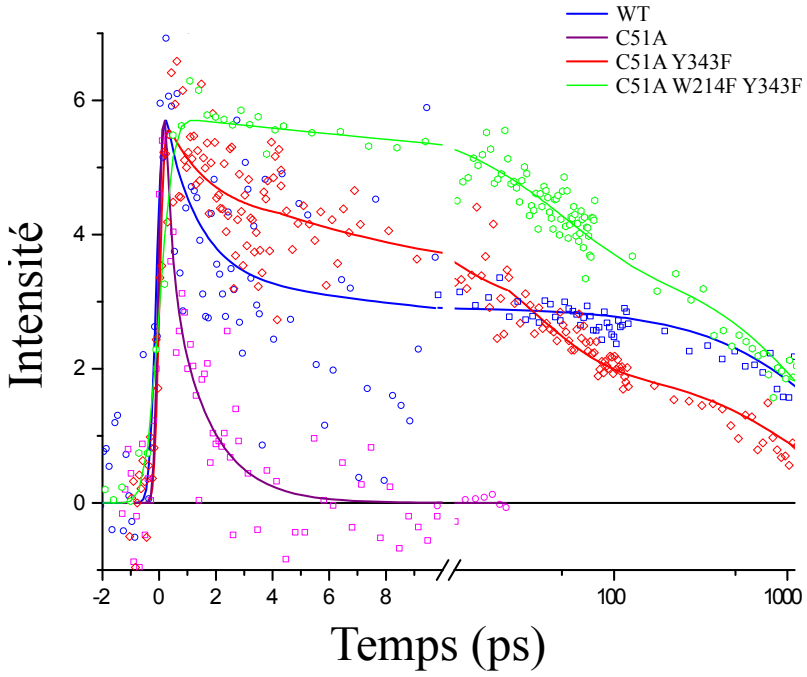

FIGURE 4.20 – **Décroissance de fluorescence en fonction du temps.** On peut distunguer deux phases de décroissance de fluorescence de la protéine de type sauvage, une rapide (environ 1 ps) et une lente (environ 2 ns). Pour le mutant C51A, on observe principalement une phase de décroissance rapide (1 ps), indiquant une extinction de fluorescence par le micro-environnement de la flavine. Le double mutant Y343F C51A trouve sa fluorescence restaurée, ce qui montre que Y343 est un *quencher* de fluorescence, mais une phase rapide persiste. La décroissance de fluorescence du triple mutant C51A W214F Y343F n'a pas de phase rapide ; le résidu W214 est donc le deuxième *quencher* de fluorescence.

Après observation des données structurales, la tyrosine 394 située à 9 Å du noyau isoalloxazine a été mutée en phénylalanine et aucune différence significative n'a pu être observé avec la protéine de type sauvage ; il en a été de même pour le double

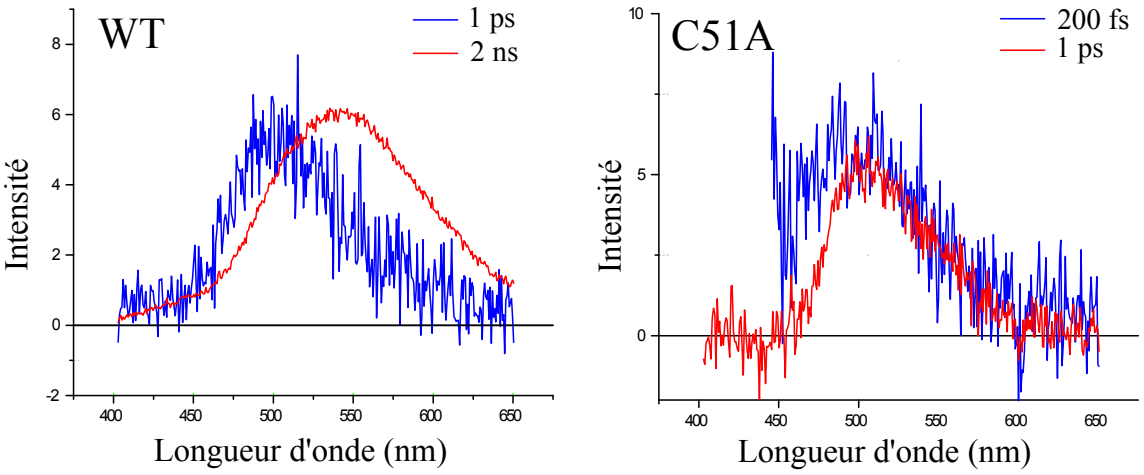

FIGURE 4.21 – **Spectres associés à la décroissance de fluorescence de la protéine de type sauvage (WT) et du mutant C51A.** La décroissance de fluorescence de la protéine de type sauvage a été décomposée en une phase rapide (1 ps) dont le spectre d'émission correspond à celui d'une flavine oxydée, et une phase lente (2 ns) correspondant à l'iminium suspecté.

mutant Y343F Y394F comparé au simple mutant Y343F. En outre un tryptophane situé au milieu d'une boucle non structurée en position 214 semblait un candidat. Dans le cas du triple mutant Y343F W214F C51A, la phase de décroissance rapide n'est pas observable, ce qui indique que le résidu W214 est un second *quencher* de la fluorescence de la flavine.

Dans l'objectif de cerner la nature de l'espèce de flavine absorbant à 400 nm, la protéine de type sauvage a été dialysée et précipitée thermiquement en vue d'une analyse du surnageant par spectrométrie de masses.

4.4 Chromatographie en phase liquide couplée à un spectromètre de masses

4.4.1 Principe

La spectrométrie de masses permet de déterminer les masses des molécules d'une solutions, via la mesure de leurs rapports masses/charges (m/z). Cette technique se base sur le lien entre la masse et l'énergie cinétique, ou sur celui entre la masse et l'oscillation dans un champ électrique ou magnétique. Les molécules, dans un premier temps, seront chargées (étape d'injection) par MALDI (*Matrix assisted Desorption*

ionization) ou ou par electronébulisation (ESI pour *Electro Spray Ionisation*). Ces molécules seront ensuite accélérées dans un vide poussé (10^{-7} Torr) pour enfin être analysées par temps de vol (TOF pour *time of flight*) ou résonnance cyclotronique dans les cas de la FTICR (*Fourrier Transform Ion Cyclotron Resonnance*) et de l'orbitrap.

L'injection

— Le MALDI consiste en une cristallisation des échantillons mélangés à une matrice excitable dans l'ultraviolet par séchage sur plaque. Une fois cette plaque introduite dans un spectromètre ou le vide est fait, elle sera irradiée par un LASER, induisant l'évaporation de la matrice et la libération des molécules contenues dans l'échantillon. A cette étape, la matrice cèdera des protons aux molécules alors mono-chargées (Fig. 4.22).

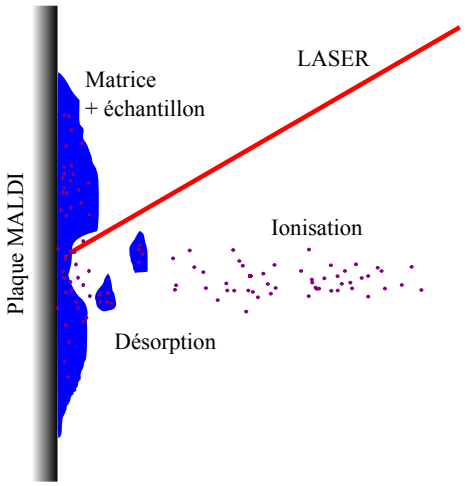

FIGURE 4.22 – **Schéma d'une source MALDI.** L'irradiation par le LASER induit la désorption du mélange matrice + échantillon, puis l'évaporation de la matrice protonne les molécules de l'échantillon.

— L'electronébulisation se fait dans des capillaires de section micrométrique soumis à une tension électrique. L'échantillon est pulvérisé sous forme d'aérosol et les molécules qu'il contient sont désolvatées et multi-chargées (Fig. 4.23).

L'analyse

— Le temps de vol (TOF pour *Time Of Flight*) permet de discriminer les ions en fonction de leur masses. Les ions reçoivent tous une quantité d'énergie

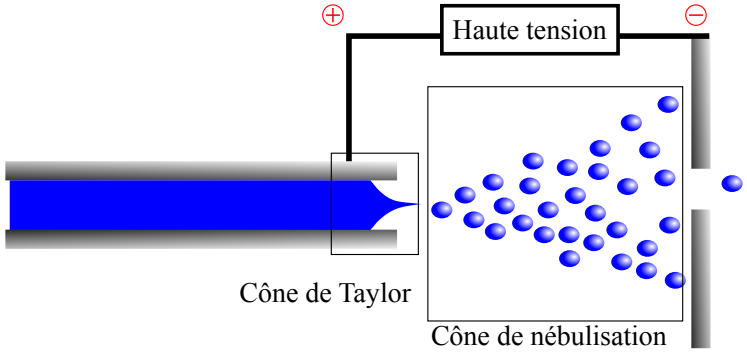

FIGURE 4.23 – **Schéma d'une source électronébulisante.**La haute tension entre le capillaire et l'analyseur sous vide permet la désolvatation des molécules alors chargées.

cinétique égale ($Ec = 1/2mv^2$), et la mesure de leurs temps d'arrivée au détecteur permet d'en déduire leurs masses.

— Les quadrupôles et pièges ioniques filtrent les ions en leurs imposant un champ électrique. Les ions y oscillent et sont éjectés si leurs rapport masse sur charge est trop éloigné de leurs zone de stabilité dans le champ. Ces deux systèmes sont couplés à des détecteurs à dynodes en cascade, permettant une amplification du signal lors d'un choc avec un ion.

— FTICR et l'orbitrap piègent les ions à l'aide de champs, magnétiques pour le FTICR (*Fourier Transform Ion Cyclotron Resonance*), et électriques pour l'orbitrap. Les ions ainsi captés par l'analyseur y oscillent et c'est leur fréquence de rotation qui permet d'établir leur rapport masse sur charge.

4.4.2 Indication d'un intermédiaire covalent FAD-ARNt par LC-MS

Après dénaturation, le spectre d'absorption de la flavine contenue dans le surnageant était caractéristique de celui d'une flavine oxydée. Ces dernières observations appuyaient l'hypothèse selon laquelle l'ion iminium postulé est stabilisé par le micro-environnement de la protéine. En outre, le rapport des absorbances à 260 et 450 nm était significativement plus important que selui d'une solution de FAD, suggérant la présence d'acides nucléiques dans le surnageant. La chromatographie en phase liquide couplée à un spectromètre de masses (LC-MS) a permis d'identifier des molécules de la masse des différents ribonucléotides constituants l'ARN, ainsi que de la ribothymidine (Fig .4.24).

On constate des pics pour les masses du CMP, GMP, et de l'UMP pour lesquels l'élution coïncide. En outre, l'AMP présente un profil d'élution différent avec un pic à 2,9 minutes ainsi que d'autres pics entre 8 et 9 minutes d'élution ; une partie de l'AMP est co-éluée avec les différents ribonucléotides et l'autre partie est co-éluée avec la flavine. Contrairement aux autres nucléotides, l'élution du TMP coïncide avec celle de la flavine. Cette dernière observation nous a poussé à croire en la présence d'un intermédiaire covalent entre la flavine et l'ARN, stable à la température de dénaturation et rompu lors de l'ionisation.

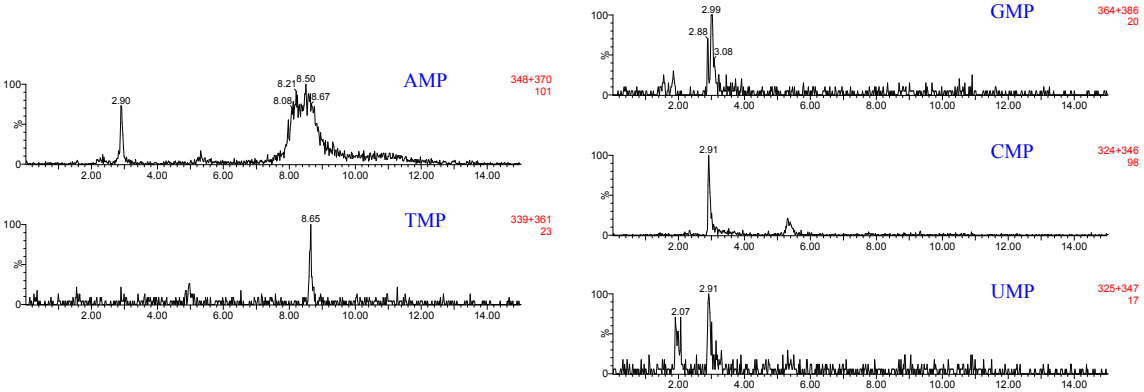

FIGURE 4.24 – **Résultats de LC-MS du surnageant obtenu après précipitation d'une préparation de TrmFO.** Les ribonucléotides caractéristiques de l'ARN sont élués à 2,9 min, l'AMP est aussi présente après 8 min d'élution, correspondant à la flavine. Du TMP est aussi présent et élué à 8,6 minutes (en même temps que la flavine), laissant suggérer un intermédiaire covalent entre la flavine et une thymidine.

4.5 Conclusions et ouvertures

Ce chapitre a pu résumer des travaux et réflexions sur les flavoprotéines méthyletransférases en général, et TrmFO plus particulièrement. Bien que les enzymes ThyX et TrmFO catalysent des réactions relativement similaires, quelques différences sont à noter.

Des données structurales de ThyX co-cristallisée avec dUMP et MTHF sont disponibles (Fig. 4.25). On peut constater un empilement entre dUMP, FAD, MTHF et une histidine en position 53 chez *Thermotoga maritima*, ce qui suggère la possibilité d'un transfert direct du MTHF vers le dUMP via le noyeau isoalloxazine de la flavine.

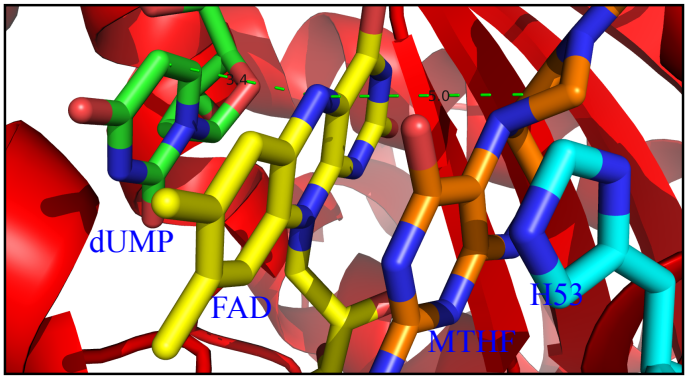

FIGURE 4.25 – **Structure du site actif de l'enzyme ThyX de *Thermotoga maritima* co-cristallisée avec FAD, dUMP et MTHF.** La co-structure (code PDB : 4GT9) a été alignée à celle du tétramère (code PDB : 1O24) et le résidu H53 du tétramère a été réorienté pour mettre en évidence son interaction d'empilement avec le MTHF. La structure montre que les deux substrat peuvent coexister au sein du site actif suggérant un mécanisme bibi aléatoire.

En ce qui concerne TrmFO, aucune donnée structurale de l'enzyme en présence du substrat n'est, pour le moment, disponible. La structure de TrmFO co-cristallisée avec MTHF et FAD ne laisse pas de place pour une uridine supplémentaire (Fig. 4.27 (A)). On peut spéculer un mécanisme de type ping-pong pour l'enzyme TrmFO (Fig. 4.26), à savoir : un transfert de méthylène du MTHF vers la flavine dans un premier temps suivi de la sortie du THF du site actif de l'enzyme avant le transfert du méthylène vers l'uridine 54. Cette proposition est appuyée par une étude préalable sur TrmFO de *Bacillus subtilis*, selon laquelle la protéine fraîchement purifiée peut méthyler efficacement l'ARNt sans ajout de MTHF [48].

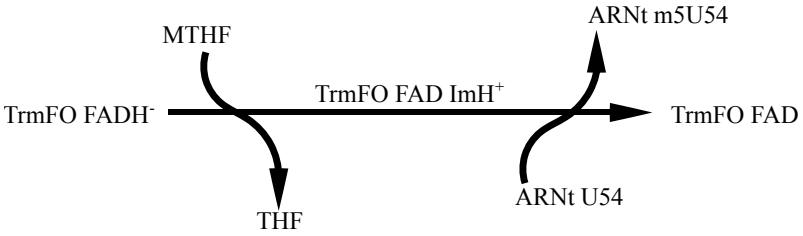

FIGURE 4.26 – **Mécanisme enzymatique de TrmFO.** Dans un premier temps a lieu le transfert d'un groupement méthylène sur le noyau isoalloxazine de la flavine, menant à la formation d'un ion iminium ; dans un deuxième temps, le groupement méthylène sera transféré sur l'ARNt

Pour ce faire, le groupement méthyle transféré au noyeau isoalloxazine de la flavine doit être stabilisé par le micro-environnement de l'enzyme. Ce travail a pu

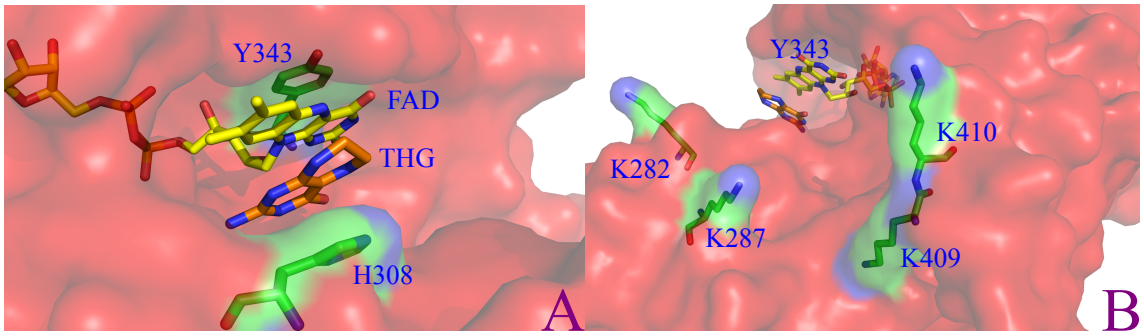

FIGURE 4.27 – **Structure de l'enzyme TrmFO de *Thermus thermophilus* co-cristallisée avec MTHF et FAD** (code PDB : 3G5R). La configuration structurale suggère fortement une impossibilité de co-existance de l'uridine de l'ARNt et du MTHF au sein du site actif. Il en découle que TrmFO catalyserait la méthylation de l'uridine 54 selon un mécanisme de type ping-pong (A). Après transfert du groupement méthylène sur le noyeau isoalloxazine, l'ARNt peut être fixé par les lysines 409 et 410, puis ajusté par les lysines 282 et 287 (B).

mettre en évidence l'importance des résidus H408, Y343 et C51 de l'enzyme TrmFO de *Thermus thermophilus* pour la formation, ou la stabilisation, d'une espèce chimique inhabituelle de flavine absorbant à 400 nm que l'on suppose être un ion iminium. Cette hypothèse impliquerait l'existence d'un équilibre entre une forme d'adduit FAD-enzyme via un pont méthylène entre l'azote 5 et la cystéine 51, et une forme réactive chargée positivement, stabilisée par la tyrosine 343.

En outre, ThyX catalyse la méthylation du dUMP libre tandis que l'uridine ciblée par TrmFO se trouve au sein de la boucle T-Ψ de L'ARNt et intéragit avec une adénosine. On peut supposer la nécessité d'un ajustement conformationnel du substrat pour rendre accessible le carbone 5 de l'uridine au site actif de l'enzyme TrmFO.

Ce travail a permis d'établir clairement le rôle des résidus K409 et K410 dans la fixation du substrat par TrmFO, et soulever la question du rôle des résidus K282 et K287 (Fig. 4.27 (B)). En effet, la mutation de ces deux derniers en alanine n'a pas eu d'influence notable sur l'interaction entre la protéine et son substrat minimum, tout en annihilant l'activité catalytique. Ces dernières observations font des résidus K282 et K287 des candidats intéressants au rôle d'ajustement structurel du substrat.

Des informations structurales de TrmFO co-cristallisée avec son substrat ribonucléique permettraient de confirmer ou d'infirmer l'hypothèse énoncée précédemment. Les résultats de LC-MS présentées dans ce chapitre suggèrent la présence d'un intermédiaire covalent stable entre la flavine et un substrat ARN dans une préparation

de TrmFO de *Thermus thermophilus*. Il devrait être possible de piéger cet intermédiaire covalent (éventuellement avec le substrat minimal) afin d'en résoudre la structure.

Conclusions générales et perspectives

Ce travail, d'abord motivé par une démarche de biologie synthétique visant à explorer des voies artificielles de biosynthèse du thymidylate, s'est vite heurté à des limites. Les voies de biosynthèse du thymidylate constituent un système ancestral, particulièrement rigide, qui s'est avéré difficile à contourner.

Métabolisme des pyrimidines chez *E.coli*

Dans un premier temps, la séquentialité des réactions métaboliques menant à la synthèse *de novo* du thymidylate ne semblait pas optimale. Cette séquentialité ne trouvant *a priori* pour seule explication que l'ordre d'apparition de ses différents composants au cours de l'évolution, une voie alternative a été proposée. Cette potentielle voie orthogonale consiste en une méthylation de l'uridine antérieure à sa réduction en 2'désoxyuridine. Pour tester cette voie, un sauvetage du phénotype d'auxotrophie pour la thymidine par la ribothymidine a été réalisé.

J'ai pu identifier la voie métabolique permettant ce sauvetage comme impliquant la désoxythymidine phosphorylase. Le mécanisme de sauvetage par la ribothymidine se fait donc en deux temps : une phosphorolyse de la ribo-thymidine en ribose-1-phosphate et en thymine, puis la formation d'une liaison n-glycosidique entre thymine et 2'désoxyribose-1-phosphate dont le produit est la thymidine. Cette observation indique une incapacité de la thymidine kinase, ou des ribonucléotides réductases, à prendre en charge la ribothymidine.

La spécificité des ces enzymes à l'encontre de la ribothymidine pourrait s'expliquer par le risque d'incorporer de la thymidine dans l'ARN au cours de la traduction. On peut facilement envisager les effets particulièrement délétères d'une éventuelle incorporation de thymidine dans l'ARN : si la thymidine pouvait s'incor-

porer dans l'ARN au même titre que l'uridine et sans discrimination, c'est toute la structure macromoléculaire des ARN qui risquerait de s'en trouver perturbée. En effet, contrairement à la thymidine, l'uridine peut être impliquée dans de nombreux types d'interactions stabilisant la structure des ARN [123].

Néanmoins, dans certains cas d'interaction rares et spécifiques, la présence d'une thymidine est plus favorable que celle d'une uridine. La spécificité de ces cas explique pourquoi la thymidine dans l'ARN n'est synthétisée que post-transcriptionnellement par des enzymes dédiées à ces fonctions.

Évolution dirigée de TrmFO en thymidylate synthase

Au cours de ce travail, j'ai eu l'occasion de tenter de modifier une de ces enzymes, TrmFO, pour en annihiler la spécificité envers les ARNt et la rendre catalytiquement active sur de l'uridine libre. Un système de sélection a pu être développé à cette fin, consistant en un clonage de gènes issus de mutagénèse aléatoire, d'amplification de la banque de mutants et, de transformation d'une souche d'*E.coli* auxotrophe pour la thymidine étalée sur milieu sélectif. Les essais, bien que n'ayant pas aboutis, n'impliquent pas nécessairement l'impossibilité de leur aboutissement. Ainsi, les limites du protocole ont pu être mises en évidence. Un clonage plus efficace permettrai la sélection du gène complémentant le phénotype d'auxotrophie pour la thymidine parmi un plus grand nombre de mutants. D'autre part, et de manière plus générale, une somme de contraintes qui s'appliquent aux expériences de mutagénèse aléatoire pourraient être contournées. L'encodage d'un gène peut avoir une influence significative sur l'issue d'une campagne d'évolution dirigée. Ce travail a présenté l'occasion de proposer un programme permettant d'optimiser l'encodage d'un gène en vue d'un campagne d'évolution dirigée (écrit par Simon Dadoun et dont le code source est exposé en annexe). Au vu de la baisse continuelle des tarifs de synthèse de gènes, ce programme pourrait apporter une valeur ajoutée non négligeable aux campagnes d'évolution dirigée, et éventuellement aider à l'évolution *in vitro* du gène *trmFO* en un gène complémentant le phénotype d'auxotrophie pour la thymidine.

Caractérisation biochimique et spectrale de TrmFO

Ceci dit, une manière plus rationnelle d'arriver à cet objectif a été de se pencher sur les différences mécanistiques entre TrmFO et la thymidylate synthase ThyX. TrmFO et ThyX sont les représentants d'une classe d'enzymes récemment

découvertes, les méthyle-transférases flavine-dépendantes. Bien que les réactifs soient similaires (MTHF comme donneur de carbone, NADPH comme réducteur), certaines différences restent notables. TrmFO méthyle l'udirine d'un ARN tandis que ThyX méthyle du dUMP libre. Il en découle que TrmFO utilise un mécanisme permettant de discriminer le substrat. Ce travail a permis d'identifier l'implication de deux résidus, les lysines 409 et 410, de l'enzyme TrmFO de *Thermus thermophilus*, dans la fixation d'un substrat minimal. Les rôles de deux autres résidus, critiques pour l'activité catalytique, les lysines 282 et 287, demeurent indéfinis ; ces deux derniers n'étant pas impliqués dans la fixation du substrat, ils pourraient l'être dans son ajustement conformationnel. Une co-structure de TrmFO et son substrat permettrait probablement d'appuyer ou réfuter cette hypothèse. Dans un avenir plus proche, un modèle de *docking* de TrmFO à son substrat ribonucléique (un ARNt) offrira des perspectives intéressantes [4].

Une autre différence entre ThyX et TrmFO réside en leurs mécanismes catalytiques. Une structure de ThyX co-cristallisée avec dUMP, MTHF et FAD montre une co-existence des trois substrats au sein du site actif, suggérant la possibilité d'un transfert "direct" du méthyle depuis le MTHF vers le dUMP via la flavine. Dans le cas de TrmFO de *Bacillus subtilis*, un intermédiaire capable de méthyler de l'ARN sans ajout de MTHF, a pu être isolé au cours d'une étude préalable [47]. Cette observation ouvre la possibilité d'un mécanisme de type ping-pong pour TrmFO. Les observations recueillies au cours de ce travail vont dans ce sens. La protéine TrmFO de *Thermus thermophilus* exprimée par *Escherichia coli* absorbe la lumière entre 200 nm 600 nm de façon singulière. Une composante du spectre d'absorption à 400 nm, inhabituelle chez les flavoprotéines, a été observée.

Devant cette particularité, l'hypothèse suggérée est celle d'un ion iminium, formé au cours de l'expression de la protéine. L'iminium, (groupement méthylène sur l'azote 5 du noyau isoalloxazine) constitue un intermédiaire réactionnel postulé pour chacune des méthyltransférases flavine-dépendantes, selon les mécanismes catalytiques proposés à ce jour. Cet iminium pourrait être l'intermédiaire réactionnel formé après action du MTHF. Selon nos observations, cet intermédiaire serait stabilisé par deux résidus du site actif, la cystéine 51 et la tyrosine 343. J'ai mis en évidence une photo-conversion de cet intermédiaire en une espèce de flavine alkylée, ceci selon un mécanisme qui reste à élucider. L'intermédiaire postulé a été caractérisé en spectro-

4. Un modèle de docking à déjà été proposé [124] ne faisant pas intervenir les résidus 409 et 410. Les observation rapportées par ce travail permettront d'affiner ce modèle.

scopie d'absorption, en spectroscopie de fluorescence à l'état stationnaire ou résolue en temps. La spectroscopie ultra-rapide a, en outre, fourni certaines informations sur la dynamique moléculaire de TrmFO. La fluorescence de la flavine oxydée de TrmFO peut être éteinte par la tyrosine 343 et par le tryptophane 214 situé sur une boucle flexible (qui n'apparaît pas dans les données structurales) de TrmFO. Cette boucle flexible peut donc contacter le site actif de TrmFO.

Des simulations dynamiques sont en cours et permettront certaines précisions sur le rôle des différents résidus de l'enzyme TrmFO, et, plus généralement, sur son mécanisme catalytique. La conaissance de TrmFO permettra de mieux caractériser l'activité de transfert de méthyle, une partie récemment découverte de l'éventail des réactions catalysées par les flavo-enzymes.

Bibliographie

[1] Erwin Schrödinger. *Qu'est-ce que la vie ? : de la physique à la biologie*. Points Sciences, Paris, 1986. OCLC : 29814507.

[2] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids ; a structure for deoxyribose nucleic acid. *Nature*, 171(4356) :737–738, April 1953.

[3] Walter Gilbert. Origin of life : The RNA world. *Nature*, 319(6055) :618–618, February 1986.

[4] K. B. Chapman and J. W. Szostak. In vitro selection of catalytic RNAs. *Curr. Opin. Struct. Biol.*, 4 :618–622, 1994.

[5] Jack W Szostak. The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry*, 3(1) :2, 2012.

[6] Rut Carballido-López. The bacterial actin-like cytoskeleton. *Microbiol. Mol. Biol. Rev.*, 70(4) :888–909, December 2006.

[7] Qingpin Xiao, Xiaohui Hu, Zhiyi Wei, and Kin Yip Tam. Cytoskeleton Molecular Motors : Structures and Their Functions in Neuron. *Int. J. Biol. Sci.*, 12(9) :1083–1092, 2016.

[8] Philip Woodman. ESCRT-III on endosomes : new functions, new activation pathway. *Biochem. J.*, 473(2) :e5–8, January 2016.

[9] Linda Johansson, Guro Gafvelin, and Elias S. J. Arnér. Selenocysteine in proteins-properties and biotechnological use. *Biochim. Biophys. Acta*, 1726(1) :1–13, October 2005.

[10] Bing Hao, Weimin Gong, Tsuneo K. Ferguson, Carey M. James, Joseph A. Krzycki, and Michael K. Chan. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*, 296(5572) :1462–1466, May 2002.

[11] Gayathri Srinivasan, Carey M. James, and Joseph A. Krzycki. Pyrrolysine encoded by UAG in Archaea : charging of a UAG-decoding specialized tRNA. *Science*, 296(5572) :1459–1462, May 2002.

[12] M. Nirenberg, P. Leder, M. Bernfield, R. Brimacombe, J. Trupin, F. Rottman, and C. O'Neal. RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proc. Natl. Acad. Sci. U.S.A.*, 53(5) :1161–1168, May 1965.

[13] T. H. Jukes and S. Osawa. The genetic code in mitochondria and chloroplasts. *Experientia*, 46(11-12) :1117–1126, December 1990.

[14] Peter B. Moore and Thomas A. Steitz. The involvement of RNA in ribosome function. *Nature*, 418(6894) :229–235, July 2002.

[15] U. Baumann and J. Oro. Three stages in the evolution of the genetic code. *BioSystems*, 29(2-3) :133–141, 1993.

[16] F. H. Crick. The origin of the genetic code. *J. Mol. Biol.*, 38(3) :367–379, December 1968.

[17] S. L. Miller. A production of amino acids under possible primitive earth conditions. *Science*, 117(3046) :528–529, May 1953.

[18] Daniel Lundin, Gustav Berggren, Derek T. Logan, and Britt-Marie Sjöberg. The origin and evolution of ribonucleotide reduction. *Life (Basel)*, 5(1) :604–636, 2015.

[19] Eduard Torrents, Margareta Sahlin, and Britt-Marie Drottz-Sjöberg. *The ribonucleotide reductase family : genetics and genomics*. Nova Biomedical Books, New York, 2009. OCLC : 701053859.

[20] A. Poole, D. Penny, and B. M. Sjöberg. Confounded cytosine ! Tinkering and the evolution of DNA. *Nat. Rev. Mol. Cell Biol.*, 2(2) :147–151, February 2001.

[21] Hannu Myllykallio, Gerard Lipowski, Damien Leduc, Jonathan Filee, Patrick Forterre, and Ursula Liebl. An alternative flavin-dependent mechanism for thymidylate synthesis. *Science*, 297(5578) :105–107, July 2002.

[22] Frédéric Escartin, Stéphane Skouloubris, Ursula Liebl, and Hannu Myllykallio. Flavin-dependent thymidylate synthase X limits chromosomal DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, 105(29) :9948–9952, July 2008.

[23] Hannu Myllykallio, Damien Leduc, Jonathan Filee, and Ursula Liebl. Life without dihydrofolate reductase FolA. *Trends Microbiol.*, 11(5) :220–223, May 2003.

[24] M. Friedkin and D. Roberts. Conversion of uracil deoxyriboside to thymidine of deoxyribonucleic acid. *J. Biol. Chem.*, 220(2) :653–660, June 1956.

[25] Janet S. Finer-Moore, Daniel V. Santi, and Robert M. Stroud. Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme : thymidylate synthase mutagenesis, function, and structure. *Biochemistry*, 42(2) :248–256, January 2003.

[26] L. Eichinger, J. A. Pachebat, G. Glöckner, M.-A. Rajandream, R. Sucgang, M. Berriman, J. Song, R. Olsen, K. Szafranski, Q. Xu, B. Tunggal, S. Kummerfeld, M. Madera, B. A. Konfortov, F. Rivero, A. T. Bankier, R. Lehmann, N. Hamlin, R. Davies, P. Gaudet, P. Fey, K. Pilcher, G. Chen, D. Saunders, E. Sodergren, P. Davis, A. Kerhornou, X. Nie, N. Hall, C. Anjard, L. Hemphill, N. Bason, P. Farbrother, B. Desany, E. Just, T. Morio, R. Rost, C. Churcher,

J. Cooper, S. Haydock, N. van Driessche, A. Cronin, I. Goodhead, D. Muzny, T. Mourier, A. Pain, M. Lu, D. Harper, R. Lindsay, H. Hauser, K. James, M. Quiles, M. Madan Babu, T. Saito, C. Buchrieser, A. Wardroper, M. Felder, M. Thangavelu, D. Johnson, A. Knights, H. Loulseged, K. Mungall, K. Oliver, C. Price, M. A. Quail, H. Urushihara, J. Hernandez, E. Rabbinowitsch, D. Steffen, M. Sanders, J. Ma, Y. Kohara, S. Sharp, M. Simmonds, S. Spiegler, A. Tivey, S. Sugano, B. White, D. Walker, J. Woodward, T. Winckler, Y. Tanaka, G. Shaulsky, M. Schleicher, G. Weinstock, A. Rosenthal, E. C. Cox, R. L. Chisholm, R. Gibbs, W. F. Loomis, M. Platzer, R. R. Kay, J. Williams, P. H. Dear, A. A. Noegel, B. Barrell, and A. Kuspa. The genome of the social amoeba Dictyostelium discoideum. *Nature*, 435(7038) :43–57, May 2005.

[27] Kamel Djaout, Vinayak Singh, Yap Boum, Victoria Katawera, Hubert F. Becker, Natassja G. Bush, Stephen J. Hearnshaw, Jennifer E. Pritchard, Pauline Bourbon, Peter B. Madrid, Anthony Maxwell, Valerie Mizrahi, Hannu Myllykallio, and Sean Ekins. Predictive modeling targets thymidylate synthase ThyX in Mycobacterium tuberculosis. *Sci Rep*, 6 :27792, June 2016.

[28] Stéphane Skouloubris, Kamel Djaout, Isabelle Lamarre, Jean-Christophe Lambry, Karine Anger, Julien Briffotaux, Ursula Liebl, Hilde de Reuse, and Hannu Myllykallio. Targeting of Helicobacter pylori thymidylate synthase ThyX by non-mitotoxic hydroxy-naphthoquinones. *Open Biol*, 5(6) :150015, June 2015.

[29] Amanda S. Fivian-Hughes, Joanna Houghton, and Elaine O. Davis. Mycobacterium tuberculosis thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid. *Microbiology (Reading, Engl.)*, 158(Pt 2) :308–318, February 2012.

[30] Joshua H. Hunter, Ramesh Gujjar, Cullen K. T. Pang, and Pradipsinh K. Rathod. Kinetics and ligand-binding preferences of Mycobacterium tuberculosis thymidylate synthases, ThyA and ThyX. *PLoS ONE*, 3(5) :e2237, May 2008.

[31] John A. Conrad, Mariliz Ortiz-Maldonado, Samuel W. Hoppe, and Bruce A. Palfey. Detection of intermediates in the oxidative half-reaction of the FAD-dependent thymidylate synthase from Thermotoga maritima : carbon transfer without covalent pyrimidine activation. *Biochemistry*, 53(32) :5199–5207, August 2014.

[32] Tatiana V. Mishanina, Liping Yu, Kalani Karunaratne, Dibyendu Mondal, John M. Corcoran, Michael A. Choi, and Amnon Kohen. An unprecedented mechanism of nucleotide methylation in organisms containing thyX. *Science*, 351(6272) :507–510, January 2016.

[33] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. Cate, and H. F. Noller. Crystal structure of the ribosome at 5.5 A resolution. *Science*, 292(5518) :883–896, May 2001.

[34] Clive Persaud, Ying Lu, Antón Vila-Sanjurjo, Jamie L. Campbell, Jason Finley, and Michael O'Connor. Mutagenesis of the modified bases, m(5)U1939

and psi2504, in Escherichia coli 23s rRNA. *Biochem. Biophys. Res. Commun.*, 392(2) :223–227, February 2010.

[35] J. T. Kealey, X. Gu, and D. V. Santi. Enzymatic mechanism of tRNA (m5u54)methyltransferase. *Biochimie*, 76(12) :1133–1142, 1994.

[36] Akram Alian, Tom T. Lee, Sarah L. Griner, Robert M. Stroud, and Janet Finer-Moore. Structure of a TrmA-RNA complex : A consensus RNA fold contributes to substrate selectivity and catalysis in m5u methyltransferases. *Proc. Natl. Acad. Sci. U.S.A.*, 105(19) :6876–6881, May 2008.

[37] Christian Toft Madsen, Jonas Mengel-Jørgensen, Finn Kirpekar, and Stephen Douthwaite. Identifying the methyltransferases for m(5)U747 and m(5)U1939 in 23s rRNA using MALDI mass spectrometry. *Nucleic Acids Res.*, 31(16) :4738–4746, August 2003.

[38] Tom T. Lee, Sanjay Agarwalla, and Robert M. Stroud. Crystal structure of RumA, an iron-sulfur cluster containing E. coli ribosomal RNA 5-methyluridine methyltransferase. *Structure*, 12(3) :397–407, March 2004.

[39] Benoit Desmolaize, Céline Fabret, Damien Brégeon, Simon Rose, Henri Grosjean, and Stephen Douthwaite. A single methyltransferase YefA (RlmCD) catalyses both m5u747 and m5u1939 modifications in Bacillus subtilis 23s rRNA. *Nucleic Acids Res.*, 39(21) :9368–9375, November 2011.

[40] R. L. Tatusov, E. V. Koonin, and D. J. Lipman. A genomic perspective on protein families. *Science*, 278(5338) :631–637, October 1997.

[41] Roman L. Tatusov, Natalie D. Fedorova, John D. Jackson, Aviva R. Jacobs, Boris Kiryutin, Eugene V. Koonin, Dmitri M. Krylov, Raja Mazumder, Sergei L. Mekhedov, Anastasia N. Nikolskaya, B. Sridhar Rao, Sergei Smirnov, Alexander V. Sverdlov, Sona Vasudevan, Yuri I. Wolf, Jodie J. Yin, and Darren A. Natale. The COG database : an updated version includes eukaryotes. *BMC Bioinformatics*, 4 :41, September 2003.

[42] Henri Grosjean, editor. *DNA and RNA modification enzymes : structure, mechanism, function and evolution*. Molecular biology intelligence unit. Landes Bioscience, Austin, Tex, 2009. OCLC : 845517152.

[43] Jaunius Urbonavicius, Céline Brochier-Armanet, Stéphane Skouloubris, Hannu Myllykallio, and Henri Grosjean. In vitro detection of the enzymatic activity of folate-dependent tRNA (Uracil-54,-C5)-methyltransferase : evolutionary implications. *Meth. Enzymol.*, 425 :103–119, 2007.

[44] Jaunius Urbonavicius, Stéphane Skouloubris, Hannu Myllykallio, and Henri Grosjean. Identification of a novel gene encoding a flavin-dependent tRNA :m5u methyltransferase in bacteria–evolutionary implications. *Nucleic Acids Res.*, 33(13) :3955–3964, 2005.

[45] Hiroshi Nishimasu, Ryuichiro Ishitani, Koki Yamashita, Chikako Iwashita, Akira Hirata, Hiroyuki Hori, and Osamu Nureki. Atomic structure of a

folate/FAD-dependent tRNA T54 methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.*, 106(20) :8180–8185, May 2009.

[46] Djemel Hamdane, Manuela Argentini, David Cornu, Hannu Myllykallio, Stéphane Skouloubris, Gaston Hui-Bon-Hoa, and Béatrice Golinelli-Pimpaneau. Insights into folate/FAD-dependent tRNA methyltransferase mechanism : role of two highly conserved cysteines in catalysis. *J. Biol. Chem.*, 286(42) :36268–36280, October 2011.

[47] Djemel Hamdane, Eduardo Bruch, Sun Un, Martin Field, and Marc Fontecave. Activation of a unique flavin-dependent tRNA-methylating agent. *Biochemistry*, 52(49) :8949–8956, December 2013.

[48] Djemel Hamdane, Manuela Argentini, David Cornu, Béatrice Golinelli-Pimpaneau, and Marc Fontecave. FAD/folate-dependent tRNA methyltransferase : flavin as a new methyl-transfer agent. *J. Am. Chem. Soc.*, 134(48) :19739–19745, December 2012.

[49] Carole Lartigue, Anne Lebaudy, Alain Blanchard, Basma El Yacoubi, Simon Rose, Henri Grosjean, and Stephen Douthwaite. The flavoprotein Mcap0476 (RlmFO) catalyzes m5u1939 modification in Mycoplasma capricolum 23s rRNA. *Nucleic Acids Res.*, 42(12) :8073–8082, July 2014.

[50] Takuo Osawa, Koichi Ito, Hideko Inanaga, Osamu Nureki, Kozo Tomita, and Tomoyuki Numata. Conserved cysteine residues of GidA are essential for biogenesis of 5-carboxymethylaminomethyluridine at tRNA anticodon. *Structure*, 17(5) :713–724, May 2009.

[51] S. Ghisla and V. Massey. Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.*, 181(1) :1–17, April 1989.

[52] B. G. Barman and G. Tollin. Flavine-protein interactions in flavoenzymes. Thermodynamics and kinetics of reduction of Azotobacter flavodoxin. *Biochemistry*, 11(25) :4755–4759, December 1972.

[53] C. Gomez-Moreno, M. Choy, and D. E. Edmondson. Purification and properties of the bacterial flavoprotein : thiamin dehydrogenase. *J. Biol. Chem.*, 254(16) :7630–7635, August 1979.

[54] F. Müller. Flavin radicals : chemistry and biochemistry. *Free Radic. Biol. Med.*, 3(3) :215–230, 1987.

[55] Chad A. Brautigam, Barbara S. Smith, Zhiquan Ma, Maya Palnitkar, Diana R. Tomchick, Mischa Machius, and Johann Deisenhofer. Structure of the photolyase-like domain of cryptochrome 1 from Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.*, 101(33) :12142–12147, August 2004.

[56] V. Massey and G. Palmer. On the existence of spectrally distinct classes of flavoprotein semiquinones. A new method for the quantitative production of flavoprotein semiquinones. *Biochemistry*, 5(10) :3181–3189, October 1966.

[57] Ana Maria Edwards. Structure and general properties of flavins. *Methods Mol. Biol.*, 1146 :3–13, 2014.

[58] W. J. H. van Berkel, N. M. Kamerbeek, and M. W. Fraaije. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J. Biotechnol.*, 124(4) :670–689, August 2006.

[59] Barrie Entsch, Lindsay J. Cole, and David P. Ballou. Protein dynamics and electrostatics in the function of p-hydroxybenzoate hydroxylase. *Arch. Biochem. Biophys.*, 433(1) :297–311, January 2005.

[60] Thomas Heine, Kathryn Tucker, Nonye Okonkwo, Berhanegebriel Assefa, Catleen Conrad, Anika Scholtissek, Michael Schlömann, George Gassner, and Dirk Tischler. Engineering Styrene Monooxygenase for Biocatalysis : Reductase-Epoxidase Fusion Proteins. *Appl. Biochem. Biotechnol.*, November 2016.

[61] Marek Bučko, Peter Gemeiner, Andrea Schenkmayerová, Tomáš Krajčovič, Florian Rudroff, and Marko D. Mihovilovič. Baeyer-Villiger oxidations : biotechnological approach. *Appl. Microbiol. Biotechnol.*, 100(15) :6585–6599, August 2016.

[62] Adolf Baeyer and Victor Villiger. Einwirkung des Caro'schen Reagens auf Ketone. *Berichte der deutschen chemischen Gesellschaft*, 32(3) :3625–3633, October 1899.

[63] Adolf Baeyer and Victor Villiger. Ueber die Einwirkung des Caro'schen Reagens auf Ketone. *Berichte der deutschen chemischen Gesellschaft*, 33(1) :858–864, January 1900.

[64] Elvira Romero, J. Rubén Gómez Castellanos, Andrea Mattevi, and Marco W. Fraaije. Characterization and Crystal Structure of a Robust Cyclohexanone Monooxygenase. *Angew. Chem. Int. Ed. Engl.*, 55(51) :15852–15855, December 2016.

[65] Cátia A. Bonito, Paula Leandro, Fátima V. Ventura, and Rita C. Guedes. Insights into Medium-chain Acyl-CoA Dehydrogenase Structure by Molecular Dynamics Simulations. *Chem Biol Drug Des*, 88(2) :281–292, August 2016.

[66] Karen S. Conrad, Craig C. Manahan, and Brian R. Crane. Photochemistry of flavoprotein light sensors. *Nat. Chem. Biol.*, 10(10) :801–809, October 2014.

[67] G. F. Gardner and J. F. Feldman. The frq locus in Neurospora crassa : a key element in circadian clock organization. *Genetics*, 96(4) :877–886, December 1980.

[68] Brian D. Zoltowski and Kevin H. Gardner. Tripping the light fantastic : blue-light photoreceptors as examples of environmentally modulated protein-protein interactions. *Biochemistry*, 50(1) :4–16, January 2011.

[69] Aba Losi and Wolfgang Gärtner. The Evolution of Flavin-Binding Photoreceptors : An Ancient Chromophore Serving Trendy Blue-Light Sensors. *Annual Review of Plant Biology*, 63(1) :49–72, June 2012.

[70] Julien Herrou and Sean Crosson. Function, structure and mechanism of bacterial photosensory LOV proteins. *Nat. Rev. Microbiol.*, 9(10) :713–723, August 2011.

[71] Maxime T. A. Alexandre, Tatiana Domratcheva, Cosimo Bonetti, Luuk J. G. W. van Wilderen, Rienk van Grondelle, Marie-Louise Groot, Klaas J. Hellingwerf, and John T. M. Kennis. Primary reactions of the LOV2 domain of phototropin studied with ultrafast mid-infrared spectroscopy and quantum chemistry. *Biophys. J.*, 97(1) :227–237, July 2009.

[72] Anna Pfeifer, Teresa Majerus, Kazunori Zikihara, Daisuke Matsuoka, Satoru Tokutomi, Joachim Heberle, and Tilman Kottke. Time-resolved Fourier transform infrared study on photoadduct formation and secondary structural changes within the phototropin LOV domain. *Biophys. J.*, 96(4) :1462–1470, February 2009.

[73] Stephan Braatsch, Mark Gomelsky, Silke Kuphal, and Gabriele Klug. A single flavoprotein, AppA, integrates both redox and light signals in Rhodobacter sphaeroides. *Mol. Microbiol.*, 45(3) :827–836, August 2002.

[74] Shinji Masuda and Carl E. Bauer. AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in Rhodobacter sphaeroides. *Cell*, 110(5) :613–623, September 2002.

[75] Mineo Iseki, Shigeru Matsunaga, Akio Murakami, Kaoru Ohno, Kiyoshi Shiga, Kazuichi Yoshida, Michizo Sugai, Tetsuo Takahashi, Terumitsu Hori, and Masakatsu Watanabe. A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis. *Nature*, 415(6875) :1047–1051, February 2002.

[76] Shinji Masuda. Light detection and signal transduction in the BLUF photoreceptors. *Plant Cell Physiol.*, 54(2) :171–179, February 2013.

[77] Vladimira Dragnea, Matthias Waegele, Septimiu Balascuta, Carl Bauer, and Bogdan Dragnea. Time-resolved spectroscopic studies of the AppA blue-light receptor BLUF domain from Rhodobacter sphaeroides. *Biochemistry*, 44(49) :15978–15985, December 2005.

[78] Magdalena Gauden, Ivo H. M. van Stokkum, Jason M. Key, Daniel Ch Lührs, Rienk van Grondelle, Peter Hegemann, and John T. M. Kennis. Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. *Proc. Natl. Acad. Sci. U.S.A.*, 103(29) :10895–10900, July 2006.

[79] Hiroshi Ishikita. Light-induced hydrogen bonding pattern and driving force of electron transfer in AppA BLUF domain photoreceptor. *J. Biol. Chem.*, 283(45) :30618–30623, November 2008.

[80] María A. Mussi, Jennifer A. Gaddy, Matías Cabruja, Brock A. Arivett, Alejandro M. Viale, Rodolfo Rasia, and Luis A. Actis. The opportunistic human pathogen Acinetobacter baumannii senses and responds to light. *J. Bacteriol.*, 192(24) :6336–6345, December 2010.

[81] Min-Hyung Ryu, Oleg V. Moskvin, Jessica Siltberg-Liberles, and Mark Gomelsky. Natural and engineered photoactivated nucleotidyl cyclases for optogenetic applications. *J. Biol. Chem.*, 285(53) :41501–41508, December 2010.

[82] Manuela Stierl, Patrick Stumpf, Daniel Udwari, Ronnie Gueta, Rolf Hagedorn, Aba Losi, Wolfgang Gärtner, Linda Petereit, Marina Efetova, Martin Schwarzel, Thomas G. Oertner, Georg Nagel, and Peter Hegemann. Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium Beggiatoa. *J. Biol. Chem.*, 286(2) :1181–1188, January 2011.

[83] Aziz Sancar. Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev.*, 103(6) :2203–2237, June 2003.

[84] Robert J. Gegear, Amy Casselman, Scott Waddell, and Steven M. Reppert. Cryptochrome mediates light-dependent magnetosensitivity in Drosophila. *Nature*, 454(7207) :1014–1018, August 2008.

[85] Dominik Heyers, Martina Manns, Harald Luksch, Onur Güntürkün, and Henrik Mouritsen. A visual pathway links brain structures active during magnetic compass orientation in migratory birds. *PLoS ONE*, 2(9) :e937, September 2007.

[86] Emre Deniz and Batu Erman. Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Funct. Integr. Genomics*, September 2016.

[87] F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. The complete genome sequence of Escherichia coli K-12. *Science*, 277(5331) :1453–1462, September 1997.

[88] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A. Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants : the Keio collection. *Mol. Syst. Biol.*, 2 :2006.0008, 2006.

[89] F. W. Studier and B. A. Moffatt. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, 189(1) :113–130, May 1986.

[90] Giuseppe Bertani. Lysogeny at mid-twentieth century : P1, P2, and other experimental systems. *J. Bacteriol.*, 186(3) :595–600, February 2004.

[91] D. Hanahan. Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.*, 166(4) :557–580, June 1983.

[92] M. K. Chaveroche, J. M. Ghigo, and C. d'Enfert. A rapid method for efficient gene replacement in the filamentous fungus Aspergillus nidulans. *Nucleic Acids Res.*, 28(22) :E97, November 2000.

[93] M. M. Bradford. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, 72 :248–254, May 1976.

[94] Arthur Kornberg and Tania A Baker. *DNA replication*. University Science ; [Palgrave, distributor, Sausalito, Calif. ; [Basingstoke, 2005. OCLC : 60419190.

[95] P. P. Cherepanov and W. Wackernagel. Gene disruption in Escherichia coli : TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, 158(1) :9–14, May 1995.

[96] Anne Derbise, Biliana Lesic, Denis Dacheux, Jean Marc Ghigo, and Elisabeth Carniel. A rapid and simple method for inactivating chromosomal genes in Yersinia. *FEMS Immunol. Med. Microbiol.*, 38(2) :113–116, September 2003.

[97] J. A. Mosberg, M. J. Lajoie, and G. M. Church. Lambda red recombineering in Escherichia coli occurs through a fully single-stranded intermediate. *Genetics*, 186(3) :791–799, November 2010.

[98] James A. Sawitzke, Lynn C. Thomason, Nina Costantino, Mikhail Bubunenko, Simanti Datta, and Donald L. Court. Recombineering : in vivo genetic engineering in E. coli, S. enterica, and beyond. *Meth. Enzymol.*, 421 :171–199, 2007.

[99] Ronny Martínez and Ulrich Schwaneberg. A roadmap to directed enzyme evolution and screening systems for biotechnological applications. *Biol. Res.*, 46(4) :395–405, 2013.

[100] J. Cohen. How DNA shuffling works. *Science*, 293(5528) :237, July 2001.

[101] J. L. Lin-Goerke, D. J. Robbins, and J. D. Burczak. PCR-based random mutagenesis using manganese and reduced dNTP concentration. *BioTechniques*, 23(3) :409–412, September 1997.

[102] R. C. Cadwell and G. F. Joyce. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.*, 2(1) :28–33, August 1992.

[103] M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21) :2947–2948, November 2007.

[104] Rajni Verma, Ulrich Schwaneberg, and Danilo Roccatano. MAP(2.0)3d : a sequence/structure based server for protein engineering. *ACS Synth Biol*, 1(4) :139–150, April 2012.

[105] Wenqing Chen, Yan Li, Jie Li, Lian Wu, Yan Li, Renxiao Wang, Zixin Deng, and Jiahai Zhou. An unusual UMP C-5 methylase in nucleoside antibiotic polyoxin biosynthesis. *Protein Cell*, 7(9) :673–683, September 2016.

[106] V. Massey. The chemical and biological versatility of riboflavin. *Biochem. Soc. Trans.*, 28(4) :283–296, 2000.

[107] Antonio Emidio Fortunato, Rossella Annunziata, Marianne Jaubert, Jean-Pierre Bouly, and Angela Falciatore. Dealing with light : the widespread and multitasking cryptochrome/photolyase family in photosynthetic organisms. *J. Plant Physiol.*, 172 :42–54, January 2015.

[108] J. R. Totter and M. J. Cormier. The relation of bacterial luciferase to alternative pathways of dihydroflavin mononucleotide oxidation. *J. Biol. Chem.*, 216(2) :801–811, October 1955.

[109] Kirsty J. McLean, Dominika Luciakova, James Belcher, Kang Lan Tee, and Andrew W. Munro. Biological diversity of cytochrome P450 redox partner systems. *Adv. Exp. Med. Biol.*, 851 :299–317, 2015.

[110] E. W. Evans, C. A. Dodson, K. Maeda, T. Biskup, C. J. Wedge, and C. R. Timmel. Magnetic field effects in flavoproteins and related systems. *Interface Focus*, 3(5) :20130037–20130037, August 2013.

[111] Sergey P. Laptenok, Latifa Bouzhir-Sima, Jean-Christophe Lambry, Hannu Myllykallio, Ursula Liebl, and Marten H. Vos. Ultrafast real-time visualization of active site flexibility of flavoenzyme thymidylate synthase ThyX. *Proc. Natl. Acad. Sci. U.S.A.*, 110(22) :8924–8929, May 2013.

[112] Djemel Hamdane, Vincent Guerineau, Sun Un, and Beatrice Golinelli-Pimpaneau. A catalytic intermediate and several flavin redox states stabilized by folate-dependent tRNA methyltransferase from Bacillus subtilis. *Biochemistry*, 50(23) :5208–5219, June 2011.

[113] Ryota Yamagami, Koki Yamashita, Hiroshi Nishimasu, Chie Tomikawa, Anna Ochi, Chikako Iwashita, Akira Hirata, Ryuichiro Ishitani, Osamu Nureki, and Hiroyuki Hori. The tRNA recognition mechanism of folate/FAD-dependent tRNA methyltransferase (TrmFO). *J. Biol. Chem.*, 287(51) :42480–42494, December 2012.

[114] *Petit Larousse illustré 1979 : 71.000 articles, 5.535 illustrations, 215 cartes en noir, 56 pages en couleurs, dont 26 hors-texte cartographiques, et un atlas en couleurs à la fin de l'ouvrage.* Dictionnaire encyclopédique pour tous. Larousse, Paris, 1979. OCLC : 612351790.

[115] T. Heyduk and J. C. Lee. Application of fluorescence energy transfer and polarization to monitor Escherichia coli cAMP receptor protein and lac promoter interaction. *Proc. Natl. Acad. Sci. U.S.A.*, 87(5) :1744–1748, March 1990.

[116] Joseph R. Lakowicz. *Principles of fluorescence spectroscopy.* Springer, New York, NY, 3. ed., corr. 4. print edition, 2010. OCLC : 700510097.

[117] Kyosuke Sato, Yasuzo Nishina, and Kiyoshi Shiga. Purification of electron-transferring flavoprotein from Megasphaera elsdenii and binding of additional FAD with an unusual absorption spectrum. *J. Biochem.*, 134(5) :719–729, November 2003.

[118] Dominic P. H. M. Heuts, Remko T. Winter, Gerke E. Damsma, Dick B. Janssen, and Marco W. Fraaije. The role of double covalent flavin binding in chito-oligosaccharide oxidase from Fusarium graminearum. *Biochem. J.*, 413(1) :175–183, July 2008.

[119] Russ Hille and Susan M. Miller, editors. *Oxidases, dehydrogenases and related systems.* Number ed. by Russ Hille ... ; Vol. 1 in Handbook of flavoproteins. De Gruyter, Berlin, 2013. OCLC : 930796635.

[120] Guohua Zhao, Robert C. Bruckner, and Marilyn Schuman Jorns. Identification of the oxygen activation site in monomeric sarcosine oxidase : role of Lys265 in catalysis. *Biochemistry*, 47(35) :9124–9135, September 2008.

[121] Sergey P. Laptenok, Patrick Nuernberger, Andras Lukacs, and Marten H. Vos. Subpicosecond Kerr-gate spectrofluorometry. *Methods Mol. Biol.*, 1076 :321–336, 2014.

[122] Etienne-Jules Marey. *Le vol des oiseaux*. G.Masson, 1890.

[123] Jörgen Jonsson, Maria Sandberg, and Svante Wold. The evolutionary transition from uracil to thymine balances the genetic code. *Journal of Chemometrics*, 10(2) :163–170, March 1996.

[124] Ryota Yamagami, Chie Tomikawa, Naoki Shigi, Ai Kazayama, Shin-Ichi Asai, Hiroyuki Takuma, Akira Hirata, Dominique Fourmy, Haruichi Asahara, Kimitsuna Watanabe, Satoko Yoshizawa, and Hiroyuki Hori. Folate-/FAD-dependent tRNA methyltransferase from Thermus thermophilus regulates other modifications in tRNA at low temperatures. *Genes Cells*, May 2016.

Liste des tableaux

1.1 Liste des vingt acides aminés standards et leur nomenclature correspondante. 5
1.2 Principales caractéristiques mécanistiques des différentes enzymes de méthylation des uridines. 30

2.1 Liste des souches d'*E.coli* 31

3.1 Complémentation de la souche FE013 par des analogues nucléosidiques. 48
3.2 Récapitulatif des résultats de complémentation du phénotype d'auxotrophie pour la thymidine 53
3.3 Résidus substitués dans la banque de mutants combinatoire de TrmFO de *Bacillus subtilis*. 57
3.4 Biais de mutagénèse de différentes polymérases utilisées pour la PCR mutagène. 62

Liste des figures

1.1 Schéma représentant la structure de la double hélice d'ADN. 2
1.2 Comparaison des structures de l'ADN et de l'ARN 3
1.3 Structure des différents ribonucléosides, constituants de l'ARN 4
1.4 Tableau matérialisant le code génétique. 6
1.5 Postulat d'un code génétique primitif. 6
1.6 Mécanisme général des ribonucléotides réductases. 7
1.7 Instabilité intrinsèque de la liaison phosphodiester des ARN et mécanisme de son hydrolyse en condition alcaline. 7
1.8 Réaction de désamination spontanée des cytidines. 8
1.9 Réaction catalysée par les thymidylate synthases 9
1.10 Structure de l'enzyme ThyA de *Bacillus subtilis* 10
1.11 Mécanisme catalytique des enzymes de la famille ThyA. 11
1.12 Différentes bactéries pathogènes utilisant une thymidylate synthase ThyX. 12
1.13 Structure de ThyX de *Thermotoga maritima* co-cristallisée avec FAD et dUMP 13
1.14 Comparaison des conformations du dUMP au sein de ThyA d'*Escherichia coli* (code PDB : 1BID) et de ThyX de *Thermotoga maritima* (code PDB : 1O25). 14
1.15 Mécanisme de polarisation du dUMP par les thymidylate synthases ThyX. 14
1.16 Mécanisme catalytique proposé pour les thymidylate synthases ThyX. 15
1.17 Schéma représentant la synthèse d'une protéine en cours d'élongation. 16
1.18 Structures secondaire et tertiaire des ARNt 18
1.19 Structure de l'enzyme TrmA d'*E.coli* co-cristallisée avec un substrat minimal ribonucléique (code PDB : 3BT7). 19
1.20 Mécanisme catalytique des méthyltransférases SAM dépendantes. . . 19
1.21 Structure de TrmFO de *Thermus thermophilus*. 21
1.22 Postulat de mécanisme catalytique pour TrmFO. 22
1.23 Alignement partiel de séquences des protéines MnmG de *Aquifex aeolicus* et TrmFO de *Thermus thermophilus*. 23
1.24 Comparaison des sites actifs de TrmFO et MnmG 23
1.25 Formule des cofacteurs flaviniques. 24

1.26 Formules des différentes formes (cationique, neutre et anionique) de flavines pour chacun de leurs états redox 25
1.27 Spectres d'absorption des flavines en fonction de leur différents états d'oxydoreduction. 26
1.28 Mécanisme général des monooxygénases flavine dépendantes. 27

2.1 Carte du plasmide pKOBEG 33
2.2 Carte du plasmide pCP-20 34
2.3 Carte du plasmide d'expression pQE-80L. 35
2.4 Hybridation des différentes amorces sur le chromosome d'*E.coli* de type sauvage. 40
2.5 Séquence codant pour TrmFO de *Thermus thermophilus* 42
2.6 Schéma du montage de spectroscopie ultra-rapide. 45

3.1 Représentation schématique du métabolisme des pyrimidines. 46
3.2 Structures des analogues nucléosidiques 47
3.3 Réaction catalysée par la désoxythymidine phosphorylase. 48
3.4 Stratégie de construction du double mutant $\Delta deoA\Delta thyA$. 49
3.5 Stratégie de vérification de l'excision de la cassette *kan* par PCR. . . 50
3.6 Vérification de l'excision de la cassette *kan* par migration sur gel d'agarose des produits de PCR. 51
3.7 Amplification de la cassette *aphA-3* à partir de l'ADN génomique de la souche FE013 flanquée par les régions en aval et en amont du gène *thyA* 51
3.8 Stratégie de vérification de l'insertion de la cassette *aphA-3* dans le gène *thyA* 52
3.9 Vérification de l'inactivation du gène $\Delta thyA$ par la cassette *aphA-3* . 52
3.10 Complémentation des différentes souches d'*Escherichia coli* par des analogues nucléosidiques. 54
3.11 Schéma d'une procédure de recombinaison de fragments par homologie de séquences. 55
3.12 Structure de TrmFO de *Thermus thermophilus* (PDB : 3G5R) 57
3.13 Alignement partiel des séquences protéiques de TrmFO de *Bacillus subtilis* et *Thermus thermophilus*. 58
3.14 Probabilité d'avoir la banque de mutants entièrement représentée en fonction du nombre de transformants 59
3.15 Protocole d'évolution dirigée par PCR mutagène de *trmFO* en gène de complémentation du phénotype $\Delta thyA$ 60
3.16 Représentation du code génétique. 61
3.17 Modélisation de l'UMP dans le site actif de ThyX de *Thermotoga maritima* 63
3.18 Schéma récapitulatif des voies classique et hypothétique de biosynthèse du thymidylate. 65

4.1 Différents processus biologiques impliquant des flavoprotéines. 66
4.2 Les trois principaux types de cofacteurs flaviniques. 67
4.3 Les différents spectres d'absorption des flavines en fonction de leurs états d'oxydation ou de réduction. 68
4.4 Structure de la boucle T-Ψ de l'ARNt^{Phe} de levure 68
4.5 Emplacement des lysines suspectées d'être impliquées dans la fixation du substrat par TrmFO de *Thermus thermophilus* (code PDB : 3G5R). 70
4.6 Structures secondaires de l'ARNt^{Phe} de *Thermus thermophilus* et du mini-substrat utilisé en anisotropie de fluorescence. 70
4.7 Analyse SDS-PAGE de TrmFO de type sauvage des mutants C51V, C223A, C51S et S52A. 71
4.8 Schéma du montage permettant l'anisotropie de fluorescence. 72
4.9 Courbes de fixation de TrmFO et différents mutants à son substrat minimal 73
4.10 Fixation de TrmFO de type sauvage à un ARN non structuré. 74
4.11 Étude de l'effet du substrat minimal ribonucléique sur le spectre d'absorption de TrmFO. 75
4.12 Évolution du spectre d'absorption de TrmFO au fil de son illumination. 76
4.13 Représentation du site actif de TrmFO de *Thermus thermophilus* (PDB : 3G5R) 77
4.14 Spectres d'absorption des mutants Y343F et C51A (A) et photoréduction du mutant Y343F (B). 78
4.15 Spectres d'absorption des protéines de type sauvage (WT) et mutante (C223A) avant et après illumination. 79
4.16 Stabilisation de l'intermédiaire iminium par les résidus C51 et Y343. . 79
4.17 Spectres d'émission de fluorescence de différents mutants de la protéine TrmFO. 81
4.18 Spectres d'émission de fluorescence de trois différents mutants pour la cystéine 51. 82
4.19 Chronophotographie du vol du pélican. 83
4.20 Décroissance de fluorescence en fonction du temps. 84
4.21 Spectres associés à la décroissance de fluorescence de la protéine de type sauvage (WT) et du mutant C51A. 85
4.22 Schéma d'une source MALDI 86
4.23 Schéma d'une source électronébulisante. 87
4.24 Résultats de LC-MS du surnageant obtenu après précipitation d'une préparation de TrmFO. 88
4.25 Structure du site actif de l'enzyme ThyX de *Thermotoga maritima* co-cristallisée avec FAD, dUMP et MTHF. 89
4.26 Mécanisme enzymatique de TrmFO. 89
4.27 Structure de l'enzyme TrmFO de *Thermus thermophilus* co-cristallisée avec MTHF et FAD. 90

Annexe : Folate-Dependent Thymidylate-Forming Enzymes : "All roads lead to T"

Folate-Dependent Thymidylate-Forming Enzymes*: „All roads lead to T“*

Pierre Sournia[1], Ursula Liebl[1]* and Hannu Myllykallio[1]*

[1]LOB, Ecole polytechnique, CNRS, INSERM, Université Paris-Saclay, 91128 Palaiseau Cedex, France

*corresponding authors: hannu.myllykallio@polytechnique.edu; tel. +33169335010

ursula.liebl@polytechnique.edu; tel. +33169335048

Introduction

It is now widely believed that ribonucleic acid (RNA) acted as genetic material prior to deoxynucleic acid (DNA). Consequently, transition from the RNA to the DNA world was a key event in the history of Life. Establishment of DNA as genetic material, found nowadays in all free-living cellular organisms and many viruses, required the invention of several enzymatic activities, including those required for converting RNA precursors to DNA precursors, retro-transcription of RNA templates to DNA and replication machineries, required for duplication of the resulting DNA molecules (Forterre *et al.*, 2004).

Modern day metabolic pathways for DNA precursor synthesis provide important clues for understanding how the RNA/DNA transition has taken place. For instance, ribonucleotide reductases (RNRs) are the only enzyme family known to perform *de novo* synthesis of deoxyribonucleotides, the precursors of DNA, via the reduction of ribonucleotides, RNA's building blocks (Torrents *et al.*, 2014). Three different classes of modern ribonucleotide reductases exist in present day organisms. They differ regarding their ways of producing and storing the catalytically necessary cysteinyl radical and are thought to be originating from one common ancestor of modern day RNRs. The homology of the three RNR families was revealed by structural work showing a conserved fold and the conservation of key residues (Lundin *et al.*, 2015). Thus, conversion of ribonucleotides to deoxyribonucleotides is an excellent example of

divergent evolution, where the evolution of an ancestral protein led to three different RNR classes with more specialized mechanisms and functions, used under different growth conditions.

The other important difference between DNA and RNA is the systematic presence of T, and not of U, in DNA, albeit T is also found at specific positions in RNA molecules, where it plays a structural role. The postulated replacement of U with T in DNA via evolutionary tinkering allowed the detection of mutations arising from frequently occurring cytosine deamination (Poole *et al*., 2001). Several different ways of producing T could be imagined, but at least in modern day organisms, free dTMP used for DNA replication is only formed from dUMP by thymidylate synthases. Production of freely soluble ribothymidylate (TMP) would be counter-selected, presumably because it could be potentially integrated into RNA directly, thus creating RNAs with methyl groups (Figure 1), which would be destabilized and likely non-functional. Having said this, UMP methylation at the level of ribonucleotides is nevertheless possible. A paralog of the ThyX enzyme, which is involved in antibiotics synthesis in *Streptomyces viridochromogenes*, can methylate *both* dUMP and UMP in solution (Chen *et al*., 2016).

Here we review recent literature and present our own findings on mechanistic studies of uracil methylation, which has occurred independently several times during evolution. After a brief presentation of the canonical ways of forming T (thymidylate synthase ThyA and tRNA methylase TrmA), special emphasis will be given for studies that have revealed mechanistic similarities and differences between two unrelated, flavin-dependent methyltransferases, ThyX and TrmFO that both use methylenetetrahydrofolate as carbon donor. Despite these enzymes catalyzing an essentially identical chemical reaction, the transfer of a carbon to position 5 of uracil, structural and mechanistic studies have indicated distinct evolutionary origins for these two flavoproteins. Finally, as essential ThyX enzymes are found in many human pathogenic bacteria, but are absent in humans, we also discuss the considerable efforts from several laboratories (including our own) in targeting ThyX proteins in view of developing new anti-microbial compounds.

Evolutionary convergence of thymidylate forming enzymes

A fundamental enzymatic reaction occurring during the metabolism of DNA and RNA in all organisms is the methylation at the C5 position of the uracil ring by site-specific methylases. These modification reactions occur at high frequency per cell division and the resulting thymidine (5-methyluridine)-derivatives are essential for DNA biosynthesis and RNA modification.

The first enzyme found to catalyze this reaction was thymidylate synthase ThyA (EC 2.1.1.45), encoded by *thyA* in most prokarya and by the TYMS gene in humans (Carreras & Santi**,** 1995; Leduc *et al.*, 2004-1). It was demonstrated that ThyA converts 2′-deoxyridine-5′-monophosphate [dUMP] to the essential DNA precursor 2′-deoxythymidine-5′-monophosphate [dTMP].

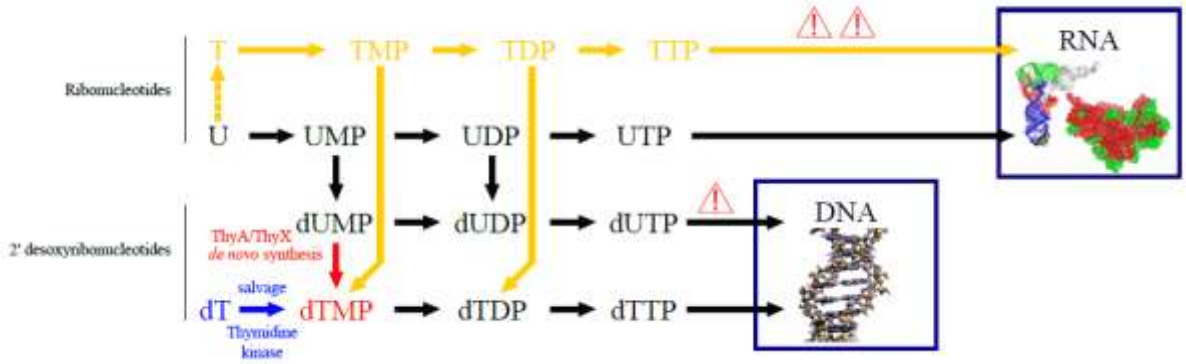

Figure 1: Metabolic pathways required for converting ribonucleotides to deoxyribonucleotides. Notice that three evolutionary related families of ribonucleotide reductases exist, capable of converting all ribonucleotides (A, U, C, G) to corresponding deoxynucleotides (not shown here). A hypothetic pathway converting "free" U to "free" T, would potentially result in the replacement of U by T, which would not be tolerated in structured RNAs, like different tRNAs or rRNAs. ThyA and the flavoenzyme ThyX convert dUMP to dTMP without sharing sequential, structural or mechanistic similarities.

Homodimeric ThyA enzymes catalyze the reductive transfer of the methylene group from methylene-5,6,7,8-tetrahydrofolate (CH_2-H_4folate), which is subsequently reduced by H_4folate to form a methyl group, thus resulting in formation of H_2folate. Independently, an S-adenosyl-l-methionine (SAM)-dependent ribothymidylate synthase was discovered in *Escherichia coli*, initially designated as RumT (Santi *et al.*, 1987) and called TrmA in bacterial systems (Björk GR, 1975; Ny & Björk, 1980). TrmA (and its orthologs in eukarya and archaea) methylate C5 of a uridine that is invariably found at

position 54 of the T-Psi loop of almost all functional tRNAs. Differently from ThyA, TrmA *directly* transfers an activated methyl group from SAM to C5 of U54 in tRNA, thus stabilizing its L-shaped conformation and facilitating decoding of mRNA on the ribosome. Moreover, at least two other SAM-dependent, homologous methyltransferases are known: RlmD (Agarwalla *et al.*, 2002; 2004), catalyzing the site-specific formation of m^5U at position 1939 in 23s rRNA, and RlmC, which catalyzes the site-specific formation of m^5U at position 747 in the same 23s rRNA (Madsen *et al*., 2003). Interestingly, despite the similar outcome of the methylation reactions catalyzed by ThyA and TrmA, the two enzymes use very different reaction mechanisms, as was subsequently confirmed by detailed biochemical and structural characterizations of ThyA and TrmA proteins.

For decades, the canonical enzyme family members ThyA and TrmA were considered the only enzymes capable of (deoxy)ribothymidylate formation. This held true until the discovery of two novel flavoproteins that catalyze the formation of thymidyl groups using CH_2-H_4folate as the carbon donor. The first one discovered, deoxythymidylate synthase ThyX (EC 2.1.1.148; Myllykallio *et al*., 2002), is an essentially prokaryotic enzyme, with only very few eukaryotic representatives known to date. This homotetrameric flavoenzyme uses methylene from CH_2-H_4folate and acquires the reducing hydride from nicotinamide adenine dinucleotide phosphate (NADPH) (Myllykallio et al., 2002; Leduc et al., 2004-2; Koehn *et al.*, 2009). In the case of the RNA metabolism, a gene encoding a new type of bacterial tRNA:m^5U-methylase, dubbed TrmFO, was identified a decade ago (Urbonavicius *et al*., 2005; 2007). Similarly to ThyX proteins, TrmFO flavoproteins also use CH_2-H_4folate as donor of a carbon group that is transferred to the uracil ring. Note that the presence of a tRNA methylation pathway dependent on the cellular reduced folate pools was known several decades prior to the identification of TrmFO proteins (Delk *et al.*, 1975, 1980; Romeo *et al.*, 1974).

Strikingly, bioinformatics and structural characterization of these enzyme families have revealed four independent evolutionary origins for thymidylate forming enzymes, thus providing an excellent example of the evolutionary convergence and versatility of the DNA/RNA modification machinery (Table 1).

Table 1: Mechanistic overview of T-forming methyltransferases. See text for details.

	ThyA	ThyX	TrmFO	TrmA
Substrate	dUMP	dUMP	t-RNA U_{54}	t-RNA U_{54}
Carbon source	CH_2H_4fol	CH_2H_4fol	CH_2H_4fol	SAM
Substrate activation	Nucleophilic (Cys)	Electrostatic polarisation	Nucleophilic (Cys)	Activated methyl group
Reductant	H_4folate	FAD/NADH	FAD/NADH	-

Mechanism of flavin-dependent methylation reactions: Bridging the DNA and RNA worlds

Both, TrmFO and ThyX catalyse the transfer of a carbon group from CH_2H_4folate to uracil rings, resulting in the formation of thymidylate. How is this carbon transfer achieved and what is the role of FAD in this process?

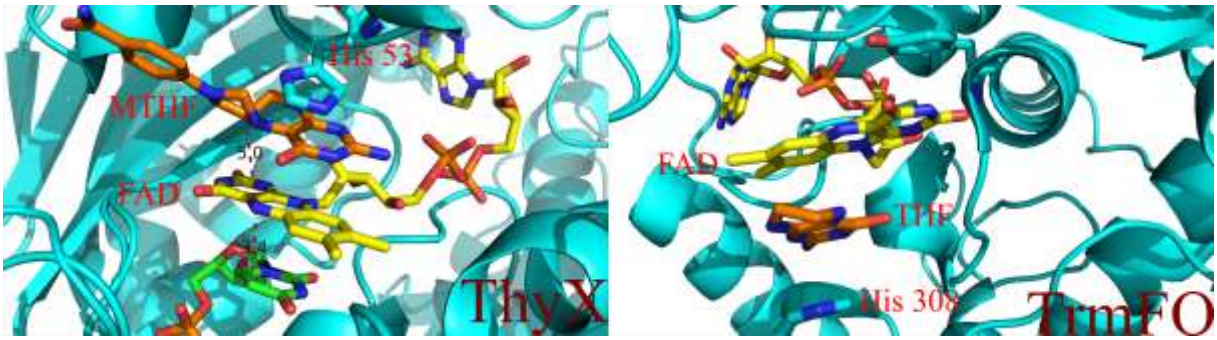

Figure 2. Active site configurations of ThyX (left) and TrmFO (right) proteins as discussed in the text.

TrmFO is a folate-and FAD-dependent enzyme found in many Gram-positive bacteria. Isolated *Bacillus subtilis* TrmFO is a mixture of several different species, one of which is capable of tRNA methylation without added substrates (Urbonavicius *et al.*, 2005, Hamdane *et al.*, 2011). This species has been identified as a covalent enzyme-methylene-FAD adduct by mass spectrometry (Hamdane *et al.*, 2012). The optical spectrum of this adduct has a broad maximum at 360 nm, which also is fully consistent

with an alkylated FAD that is covalently bound to the enzyme via a cysteine residue (C53), as was demonstrated via mutagenesis studies. These results indicate that this cysteine is required for formation of a reaction intermediate, where CH_2 originating from CH_2folate is transferred to tRNA using FAD as transfer agent. Note that this unique flavin-dependent tRNA-methylation can be activated at low pH (protonation of Cys53) and that the actual methylation agent is likely not the above described intermediate, but rather its highly reactive iminium intermediate [(FADH(N5)=CH_2]$^+$ (Hamdane *et al*., 2013). Although this intermediate has not been observed in *B. subtilis* TrmFO, it is expected to absorb at 400-420 nm.

Structural data for the *Thermus thermophilus* enzyme indicate binding of folic acid in the vicinity of the FAD isoribityl ring; fully consistent with the proposed mechanism. The RNA binding site is not known, but tRNA binding is likely to bring about important configurational changes, similar to what is observed for TrmA. Moreover, it is unknown where the accepting uracil is located, but in the *B. subtilis* enzyme, Cys226 is implicated in activation of uracil, as a mini-tRNA substrate becomes covalently attached if it carries fluorine at C5 of U54 of tRNA (Hamdane *et al*., 2011).

Mechanistic studies using ThyX proteins provide an interesting parallel with TrmFO proteins. Here it is known that the complex interplay of several substrates takes place during oxidative and reductive half reactions that minimize the unwanted turnover of ThyX proteins with molecular oxygen (Gattis & Palfey, 2005; Koehn *et al*., 2009; Mishanina *et al.*, 2012; Becker *et al.*, 2014). Recent mechanistic studies have managed to trap a key reaction intermediate under alkaline conditions. This reaction intermediate was identified as a dUMP-CH_2-FAD bridged compound, suggesting that, similar to TrmFO proteins, FAD transfers carbon atoms during enzymatic formation of thymidyl groups (Mishanina *et al.*, 2014, 2016). Here it has also been demonstrated that the 5-deazaflavin reconstituted enzyme is inactive, indicating that nucleophilicity of the FAD co-factor plays a key role in this reaction. The key mechanistic difference between TrmFO and ThyX proteins is that for ThyX enzymes an electrostatic, and not covalent, activation of dUMP has been proposed (Stull *et al*., 2016; Mishanina *et al.,* 2016). This proposal is fully consistent with the structural information available for ThyX proteins, indicating that both, CH_2H_4folate and dUMP are bound in close vicinity, at opposite sides of the FAD isoribityl ring system. Consequently, ThyX and TrmFO proteins share the

same carbon relay mechanism, but differ in terms of how the receiving substrates (U54 or dUMP) are activated.

Flavin-dependent enzymes as anti-microbial targets

As described above**,** the last synthesis step of the essential DNA precursor dTMP is the methylation of dUMP, catalyzed by two distinct families of thymidylate synthases, ThyA and the flavoprotein ThyX (Myllykallio *et al*., 2002, 2003; Graziani *et al.*, 2004, 2006), which are in general mutually exclusive. As the active site configurations and catalytic mechanisms of ThyX and ThyA proteins are drastically different, selective inhibition of ThyX is a highly feasible goal and of particular interest, as ThyX enzymes are present in many enteric pathogens or bacteria causing zoonotic diseases, including *Helicobacter*, *Mycobacteria* and *Clostridium* species. ThyX proteins have attracted broad interest from different scientific communities worldwide and recent genome-wide sequencing studies have established thymidylate synthase ThyX as a virulence factor, as frameshift mutations within *Leptospira interrogans thyX* attenuate virulence in this spirochete, a causative agent of leptosprirosis (Lehmann *et al.,* 2013). These studies also revealed substantial *in vivo* upregulation of *L. interrogans thyX* when compared with *in vitro* cultured bacterial cells. Moreover, replacement of *thyX* by *thyA*, which encodes a catalytically more efficient thymidylate synthase, thus possibly increasing fitness (Escartin *et al*., 2008), has been linked to hypervirulence in *Clostridium difficile* (Knetsch *et al.,* 2011). Systematic studies have associated mutations resulting in thymidylate synthase ThyX overexpression in drug resistance (Zhang *et al*., 2013). Recently it has also been proposed that mutations occurring during the intra-patient evolution of multi-drug resistant *Mycobacterium tuberculosis* complex strains result in up-regulation of ThyX (Merker *et al.*, 2013). Altogether, these observations indicate that the thymidylate metabolism is not only required for growth, but also actively shapes virulence and fitness of pathogenic bacteria. Together with their unique biochemical reaction mechanism, these observations make ThyX proteins attractive drug targets (Myllykallio *et al*., 2002; Choi *et al*., 2016) and extensive efforts to identify ThyX inhibitors are underway. Note also, provided the fact that methylation of U54 is a non-essential tRNA

modification, TrmFO enzymes do not constitute an optimal starting point for the development of new antibiotics compounds.

To date, several classes of ThyX inhibitors have been discovered. Early efforts of targeting ThyX allowed the identification of two distinct classes of compounds, based on a thiazolidine core, as inhibitors of thymidylate synthase X with submicromolar concentrations (Esra Onen *et al*., 2008).

Activity-based screening of a small library of 2500 chemically diverse molecules using identified 1,4-naphthtoquinone (NQ) derivatives as specific inhibitors of ThyX proteins (Basta *et al.*, 2012), without affecting human thymidylate synthase. To date, the most interesting of these inhibitors with a clear dose-response against *Helicobacter pylori* have K_i-values in the range of 200-300 nM (Skouloubris *et al.,* 2015). These molecules, likely preventing formation of the catalytically relevant $FADH_2$, show competitive inhibition with respect to dUMP and a non-competitive inhibition mode with respect to the other substrates, suggesting that CH_2H_4folate and NADPH binding sites do not markedly overlap with the dUMP/inhibitor site. Testing the *in vivo* efficacy of identified non-cytotoxic, non-mitotoxic 2-OH-1,4-NQ inhibitors in a mouse model for *H. pylori* infections identified tight-binding ThyX inhibitors that were tolerated in mice and can be associated with a modest effect in reducing the number of colonizing bacteria, thus providing proof-of-concept that targeting ThyX enzymes is a highly feasible strategy for the development of therapies against *H. pylori* and other ThyX-dependent pathogenic bacteria.

Differently from the large majority of bacteria, where distribution of *thyX* and *thyA* is mutually exclusive, mycobacteria contain both, *thyA* and *thyX* genes. While both thymidylate synthases are expressed in *M. tuberculosis, m*utational studies show that *thyX* is essential, confirming the enzyme as an attractive drug target (Fivian-Hughes *et al.*, 2012).

5-aryl- and 5-alkynyl-substituted 2'-deoxyuridine monophosphate (dUMP) analogues were synthesized and evaluated as potential inhibitors of mycobacterial ThyX. This led to the identification of a derivative that lacks activity against mycobacterial ThyA and displays an IC(50) value against mycobacterial ThyX of 0.91 μM (Kögler *et al.*, 2011). In a continuation of these studies, 5-substituted 6-aza-dUMP analogues in which the C(6) of the uracil moiety is replaced by a N-atom, thus precluding

the nucleophilic attack of a hydride from $FADH_2$, were synthesized and evaluated as potential inhibitors of the two mycobacterial thymidylate synthases. This led to a derivative exhibiting weak ThyX inhibitory activity (33% inhibition at 50 μM). Introduction of alkyl and aryl groups at C(5) resulted in complete loss of inhibitory activity, whereas the attachment of a 3-(octanamido)prop-1-ynyl side chain retained a weak level of mycobacterial ThyX inhibition (40% inhibition at 50 μM). None of the synthesized derivatives displayed any significant inhibitory activity against mycobacterial ThyA (Kögler *et al*., 2012).

Moreover, the interaction of the known inhibitor 5-(3-octanamidoprop-1yn-1yl)-2'-deoxyuridine-5'-monophosphate with *Mtb* ThyX was explored by molecular modeling and NMR experiments. While the dUMP moiety was found to occupy the cavity of the natural substrate in ThyX, the rest of the ligand extends to the outside between two of its four subunits of the homotetrameric enzyme and inhibitory effects of up to 43 % were found on ThyX at 50 μM. Guided by the results of the modeling and NMR studies, and inspired by the success of acyclic antiviral nucleosides, compounds where a 5-alkynyl uracyl moiety is coupled to an acyclic nucleoside phosphonate (ANP) were synthesized and evaluated for inhibition of ThyX. While only modest activity was achieved, this is the first example of an ANP inhibiting ThyX, and a starting point for developing more potent compounds (Parchina *et al*., 2013).

The activity of several N-(3-(5-(2'-deoxyuridine))prop-2-ynyl)octanamide phosphoroamidates as potential anti-tuberculars was evaluated. Several phosphoroamidate derivatives of *N*-(3-(5-(2′-deoxyuridine-5′-monophosphate))prop-2-ynyl)octanamide were designed in order to improve permeability through the mycobacterial cell wall. Biological tests of these prodrugs showed antimycobacterial activity against *M. tuberculosis* H37Rv and *M. bovis* BCG, suggesting that the increased lipophilicity together with the correct molecular size of phosphoroamidate derivatives allowed the compounds to penetrate through the mycobacterial cell wall, thus liberating the monophosphate intracellulary and targeting ThyX (McGuigan *et al*., 2014).

Further systematic SAR and docking studies revealed 5-undecyloxymethyl-2'-deoxyuridine 5'-monophosphate, displaying an IC50 value against ThyX of 8.32 μM and lacking activity against ThyA (Alexandrova *et al*., 2015).

Very recent studies based upon a virtual screening and structural approach revealed the first X-ray crystal structure of *Thermatoga maritima* ThyX in complex with a non-substrate analog inhibitor, identifying 1H-imidazo[4,5-d]pyridazine as scaffold for the development of *Mtb*-ThyX inhibitors (Luciani *et al*., 2016).

Finally, using a combination of cheminformatics and *in vitro* screening, new *Mtb* ThyX inhibitors, 2-chloro-3-(4-methanesulfonylpiperazin-1-yl)-1,4-dihydronaphthalene -1,4-dione) and idebenone were identified, which show modest whole-cell activity. As idebenone has already passed clinical trials for unrelated uses and targets at least partially ThyX in living cells, this is very encouraging for the further development of ThyX inhibitors towards biomedical applications (Djaout *et al*., 2016).

It is obvious that mechanism-based ThyX inhibitors are expected to have a large impact, as several important human pathogens, among them Mycobacteria, Chlamydia or Rickettsia, rely on ThyX for *de novo* dTMP synthesis.

Conclusions and Perspectives

Here we have described how all the roads from U lead to T, with a special focus on flavin-dependent enzymes that relay a methylene using the FAD co-factor from CH_2H_4folate to the uracil ring. These enzymes, ThyX and TrmFO, provide an excellent example how Nature has modified the micro-environment of the FAD co-factor to modulate its chemical reactivity, thus allowing carbon transfer, as well as preventing unwanted turnover of $FADH_2$ with molecular oxygen. Strikingly, as enzymes using the analogous methylation reaction have not been identified in humans, the mechanistic and inhibitory studies discussed here have potential to aid in the development of new anti-microbial compounds. In this respect it is noteworthy that several ThyX inhibitors with anti-microbial activity already been reported, prompting further optimization by exploiting the newly discovered reaction mechanisms.

References

1. Agarwalla S, Kealey JT, Santi DV. et al. Characterization of the 23 S ribosomal RNA m5U1939 methyltransferase from Escherichia coli. J Biol Chem. 2002;277(11):8835–8840.
2. Agarwalla S, Stroud RM, Gaffney BJ. Redox reactions of the iron-sulfur cluster in a ribosomal RNA methyltransferase, RumA: optical and EPR studies. J Biol Chem. 2004;279(33):34123–34129.
3. Alexandrova LA1, Chekhov VO2, Shmalenyuk ER3, Kochetkov SN3, El-Asrar RA4, Herdewijn P4. Synthesis and evaluation of C-5 modified 2'-deoxyuridine monophosphates as inhibitors of M. tuberculosis thymidylate synthase. Bioorg Med Chem. 2015 Nov 15;23(22):7131-7. doi: 10.1016/j.bmc.2015.09.053. Epub 2015 Oct 9.
4. Basta, T., Boum, Y., Briffotaux, J., Becker, H. F., Lamarre-Jouenne, I., Lambry, J. C., Skouloubris, S., Liebl, U., Graille, M., van Tilbeurgh, H.*, et al.* 2012 Mechanistic and structural basis for inhibition of thymidylate synthase ThyX. *Open Biol*. **2**, 120120. (10.1098/rsob.120120)
5. Becker, H. F., Djaout, K., Lamarre, I., Ulmer, J. E., Schaming, D., Balland, V., Liebl, U., Myllykallio, H., Vos, M. H. 2014 Substrate interaction dynamics and oxygen control in the active site of thymidylate synthase ThyX. *Biochem J*. 459, 37-45. (10.1042/BJ20131567)
6. Björk GR. Transductional mapping of gene trmA responsible for the production of 5-methyluridine in transfer ribonucleic acid of Escherichia coli. J Bacteriol. 1975;124(1):92–98.
7. Carreras, C. W., and D. V. Santi**.** 1995. The catalytic mechanism and structure of thymidylate synthase. Annu. Rev. Biochem. 64**:**721-762.
8. Chen W *et al*., An unusual UMP C-5 methylase in nucleoside antibiotic polyoxin biosynthesis, Protein & Cell, September 2016, Volume 7, Issue 9, pp 673–683.
9. Choi, M., Karunaratne, K., Kohen, A. 2016 Flavin-Dependent Thymidylate Synthase as a New Antibiotic Target. Molecules. 2016 May 20;21(5). pii: E654. doi: 10.3390/molecules21050654.
10. Conrad JA, Ortiz-Maldonado M, Hoppe SW, Palfey BA. Detection of intermediates in the oxidative half-reaction of the FAD-dependent thymidylate synthase from Thermotoga maritima: carbon transfer without covalent pyrimidine activation. Biochemistry. 2014 Aug 19;53(32):5199-207.
11. Delk AS, Rabinowitz JC. Biosynthesis of ribosylthymine in the transfer RNA of Streptococcus faecalis: a folate-dependent methylation not involving S-adenosylmethionine. Proc Natl Acad Sci U S A. 1975 Feb;72(2):528-30.
12. Delk AS, Nagle DP Jr, Rabinowitz JC. Methylenetetrahydrofolate-dependent biosynthesis of ribothymidine in transfer RNA of Streptococcus faecalis. Evidence for reduction of the 1-carbon unit by FADH2.J Biol Chem. 1980 May 25;255(10):4387-90.
13. Djaout, K., Singh, V., Boum, Y., Katawera, V., Becker, H. F., Bush, N. G., Hearnshaw, S. J., Pritchard, J. E., Bourbon, P., Madrid, P. B.*, et al.* 2016 Predictive modeling targets thymidylate synthase ThyX in Mycobacterium tuberculosis. *Sci Rep*. 6, 27792. (10.1038/srep27792doi: 10.1099/mic.0.053983-0
14. Escartin, F., Skouloubris, S., Liebl, U., Myllykallio, H. 2008 Flavin-dependent thymidylate synthase X limits chromosomal DNA replication. *Proc Natl Acad Sci U S A*. **105**, 9948-9952. (10.1073/pnas.0801356105)
15. Esra Onen, F., Boum, Y., Jacquement, C., Spanedda, M. V., Jaber, N., Scherman, D., Myllykallio, H., Herscovici, J. 2008 Design, synthesis and evaluation of potent thymidylate synthase X inhibitors. *Bioorg Med Chem Lett*. 18, 3628-3631. (10.1016/j.bmcl.2008.04.080)
16. Fivian-Hughes AS., Joanna Houghton, and Elaine O. Davis *Mycobacterium tuberculosis* thymidylate synthase gene *thyX* is essential and potentially bifunctional, while *thyA*

deletion confers resistance to *p*-aminosalicylic acid Microbiology. 2012 Feb; 158(Pt 2): 308–318.
17. Forterre P, Filée J, Myllykallio H. Origin and Evolution of DNA and DNA Replication Machineries, *dans* L. Ribas de Pouplan (Ed.), The genetic code and the origin of life, Landes Bioscience, 2004, p. 145-168.
18. Gattis SG, Palfey BA. Direct observation of the participation of flavin in product formation by thyX-encoded thymidylate synthase. J Am Chem Soc. 2005 Jan 26;127(3):832-3.
19. Graziani, S., Bernauer, J., Skouloubris, S., Graille, M., Zhou, C. Z., Marchand, C., Decottignies, P., van Tilbeurgh, H., Myllykallio, H., Liebl, U. 2006 Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J Biol Chem*. 281, 24048-24057. (10.1074/jbc.M600745200)
20. Graziani, S., Xia, Y., Gurnon, J. R., Van Etten, J. L., Leduc, D., Skouloubris, S., Myllykallio, H., Liebl, U. 2004 Functional analysis of FAD-dependent thymidylate synthase ThyX from Paramecium bursaria Chlorella virus-1. *J Biol Chem*. 279, 54340-54347. (10.1074/jbc.M409121200)
21. Hamdane, D., Argentini, M., Cornu, D., Myllykallio, H., Skouloubris, S., Hui-Bon-Hoa, G., Golinelli-Pimpaneau, B. 2011 Insights into folate/FAD-dependent tRNA methyltransferase mechanism: role of two highly conserved cysteines in catalysis. *J Biol Chem*. 286, 36268-36280. (10.1074/jbc.M111.256966)
22. Hamdane D, Argentini M, Cornu D, Golinelli-Pimpaneau B, Fontecave M. FAD/folate-dependent tRNA methyltransferase: flavin as a new methyl-transfer agent. J Am Chem Soc. 2012 Dec 5;134(48):19739-45. doi: 10.1021/ja308145p.
23. Hamdane D, Bruch E, Un S, Field M, Fontecave M. Activation of a unique flavin-dependent tRNA-methylating agent. Biochemistry. 2013 Dec 10;52(49):8949-56. doi: 10.1021/bi401387
24. Knetsch CW, Hensgens MP, Harmanus C, van der Bijl MW, Savelkoul PH, Kuijper EJ, Corver J, van Leeuwen HC. Genetic markers for Clostridium difficile lineages linked to hypervirulence. Microbiology. 2011 Nov;157(Pt 11):3113-23. doi: 10.1099/mic.0.051953-0.
25. Koehn, E. M., Fleischmann, T., Conrad, J. A., Palfey, B. A., Lesley, S. A., Mathews, II, Kohen, A. 2009 An unusual mechanism of thymidylate biosynthesis in organisms containing the thyX gene. *Nature*. 458, 919-923. (10.1038/nature07973)
26. Koehn, E. M., Kohen, A. 2010 Flavin-dependent thymidylate synthase: a novel pathway towards thymine. *Arch Biochem Biophys*. 493, 96-102. (10.1016/j.abb.2009.07.016)
27. Kögler M, Vanderhoydonck B, De Jonghe S, Rozenski J, Van Belle K, Herman J, Louat T, Parchina A, Sibley C, Lescrinier E, Herdewijn P. Synthesis and evaluation of 5-substituted 2'-deoxyuridine monophosphate analogues as inhibitors of flavin-dependent thymidylate synthase in Mycobacterium tuberculosis. J Med Chem. 2011 Jul 14;54(13):4847-62. doi: 10.1021/jm2004688. Epub 2011 Jun 9.
28. Kögler M, Busson R, De Jonghe S, Rozenski J, Van Belle K, Louat T, Munier-Lehmann H, Herdewijn P. Synthesis and evaluation of 6-aza-2'-deoxyuridine monophosphate analogs as inhibitors of thymidylate synthases, and as substrates or inhibitors of thymidine monophosphate kinase in Mycobacterium tuberculosis. Chem Biodivers. 2012 Mar;9(3):536-56.
29. Leduc, D., Graziani, S., Meslet-Cladiere, L., Sodolescu, A., Liebl, U., Myllykallio, H. 2004-1, Two distinct pathways for thymidylate (dTMP) synthesis in (hyper)thermophilic Bacteria and Archaea. *Biochem Soc Trans*. **32**, 231-235.
30. Leduc, D., Graziani, S., Lipowski, G., Marchand, C., Le Marechal, P., Liebl, U., Myllykallio, H. 2004-2, Functional evidence for active site location of tetrameric thymidylate synthase X at the interphase of three monomers. *Proc Natl Acad Sci U S A*. 101, 7252-7257. (10.1073/pnas.0401365101)
31. Lehmann JS, Fouts DE, Haft DH, Cannella AP, Ricaldi JN, Brinkac L, Harkins D, Durkin S, Sanka R, Sutton G, Moreno A, Vinetz JM, Matthias MA. Pathogenomic inference of

virulence-associated genes in Leptospira interrogans. PLoS Negl Trop Dis. 2013 Oct 3;7(10):e2468. doi: 10.1371/journal.pntd.0002468.
32. Luciani R, Saxena P, Surade S, Santucci M, Venturelli A, Borsari C, Marverti G, Ponterini G, Ferrari S, Blundell TL, Costi MP.Virtual Screening and X-ray Crystallography Identify Non-Substrate Analog Inhibitors of Flavin-Dependent Thymidylate Synthase. J Med Chem. 2016 Oct 13;59(19):9269-9275.
33. Lundin D, Berggren G, Logan DT, Sjöberg BM. The origin and evolution of ribonucleotide reduction. Life (Basel). 2015 Feb 27;5(1):604-36. doi: 10.3390/life5010604.
34. Madsen CT, Mengel-Jorgensen J, Kirpekar F. et al. Identifying the methyltransferases for m(5)U747 and m(5)U1939 in 23S rRNA using MALDI mass spectrometry. Nucleic Acids Res. 2003;31(16):4738–4746.
35. McGuigan C, Derudas M, Gonczy B, Hinsinger K, Kandil S, Pertusati F, Serpi M, Snoeck R, Andrei G, Balzarini J, McHugh TD[4], Maitra A, Akorli E, Evangelopoulos D, Bhakta S. ProTides of N-(3-(5-(2'-deoxyuridine))prop-2-ynyl)octanamide as potential anti-tubercular and anti-viral agents. Bioorg Med Chem. 2014 May 1;22(9):2816-24. doi: 10.1016/j.bmc.2014.02.056. Epub 2014 Mar 15.
36. Merker M, Kohl TA, Roetzer A, Truebe L, Richter E, Rüsch-Gerdes S, Fattorini L, Oggioni MR, Cox H, Varaine F, Niemann S. Whole genome sequencing reveals complex evolution patterns of multidrug-resistant Mycobacterium tuberculosis Beijing strains in patients. PLoS One. 2013 Dec 6;8(12):e82551. doi: 10.1371/journal.pone.0082551.
37. Mishanina, T. V., Koehn, E. M., Conrad, J. A., Palfey, B. A., Lesley, S. A., Kohen, A. 2012 Trapping of an intermediate in the reaction catalyzed by flavin-dependent thymidylate synthase. *J Am Chem Soc*. 134, 4442-4448. (10.1021/ja2120822)
38. Mishanina, T. V., Corcoran, J. M., Kohen, A. 2014 Substrate activation in flavin-dependent thymidylate synthase. *J Am Chem Soc*. 136, 10597-10600. (10.1021/ja506108b)
39. Mishanina, T. V., Yu, L., Karunaratne, K., Mondal, D., Corcoran, J. M., Choi, M. A., Kohen, A. 2016 An unprecedented mechanism of nucleotide methylation in organisms containing thyX. *Science*. 351, 507-510. (10.1126/science.aad0300)
40. Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., Liebl, U. 2002 An alternative flavin-dependent mechanism for thymidylate synthesis. *Science*. 297, 105-107. (10.1126/science.1072113)
41. Myllykallio, H., Leduc, D., Filee, J., Liebl, U. 2003 Life without dihydrofolate reductase FolA. *Trends Microbiol*. 11, 220-223.
42. Ny T, Björk GR. Cloning and restriction mapping of the trmA gene coding for transfer ribonucleic acid (5-methyluridine)-methyltransferase in Escherichia coli K-12. J Bacteriol. 1980, 142:371-9.
43. Parchina A, Froeyen M, Margamuljana L, Rozenski J, De Jonghe S, Briers Y, Lavigne R, Herdewijn P, Lescrinier E. Discovery of an acyclic nucleoside phosphonate that inhibits Mycobacterium tuberculosis ThyX based on the binding mode of a 5-alkynyl substrate analogue. ChemMedChem. 2013 Aug;8(8):1373-83. doi: 10.1002/cmdc.201300146. Epub 2013 Jul 8.
44. Poole A, Penny D, Sjöberg BM. Confounded cytosine! Tinkering and the evolution of DNA. Nat Rev Mol Cell Biol. 2001 Feb;2(2):147-51.
45. Romeo JM, Delk AS, Rabinowitz JC. The occurrence of a transmethylation reaction not involving S-adenosylmethionine in the formation of ribothymidine in Bacillus subtilis transfer-RNA. Biochem Biophys Res Commun. 1974 Dec 23;61(4):1256-61.
46. Santi DV, Hardy LW. Catalytic mechanism and inhibition of tRNA (uracil-5-) methyltransferase: evidence for covalent catalysis. Biochemistry. 1987;26(26):8599–8606.
47. Skouloubris, S., Djaout, K., Lamarre, I., Lambry, J. C., Anger, K., Briffotaux, J., Liebl, U., de Reuse, H., Myllykallio, H. 2015 Targeting of Helicobacter pylori thymidylate synthase ThyX by non-mitotoxic hydroxy-naphthoquinones. *Open Biol*. **5**, 150015. (10.1098/rsob.150015)

48. Stull, F. W., Bernard, S. M., Sapra, A., Smith, J. L., Zuiderweg, E. R., Palfey, B. A. 2016 Deprotonations in the Reaction of Flavin-Dependent Thymidylate Synthase. *Biochemistry*. **55**, 3261-3269. (10.1021/acs.biochem.6b00510)
49. Torrents E., Ribonucleotide reductases: essential enzymes for bacterial life. Front Cell Infect Microbiol. 2014; 4: 52. doi: 10.3389/fcimb.2014.00052
50. Urbonavicius, J., Skouloubris, S., Myllykallio, H., Grosjean, H. 2005 Identification of a novel gene encoding a flavin-dependent tRNA:m5U methyltransferase in bacteria--evolutionary implications. *Nucleic Acids Res*. **33**, 3955-3964. (10.1093/nar/gki703)
51. Urbonavicius, J., Brochier-Armanet, C., Skouloubris, S., Myllykallio, H., Grosjean, H. 2007 In vitro detection of the enzymatic activity of folate-dependent tRNA (Uracil-54,-C5)-methyltransferase: evolutionary implications. *Methods Enzymol*. 425, 103-119. (10.1016/S0076-6879(07)25004-9)
52. Zhang H, Li D, Zhao L, Fleming J, Lin N, Wang T, Liu Z, Li C, Galwey N, Deng J, Zhou Y, Zhu Y, Gao Y, Wang T, Wang S, Huang Y, Wang M, Zhong Q, Zhou L, Chen T, Zhou J, Yang R, Zhu G, Hang H, Zhang J, Li F, Wan K, Wang J, Zhang XE, Bi L (2013) Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nature genetics* 45**:** 1255-1260

Annexe : présentation de posters aux différents congrès

1. Anti-microbial compounds targeting the alternative thymidylate synthase ThyX.

Hannu Myllykallio, Kamel Djaout, **Pierre Sournia**, Stéphane Skouloubris, **Ursula Liebl** FEBS-EMBO Conference, 30 August-4 September 2014, Paris, France.

ABSTRACT : The growing problem of antibiotic-resistant bacteria in clinical settings points to a need for new anti-infective therapies to address global health problems. However, the rate of new antimicrobial compounds to be developed by the pharmaceutical industry alone will not be sufficient to meet the expected need for the foreseeable future.

We have exploited the alternative thymidylate synthase ThyX proteins discovered in the laboratory as a new target for antimicrobial compounds. This is highly justified as ThyX enzymes are essential, present in many pathogenic bacteria (e.g. Helicobacter, Mycobacteria, Chlamydia, Rickettsia and Clostridium species, 30% of completed bacterial genomes carry thyX), but absent in humans. The active site configurations and catalytic reaction mechanisms differ drastically between human thymidylate synthase and ThyX proteins, thus further facilitating conception of ThyX-specific inhibitors. Moreover, recent genome-wide sequencing studies have established ThyX proteins as a virulence factor in Leptospira and revealed that overexpression of thyX is associated with resistance to the first- and second-line tuberculosis drugs.

We describe how the use of medium-throughput robotized tests for detecting ThyX activity led to the discovery of specific ThyX inhibitors that do not act on human thymidylate synthase. The most interesting ThyX inhibitors inhibit ThyX in genetically modified bacterial strains and show bactericidal activity in the micromolar range against laboratory and clinical strains of Helicobacter and Mycobacteria species. These inhibitors bind in the vicinity of the redox active co-factor flavin adenine nucleotide, in unexpected configuration.

Our molecules are the first non-substrate based ThyX inhibitors with activity against whole cells. Our biochemical, biophysical and structural studies provide means for development of more potent ThyX inhibitors in order to the fight the problem of antibiotics resistance.

2. Active site dynamics of flavin-dependent methylases.

Pierre Sournia, Marten H. Vos, Ursula Liebl, Hannu Myllykallio 40th FEBS Congress (FEBS 2015), 4–9 July 2015, Berlin, Germany

ABSTRACT : Enzymatic methylation of uridyl to form (ribo)thymidyl occurs during the metabolism of DNA and RNA in all organisms. Different pathways exist, implicating thymidylate synthase ThyA that forms the essential DNA precursor thymidylate by methylating deoxyuridine monophosphate (dUMP), and the S-adenosyl-l-methionine-dependent methyltransferase TrmA, which catalyzes the formation of 5-methyluridine at position 54 of tRNA. Recently, two novel folate/flavin-dependent methylases, thymidylate synthase ThyX and tRNA methyltransferase TrmFO, were discovered. Both enzymes use CH2-H4folate as carbon donor and rely on an FAD/NADPH couple as reductant to form a methyl group. In ThyX, all three substrates bind in close proximity to the catalytic FAD group and we have recently demonstrated the real time active site flexibility by studying the dynamics of FAD fluorescence. In order to determine if active site flexibility is a general feature of flavin-dependent methylases, we expressed TrmFO enzymes from mesophilic and thermophilic bacteria. The tRNA substrate of TrmFO proteins is much larger than dUMP, suggesting that substantial flexibility of the active site is required for enzyme function. Different from the mesophilic enzyme, the thermophilic TrmFO purified as highly stable complex with unusual spectral properties. Mass spectrometric analysis revealed very tightly, but non-covalently, bound RNA and FAD linked to the complex. We are currently using the quenching of flavin fluorescence, together with tyrosine mutagenesis and molecular dynamics simulations, to probe the dynamic properties of the active site on different time scales in TrmFO. Our data are expected to have important implications for the role of active site flexibility in multisubstrate enzymes.

3. Ultrafast electron transfer and visualisation of active site dynamics in the flavoenzyme TrmFO.

Lipsa Nag, Pierre Sournia, Ursula Liebl, Marten H. Vos. Symposium TCBIO Transferts de charges, dynamique des biomolécules et fonctions biologiques, mai 2016, Orsay, France

EuCheMS Congress, septembre 2016, Sevilla, Spain

ABSTRACT : Flavoproteins are are a class of enzymes containing flavin nucleotides (FAD and FMN) as cofactors in their active site. They are known to participate in electron transfer reactions in various biological processes including DNA synthesis and DNA repair. With the discovery of the enzyme TrmFO, flavoproteins have now been proven involved in modification of RNA. TrmFO is a folate-dependent tRNA(uracil-54,C5)-methyltransferase. This makes it an important system to study the mechanism of methylation of tRNA.

TrmFO has three aromatic substrates (NADPH, folate, and tRNA) and must be quite flexible to be able to co-accommodate these. In this study, we used electron transfer as a probe for protein dynamics. Time-resolved fluorescence and absorption spectroscopies were employed to visualize the excited state and product state respectively for the investigation the functional dynamics of the active site in TrmFO (from the bacterium Thermus thermophilus). Mutations in the active site yield strong modulations in the excited state lifetimes and their distributions and product state spectral properties. We examined wild type TrmFO along with several active site mutants using an in-house femtoscond Kerr-gate spectrofluorometer and a femtosecond pump-probe spectrometer.

A proximal tyrosine has been identified as the main fluorescence quencher of the femtosecond timescale FAD* excited state, with a contiguous Trp residue also believed to be involved in charge separation. Moreover, we encountered and identified a TyrOH+ product state, the slow re-reduction of which allows for the characterization of the reclusive TyrOH+ transient which is predicted to be an intermediate in many reactions but has not been spectrally identified before.

Annexe : code source de l'utilitaire d'optimisation d'encodage des gènes pour la mutagénèse aléatoire

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"

    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">


<html lang="en" xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">

<head>

  <meta name="generator" content="HTML Tidy for Mac OS X (vers 31 October 2006 - Apple Inc.
build 13), see www.w3.org" />

  <meta name="author" content="Simon Dadoun, Pierre Sournia, Laboratoire d'Optique et
Biosciences, Ecole Polytechnique, France">

  <title>Mutagenesis Bias Fraction</title>

  <meta http-equiv="Content-Type" content="text/html; charset=us-ascii" />

  <link rel="stylesheet" href="includes/stylesheet.css" type="text/css" />

  <script type="text/javascript" src="scripts/sms_common.js">

</script>

  <script type="text/javascript" src="scripts/sms_genetic_codes.js">

</script>

  <script type="text/javascript" src="scripts/trans_map.js">

</script>

  <script type="text/javascript" src="scripts/mutagenesis_bias_fraction.js">

</script>

</head>


<body class="main" onload="document.main_form.main_submit.focus()">

  <!-- MAIN TABLE -->


  <table border="0" cellspacing="2" cellpadding="2">

    <tbody>

      <tr>

        <td class="links">

          <a href="index.html"><img src="includes/logo.png" alt="SMS logo" width="144"
height="56" border="0" /></a> <!-- INSERT LINKS -->

  <div class="category">

    Format Conversion

  </div>


  <div class="program">

    <a href="combine_fasta.html" onmouseover="self.status='Combines multiple FASTA entries
into a single sequence.'; return true;" onmouseout="self.status=' '; return true;">-Combine
FASTA</a><br />

    <a href="embl_fasta.html" onmouseover="self.status='Returns the entire sequence contained
in an EMBL file in FASTA format.'; return true;" onmouseout="self.status=' '; return true;">-
EMBL to FASTA</a><br />
```

```
    <a href="embl_feat.html" onmouseover="self.status='Parses the feature table of an EMBL
file and returns the feature sequences.'; return true;" onmouseout="self.status=' '; return
true;">-EMBL Feature Extractor</a><br />

    <a href="embl_trans.html" onmouseover="self.status='Parses the feature table of an EMBL
file and returns the protein translations.'; return true;" onmouseout="self.status=' '; return
true;">-EMBL Trans Extractor</a><br />

    <a href="filter_dna.html" onmouseover="self.status='Removes non-DNA characters from
text.'; return true;" onmouseout="self.status=' '; return true;">-Filter DNA</a><br />

    <a href="filter_protein.html" onmouseover="self.status='Removes non-protein characters
from text.'; return true;" onmouseout="self.status=' '; return true;">-Filter Protein</a><br
/>

    <a href="genbank_fasta.html" onmouseover="self.status='Returns the entire sequence
contained in a GenBank file in FASTA format.'; return true;" onmouseout="self.status=' ';
return true;">-GenBank to FASTA</a><br />

    <a href="genbank_feat.html" onmouseover="self.status='Parses the feature table of a
GenBank file and returns the feature sequences.'; return true;" onmouseout="self.status=' ';
return true;">-GenBank Feature Extractor</a><br />

    <a href="genbank_trans.html" onmouseover="self.status='Parses the feature table of a
GenBank file and returns the protein translations.'; return true;" onmouseout="self.status='
'; return true;">-GenBank Trans Extractor</a><br />

    <a href="one_to_three.html" onmouseover="self.status='Converts single letter amino acid
codes to three letter codes.'; return true;" onmouseout="self.status=' '; return true;">-One
to Three</a><br />

    <a href="range_extract_dna.html" onmouseover="self.status='Reads a list of positions and
ranges and returns those parts of a DNA sequence.'; return true;" onmouseout="self.status=' ';
return true;">-Range Extractor DNA</a><br />

    <a href="range_extract_protein.html" onmouseover="self.status='Reads a list of positions
and ranges and returns those parts of a protein sequence.'; return true;"
onmouseout="self.status=' '; return true;">-Range Extractor Protein</a><br />

    <a href="rev_comp.html" onmouseover="self.status='Determines the reverse-complement,
reverse, or complement of the sequence you enter.'; return true;" onmouseout="self.status=' ';
return true;">-Reverse Complement</a><br />

    <a href="split_codons.html" onmouseover="self.status='Separates bases according to codon
position.'; return true;" onmouseout="self.status=' '; return true;">-Split Codons</a><br />

    <a href="split_fasta.html" onmouseover="self.status='Converts a FASTA sequence into
multiple sequences.'; return true;" onmouseout="self.status=' '; return true;">-Split
FASTA</a><br />

    <a href="three_to_one.html" onmouseover="self.status='Converts three letter amino acid
codes to one letter codes.'; return true;" onmouseout="self.status=' '; return true;">-Three
to One</a><br />

    <a href="window_extract_dna.html" onmouseover="self.status='Returns DNA sequence segments
specified by a position and window size.'; return true;" onmouseout="self.status=' '; return
true;">-Window Extractor DNA</a><br />

    <a href="window_extract_protein.html" onmouseover="self.status='Returns protein sequence
segments specified by a position and window size.'; return true;" onmouseout="self.status=' ';
return true;">-Window Extractor Protein</a>

  </div>


  <div class="category">

    Sequence Analysis

  </div>


  <div class="program">
```

```
    <a href="codon_plot.html" onmouseover="self.status='Plots codon frequency (according to
the codon table you enter) for each codon in a DNA sequence.'; return true;"
onmouseout="self.status=' '; return true;">-Codon Plot</a><br />

    <a href="codon_usage.html" onmouseover="self.status='Returns a standard codon usage
table.'; return true;" onmouseout="self.status=' '; return true;">-Codon Usage</a><br />

    <a href="cpg_islands.html" onmouseover="self.status='Returns a list of potential CpG
islands.'; return true;" onmouseout="self.status=' '; return true;">-CpG Islands</a><br />

    <a href="dna_mw.html" onmouseover="self.status='Calculates the molecular weight of DNA
sequences.'; return true;" onmouseout="self.status=' '; return true;">-DNA Molecular
Weight</a><br />

    <a href="dna_pattern.html" onmouseover="self.status='Returns positions of the patterns you
enter.'; return true;" onmouseout="self.status=' '; return true;">-DNA Pattern Find</a><br />

    <a href="dna_stats.html" onmouseover="self.status='Returns basic sequence statistics.';
return true;" onmouseout="self.status=' '; return true;">-DNA Stats</a><br />

    <a href="fuzzy_search_dna.html" onmouseover="self.status='Returns sequences that are
identical or similar to a query sequence.'; return true;" onmouseout="self.status=' '; return
true;">-Fuzzy Search DNA</a><br />

    <a href="fuzzy_search_protein.html" onmouseover="self.status='Returns sequences that are
identical or similar to a query sequence.'; return true;" onmouseout="self.status=' '; return
true;">-Fuzzy Search Protein</a><br />

    <a href="ident_sim.html" onmouseover="self.status='Accepts aligned sequences in FASTA
format and calculates the identity and similarity of each sequence pair.'; return true;"
onmouseout="self.status=' '; return true;">-Ident and Sim</a><br />

    <a href="multi_rev_trans.html" onmouseover="self.status='Can be used to predict a DNA
sequence in another species using a protein sequence alignment.'; return true;"
onmouseout="self.status=' '; return true;">-Multi Rev Trans</a><br />

    <a href="mutate_for_digest.html" onmouseover="self.status='Finds DNA sequences that can
easily be converted to a restriction site.'; return true;" onmouseout="self.status=' '; return
true;">-Mutate for Digest</a><br />

    <a href="orf_find.html" onmouseover="self.status='Determines the positions of open reading
frames.'; return true;" onmouseout="self.status=' '; return true;">-ORF Finder</a><br />

    <a href="pairwise_align_codons.html" onmouseover="self.status='Returns the optimal global
alignment for two coding DNA sequences.'; return true;" onmouseout="self.status=' '; return
true;">-Pairwise Align Codons</a><br />

    <a href="pairwise_align_dna.html" onmouseover="self.status='Returns the optimal global
alignment for two DNA sequences.'; return true;" onmouseout="self.status=' '; return true;">-
Pairwise Align DNA</a><br />

    <a href="pairwise_align_protein.html" onmouseover="self.status='Returns the optimal global
alignment for two protein sequences.'; return true;" onmouseout="self.status=' '; return
true;">-Pairwise Align Protein</a><br />

    <a href="pcr_primer_stats.html" onmouseover="self.status='Returns a report describing PCR
primer properties'; return true;" onmouseout="self.status=' '; return true;">-PCR Primer
Stats</a><br />

    <a href="pcr_products.html" onmouseover="self.status='Generates PCR products from a
template and two primer sequences.'; return true;" onmouseout="self.status=' '; return
true;">-PCR Products</a><br />

    <a href="protein_gravy.html" onmouseover="self.status='Returns the grand average of
hydropathy value of protein sequences.'; return true;" onmouseout="self.status=' '; return
true;">-Protein GRAVY</a><br />

    <a href="protein_iep.html" onmouseover="self.status='Returns the predicted isoelectric
point of protein sequences.'; return true;" onmouseout="self.status=' '; return true;">-
Protein Isoelectric Point</a><br />

    <a href="protein_mw.html" onmouseover="self.status='Calculates the molecular weight of
protein sequences.'; return true;" onmouseout="self.status=' '; return true;">-Protein
Molecular Weight</a><br />
```

```
    <a href="protein_pattern.html" onmouseover="self.status='Returns positions of the patterns
you enter.'; return true;" onmouseout="self.status=' '; return true;">-Protein Pattern
Find</a><br />

    <a href="protein_stats.html" onmouseover="self.status='Returns basic sequence
statistics.'; return true;" onmouseout="self.status=' '; return true;">-Protein Stats</a><br
/>

    <a href="rest_digest.html" onmouseover="self.status='Converts the sequence you enter into
restriction fragments.'; return true;" onmouseout="self.status=' '; return true;">-Restriction
Digest</a><br />

    <a href="rest_summary.html" onmouseover="self.status='Returns the number and positions of
restriction sites.'; return true;" onmouseout="self.status=' '; return true;">-Restriction
Summary</a><br />

    <a href="rev_trans.html" onmouseover="self.status='Can be used to convert protein into
DNA.'; return true;" onmouseout="self.status=' '; return true;">-Reverse Translate</a><br />

    <a href="translate.html" onmouseover="self.status='Returns the translation in the reading
frame you specify.'; return true;" onmouseout="self.status=' '; return true;">-Translate</a>

  </div>


  <div class="category">

    Sequence Figures

  </div>


  <div class="program">

    <a href="color_align_cons.html" onmouseover="self.status='Colors a sequence alignment
based on sequence conservation.'; return true;" onmouseout="self.status=' '; return true;">-
Color Align Conservation</a><br />

    <a href="color_align_prop.html" onmouseover="self.status='Colors a protein alignment based
on biochemical properties of residues.'; return true;" onmouseout="self.status=' '; return
true;">-Color Align Properties</a><br />

    <a href="group_dna.html" onmouseover="self.status='Numbers and groups DNA according to
your specifications.'; return true;" onmouseout="self.status=' '; return true;">-Group
DNA</a><br />

    <a href="group_protein.html" onmouseover="self.status='Numbers and groups amino acids
according to your specifications.'; return true;" onmouseout="self.status=' '; return true;">-
Group Protein</a><br />

    <a href="primer_map.html" onmouseover="self.status='Shows PCR primer annealing sites,
translations, and restriction sites.'; return true;" onmouseout="self.status=' '; return
true;">-Primer Map</a><br />

    <a href="rest_map.html" onmouseover="self.status='Shows restriction sites and protein
translations.'; return true;" onmouseout="self.status=' '; return true;">-Restriction
Map</a><br />

    <a href="trans_map.html" onmouseover="self.status='Shows protein translations.'; return
true;" onmouseout="self.status=' '; return true;">-Translation Map</a>

  </div>


  <div class="category">

    Random Sequences

  </div>


  <div class="program">
```

```
    <a href="mutate_dna.html" onmouseover="self.status='Introduces random mutations into DNA
sequences.'; return true;" onmouseout="self.status=' '; return true;">-Mutate DNA</a><br />

    <a href="mutate_protein.html" onmouseover="self.status='Introduces random mutations into
protein sequences.'; return true;" onmouseout="self.status=' '; return true;">-Mutate
Protein</a><br />

    <a href="random_coding_dna.html" onmouseover="self.status='Generates a random coding
sequence of the length you specify.'; return true;" onmouseout="self.status=' '; return
true;">-Random Coding DNA</a><br />

    <a href="random_dna.html" onmouseover="self.status='Generates a random DNA sequence of the
length you specify.'; return true;" onmouseout="self.status=' '; return true;">-Random DNA
Sequence</a><br />

    <a href="random_dna_regions.html" onmouseover="self.status='Replaces regions of the DNA
sequences you enter with random bases.'; return true;" onmouseout="self.status=' '; return
true;">-Random DNA Regions</a><br />

    <a href="random_protein.html" onmouseover="self.status='Generates a random protein
sequence of the length you specify.'; return true;" onmouseout="self.status=' '; return
true;">-Random Protein Sequence</a><br />

    <a href="random_protein_regions.html" onmouseover="self.status='Replaces regions of the
protein sequences you enter with random residues.'; return true;" onmouseout="self.status=' ';
return true;">-Random Protein Regions</a><br />

    <a href="sample_dna.html" onmouseover="self.status='Samples bases from a DNA sequence with
replacement.'; return true;" onmouseout="self.status=' '; return true;">-Sample DNA</a><br />

    <a href="sample_protein.html" onmouseover="self.status='Samples residues from a protein
sequence with replacement.'; return true;" onmouseout="self.status=' '; return true;">-Sample
Protein</a><br />

    <a href="shuffle_dna.html" onmouseover="self.status='Randomly shuffles the DNA sequences
you enter.'; return true;" onmouseout="self.status=' '; return true;">-Shuffle DNA</a><br />

    <a href="shuffle_protein.html" onmouseover="self.status='Randomly shuffles the protein
sequences you enter.'; return true;" onmouseout="self.status=' '; return true;">-Shuffle
Protein</a>

  </div>


  <div class="category">

    Miscellaneous

  </div>


  <div class="program">

    <a href="iupac.html" onmouseover="self.status='IUPAC codes for DNA and protein.'; return
true;" onmouseout="self.status=' '; return true;">-IUPAC codes</a><br />

    <a href="genetic_code.html" onmouseover="self.status='The genetic codes used in the
Sequence Manipulation Suite.'; return true;" onmouseout="self.status=' '; return true;">-
Genetic codes</a><br />

    <a href="browser_compat.html" onmouseover="self.status='Browser compatibility.'; return
true;" onmouseout="self.status=' '; return true;">-Browser compatibility</a><br />

    <a href="mirror.html" onmouseover="self.status='Mirror the Sequence Manipulation Suite.';
return true;" onmouseout="self.status=' '; return true;">-Mirror this site</a><br />

    <a href="mirror.html" onmouseover="self.status='Use the Sequence Manipulation Suite off-
line.'; return true;" onmouseout="self.status=' '; return true;">-Use this site off-
line</a><br />

    <a href="about.html" onmouseover="self.status='About the Sequence Manipulation Suite.';
return true;" onmouseout="self.status=' '; return true;">-About this site</a><br />
```

```
    <a href="acknowledgments.html" onmouseover="self.status='Acknowledgments.'; return true;"
onmouseout="self.status=' '; return true;">-Acknowledgments</a><br />

    <a href="reference.html" onmouseover="self.status='Citing the Sequence Manipulation
Suite.'; return true;" onmouseout="self.status=' '; return true;">-Reference</a>

  </div>

<!-- END INSERT LINKS -->


          <div class="copyright">

            <br />

            The Sequence Manipulation Suite Copyright &copy; 2000, 2004 Paul Stothard. Send
questions and comments to stothard@ualberta.ca

          </div>


          <center>

            <div class="program">

              <a href="index.html" target="_top">home</a>

            </div>

          </center>

        </td><!-- PROGRAM CONTENT -->


        <td class="program">

          <table border="0" cellspacing="0" cellpadding="2">

            <tbody>

              <tr>

                <td class="sms">Sequence Manipulation Suite:</td>

              </tr>


              <tr>

                <td class="title">Mutagenesis Bias Fraction</td>

              </tr>


              <tr>

                <td class="description">

                  "Mutagenesis Bias Fraction" needs a mutagenegis bias as input and weights
for each type of mutation, it generates a codon usage table that can be used as input of
"reverse translate".<br />

                  <br />


                  <form action="" name="main_form" id="main_form">

                   <br />

                   <input type="hidden" />

                   <br />
```

```
                    <div class="warning">

                      Please check the <a href="browser_compat.html">browser compatibility</a>
page before using this program.

                    </div>

<input type="button" value="Submit" name="main_submit"
onclick="try{mutagenesisBiasFraction(document)} catch(e) {alert('The following error was
encountered: ' + e);}" />

<input type="button" value="Log" name="main_log"
onclick="try{mutagenesisBiasFractionLog(document)} catch(e) {alert('The following error was
encountered: ' + e);}" />

<!--<input type="button" value="Clear" onclick="document.forms[0].elements[0].value = ' ';
document.forms[0].elements[4].value = ' '" />-->

<input type="reset" value="Reset" />

                    <ul>

                      <li>Use the <select>

                        <option selected="selected" value="transl_table=1">

                          standard (1)

                        </option>


                        <option value="transl_table=2">

                          vertebrate mitochondrial (2)

                        </option>


                        <option value="transl_table=3">

                          yeast mitochondrial (3)

                        </option>


                        <option value="transl_table=4">

                          mold mitochondrial (4)

                        </option>


                        <option value="transl_table=5">

                          invertebrate mitochondrial (5)

                        </option>


                        <option value="transl_table=6">

                          ciliate nuclear (6)

                        </option>


                        <option value="transl_table=9">

                          echinoderm mitochondrial (9)

                        </option>
```

```
<option value="transl_table=10">
  euplotid nuclear (10)
</option>

<option value="transl_table=11">
  bacterial (11)
</option>

<option value="transl_table=12">
  alternative yeast nuclear (12)
</option>

<option value="transl_table=13">
  ascidian mitochondrial (13)
</option>

<option value="transl_table=14">
  flatworm mitochondrial (14)
</option>

<option value="transl_table=15">
  Blepharisma macronuclear (15)
</option>

<option value="transl_table=16">
  chlorophycean mitochondrial (16)
</option>

<option value="transl_table=21">
  trematode mitochondrial (21)
</option>

<option value="transl_table=22">
  Scenedesmus obliquus mitochondrial (22)
</option>

<option value="transl_table=23">
  Thraustochytrium mitochondrial (23)
```

```
                        </option>

                      </select> <a href="genetic_code.html">genetic code</a>.</li>

                    </ul>

                    Enter mutation probabilities in % or select a preset


<select id="mutagenesis_preset" onclick="try{click_mutagenesis_preset(document)} catch(e)
{alert('The following error was encountered: ' + e);}">

<option selected="selected" value=""></option><option value="Mutazyme II">Mutazyme
II</option><option value="Mutazyme I">Mutazyme I</option><option
value="Taq">Taq</option></select>


                    <br />

A -> T <input type="text" value="0.01" size="3" id="mpat"/> A -> C <input type="text"
value="0.01" size="3" id="mpac"/> A -> G <input type="text" value="0.01" size="3" id="mpag"/>
T -> A <input type="text" value="0.01" size="3" id="mpta"/> T -> C <input type="text"
value="0.01" size="3" id="mptc"/> T -> G <input type="text" value="0.01" size="3" id="mptg"/>
<br />

G -> A <input type="text" value="0.01" size="3" id="mpga"/> G -> T <input type="text"
value="0.01" size="3" id="mpgt"/> G -> C <input type="text" value="0.01" size="3" id="mpgc"/>
C -> A <input type="text" value="0.01" size="3" id="mpca"/> C -> T <input type="text"
value="0.01" size="3" id="mpct"/> C -> G <input type="text" value="0.01" size="3" id="mpcg"/>
<br />

                    <br />

                    Enter weights for score : <br />

stop codon <input type="text" value="2" size="4" id="wstop"/>

silent mutation <input type="text" value="7" size="4" id="wsilent"/>

from proline <input type="text" value="2" size="4" id="wfromproline"/>

to proline <input type="text" value="2" size="4" id="wtoproline"/>

stay proline <input type="text" value="14" size="4" id="wstayproline"/>

diversity <input type="text" value="14" size="4" id="wdiversity"/> <br />

                  </form>*This page requires JavaScript. See <a
href="browser_compat.html">browser compatibility.</a><br />

                  *You can <a href="mirror.html">mirror this page</a> or <a
href="mirror.html">use it off-line</a>.<br />

                  <br />


                  <center>

                    <a href="codon_usage.html" target="_blank">new window</a> | <a
href="index.html">home</a> | <a href="reference.html">citation</a>

                  </center><span class="version">Sun Oct 28 03:52:59 2012</span>

                </td>

              </tr>


              <tr>

                <td><span class="valid">Valid XHTML 1.0; Valid CSS</span></td>

              </tr>
```

```
            </tbody>
          </table><!-- MAIN TABLE END -->
        </td>
      </tr>
    </tbody>
  </table>
</body>
</html>
```

Titre : La méthylation flavine-dépendante d'acides nucléiques : aspects évolutifs, métaboliques, biochimiques et spectroscopiques

Mots clefs : Flavoenzymes, évolution dirigée, spectroscopie optique, ARNt méthyltransferases TrmFO, (ribo)thymidine, anabolisme des nucléotides

Résumé : La méthylation de l'uridine sur son carbone 5 est apparue au cours de l'évolution sous plusieurs formes. Tout d'abord, les thymidylate synthases permettent la synthèse de novo du dTMP, un précurseur essentiel de l'ADN des trois règnes du vivant. Deux familles de thymidylate synthases sont connues à ce jour : ThyA et la flavo-enzyme ThyX, codées par des gènes hétérologues et ayant des structures et mécanismes réactionnels radicalement différents. En outre, cette méthylation de l'uridine est apparue sous forme de modifications post-transcriptionnelles des ARNt et ARNr. Cette thèse vise à questionner les contraintes évolutives ayant menés à l'établissement de cet ordre selon lequel la méthylation de l'uridine sur son carbone 5 est apparue quatre fois de manière indépendante au cours de l'évolution. Dans un premier temps, elle décrit des travaux de génétique bactérienne sur le métabolisme de *E.coli*, plus précisément, l'identification de certains analogues nucléosidiques capables de complémenter le phénotype d'auxotrophie pour la thymidine et les enzymes permettant cette complémentation. Une approche de biologie synthétique en vue d'établir une voie alternative de biosynthèse du thymidylate a aussi été mise en oeuvre. Une technique de sélection de gènes de complémentation du phénotype d'auxotrophie pour la thymidine, issus de mutagénèse aléatoire, a pu être développée. Dans un deuxième temps, des études biochimiques et sppectroscopiques ont été réalisées sur la méthyltransférase flavine-dépendante TrmFO, responsable de la méthylation post-transciptionnelle de l'uridine 54 des ARNt de certains microorganismes. L'implication de résidus spécifiques dans la fixation du substrat a pu être déterminée d'une part, et certains intermédiaires réactionnels potentiels ont été caractérisés spectralement d'autre part. Ces dernières observations s'appuient, en outre, sur des études de spectroscopie résolue en temps et des simulations de dynamique moléculaire afin de mieux comprendre les flavoprotéines en général et les méthyletransférases flavine-dépendantes en particulier.

Title : Flavin-dependent methylation of nucleic acids: evolutionary, metabolic, biochemical and spectroscopic aspects

Keywords : Flavoenzymes, directed evolution, optical spectroscopy, tRNA methyltransferase TrmFO, (ribo)thymidine, nucleotides anabolism

Abstract : Enzymes catalyzing the methylation of uridine at its carbon 5 position have appeared independently in different forms across evolution. Thymidylate synthases ThyA and the flavoprotein ThyX catalyze the de novo synthesis of dTMP, an essential DNA precursor in the three domains of life. They are encoded by heterologous genes and have drastically different structures and reaction mechanisms. On the other hand, this uridine methylation is also performed by tRNA and rRNA post-transcriptional modification enzymes. This thesis assesses the question of the evolutionary constraints that have led independently to four kinds of uridine methylation. The first part describes the identification of a metabolic pathway allowing complementation of thymidine auxotrophy by non-natural nucleoside analogs in *Escherichia coli*. A synthetic biology approach, aiming to establish an alternative pathway for thymidylate biosynthesis, was also implemented and a selection strategy for thymidine auxotrophy-complementing genes was developed. In a second part, biochemical and spectral studies where performed on the flavin-dependent methyltransferase TrmFO, responsible for the post-transcriptional methylation of uridine at the invariant position 54 of tRNA in several microorganisms. The involvement of specific amino acid residues in substrate fixation and/or stabilization of potential reaction intermediates was demonstrated. Their spectral characterization supports previously proposed reaction schemes for flavin-dependent thymidylate forming enzymes. These observations are currently being pursued by parallel approaches combining time-resolved spectroscopy and molecular dynamics simulations, aiming to further our understanding of how flavin mediates the transfer of carbon molecules from folate to uracil rings.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France