



**HAL**  
open science

# Physiological consequences of thymidylate synthase ThyX utilization.

Frédéric Escartin

► **To cite this version:**

Frédéric Escartin. Physiological consequences of thymidylate synthase ThyX utilization.. Life Sciences [q-bio]. Ecole Polytechnique X, 2008. English. NNT : . pastel-00004113

**HAL Id: pastel-00004113**

**<https://pastel.hal.science/pastel-00004113>**

Submitted on 22 Jul 2010

**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.



Thèse de doctorat de l'ÉCOLE POLYTECHNIQUE

Spécialité : **BIOLOGIE**

Présentée par

**Frédéric ESCARTIN**

Pour obtenir le titre de Docteur de l'ÉCOLE POLYTECHNIQUE

# **Etude des conséquences physiologiques de l'utilisation de la thymidylate synthase ThyX**

Sous la direction du Dr Ursula LIEBL et du Dr Hannu MYLLYKALLIO

Soutenue le 26 mai 2008

Jury composé de

Docteur Hilde DE REUSE	Examineur
Docteur Ursula LIEBL	Examineur
Docteur Hannu MYLLYKALLIO	Examineur
Docteur Marie-Agnès PETIT	Rapporteur
Docteur Pablo RADICELLA	Rapporteur
Docteur Emmanuelle SCHMITT	Examineur
Professeur Suzanne SOMMER	Examineur



*"The most exciting phrase to hear in science,  
the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...'"*

*Isaac Asimov*





# Remerciements

En premier lieu, je tiens à remercier Hilde de Reuse, Marie-Agnès Petit, Pablo Radicella, Emmanuelle Schmitt et Suzanne Sommer d'avoir accepté d'évaluer mon travail de thèse.

Durant ma thèse, mon temps aura été partagé entre deux laboratoires et deux directeurs de thèses. Cette situation m'a contraint à organiser mon temps et surtout appris à lutter pour garder mes outils de travail (paillasse, pipetman, bureau, etc.). Le fait de travailler dans deux labos m'a surtout permis de fréquenter deux fois plus de personnes (qu'il va falloir remercier maintenant).

Je souhaite remercier le petit monde du LOB pour m'avoir gentiment accueilli. En particulier, merci :

- Jean-Louis Martin pour m'avoir accueilli dans son laboratoire.
- Ursula, ma directrice « officielle » pour m'avoir encadré durant ces 3 années dans la bonne humeur. Je te remercie pour m'avoir permis de gagner en autonomie, d'encadrer des stagiaires et d'exposer mon travail en congrès. Tu as toujours été présente lors de mes problèmes administratifs répétés.
- Isabelle pour avoir eu le courage de travailler avec moi sur le projet Hp-PyrE, j'espère ne pas t'avoir trop stressée avec les purifs. C'est réellement un plaisir de travailler avec toi, mais il faut arrêter d'apporter du chocolat et des gâteaux dans le bureau, ce n'est pas bon pour ma ligne !
- Sébastien pour les discussions scientifiques et les moments ludiques hors-labo. Nos voyages en amoureux au cœur de l'Europe resteront longtemps de très bons souvenirs.
- Andreea, successivement ma stagiaire puis ma collègue, pour ta bonne humeur, ton sens de l'organisation du labo et les coups de mains que tu m'as apporté. Tu finiras bien un jour par me faire croire que les roumains peuvent être des gens biens.
- Mes compagnons de repas et de café (Claude, Latifa, Isabelle, Andreea, Liem, Claire, Christelle et Laure) pour l'absence de discussion scientifique, transformant le déjeuner en un moment privilégié.
- Mes administratives préférées, Laure et Christelle, pour m'avoir toujours permis de trouver une fin heureuse à mes déboires administratifs.
- La bio-team et en particulier Isabelle, Latifa, Andreea et Taku.
- Grégory pour ta bonne humeur et ton investissement durant ton court séjour au laboratoire.

Mon séjour à l'IGM aura également été très plaisant. Merci :

- Hannu pour m'avoir donné un sujet passionnant et encadré mes recherches. Même si mes méthodes de travail t'ont parfois laissées perplexe, j'ai eu beaucoup de plaisir à travailler avec toi. Merci de m'avoir laissé autant de liberté, tu m'as permis de tester mes propres hypothèses et de « perdre mon temps » à faire des tableaux sous Excel. Tu m'as également appris qu'il ne faut pas se limiter à un champ de recherche et ta volonté à me faire faire de la biochimie n'aura pas été veine.
- Stéphane de m'avoir fait découvrir le labo à mon arrivée et de m'avoir formé à la manipulation de *H. pylori*. Ta rigueur (paranoïa ?) est pour moi un modèle du genre et j'espère qu'elle te maintiendra dans le clan des enseignants-chercheurs « publiants ». Je compte sur toi pour maintenir les apéros improvisés sans lesquels la recherche est si peu de chose.

- Joelle pour la touche féminine et maternelle que tu insuffles dans ce labo. Ta gentillesse et ta franchise me manqueront.
- Yap pour ta bonne humeur quasi-permanente et ta programmation musicale quotidienne.
- Esma pour m'avoir permis de faire de jolies photos de mes bactéries malades.
- Suzanne de m'avoir dirigé vers ce jeune chef d'équipe d'Orsay lors de ma recherche de thèse.
- Les habitants de la cafet' (Joelle, Annie, Françoise, Pascale, Esma, Stéphane, Florence, Christiane, Martine, Claire, Cécile, Magali, ...) pour les cafés du matin et les déjeuners animés.

Un peu dans le désordre, je souhaite également remercier :

- Linda et Fabrice pour avoir été présents dans les moments compliqués.
- Christian Hulen pour m'avoir confié ses étudiants durant les TP de Biologie Moléculaire et Microbiologie à Cergy.
- Laurent Terradot avec qui j'ai eu le plaisir de collaborer durant cette thèse. J'espère vraiment que tu obtiendras enfin de beaux cristaux.
- Hilde et Kerstin pour m'avoir accueilli à Pasteur le temps de quelques manips, et pour leurs conseils avisés concernant *Helicobacter*.
- Mon grand frère pour avoir corrigé ce manuscrit et m'avoir appris à accorder le participe passé

Les Thésards anonymes pour nos soirées scientifiques et ludiques mémorables qui auront su faire rimer chips, bière et science (et ce n'est pas facile).

Je souhaite remercier tout particulièrement deux personnes qui m'ont donné le goût de la recherche : Jean-Marie Clément et Caroline Wilde. Vous avez su me donner ma chance alors que j'étais loin d'être le meilleur candidat.

Enfin un grand merci à mes deux amours : ma princesse pour ton amour, ton soutien et surtout pour ta patience ; et Sophie qui restera ma plus belle publication de thèse.

## TABLE DES MATIERES

LISTE DES ABBREVIATIONS .....	5
TABLE DES ILLUSTRATIONS.....	6
AVANT PROPOS.....	8
INTRODUCTION .....	11
Chapitre I. Le métabolisme des précurseurs des acides nucléiques.....	12
I. Les acides nucléiques .....	12
A. Un rapide historique de la découverte de l'ADN .....	12
B. La nature des constituants des acides nucléiques .....	14
II. Le métabolisme des bases pyrimidiques.....	15
A. La synthèse de novo.....	17
a. Biosynthèse des ribonucléotides UTP et CTP.....	17
b. Biosynthèse du dCTP.....	17
c. Biosynthèse du dTTP: les thymidylate synthases.....	18
(i) Caractéristiques biochimiques.....	18
1. La thymidylate synthase « classique », ThyA.....	18
2. La thymidylate synthase flavine dépendante, ThyX .....	21
(ii) Thymidylate synthases et régulation.....	24
a. Autorégulation traductionnelle de ThyA .....	24
b. ThyX est-elle aussi son propre régulateur ?.....	25
(iii) Importance biomédicale des thymidylate synthases .....	26
a. Inhibiteurs directs et indirects de ThyA .....	26
b. ThyX, une nouvelle cible antibactérienne.....	27
c. Le phénomène de « thymineless death » .....	29
B. Les voies de récupération et de recyclage des pyrimidines.....	32
C. Le métabolisme des pyrimidines chez <i>Helicobacter pylori</i> .....	34
a. Pourquoi <i>Helicobacter pylori</i> ?.....	34

b. Voies de novo et de récupération des pyrimidines .....	35
c. Les enzymes du métabolisme des pyrimidines chez <i>H. pylori</i> , de bonnes cibles antimicrobiennes ? .....	36
Chapitre II. Les génomes procaryotes.....	38
I. Structure.....	39
II. La composition en bases .....	40
III. La taille des génomes procaryotes .....	43
Problématique de la thèse .....	47
<b>RESULTATS ET DISCUSSION .....</b>	<b>49</b>
Chapitre I. Thymidylate synthases et métabolisme des folates .....	50
I. Introduction.....	50
A. Rôle clé du tétrahydrofolate et de ses dérivés .....	50
B. Problématique.....	51
II. Article I : “Flavin-dependent thymidylate synthase ThyX activity : Implications for the folate cycle in Bacteria” .....	54
III. Discussion .....	69
A. ThyX est une thymidylate synthase in vivo .....	69
B. ThyA et Fola sont fonctionnellement couplées .....	70
C. Les cas atypiques.....	70
a. Absence de Fola en présence de ThyA .....	71
b. Présence de ThyX, ThyA et Fola.....	72
c. Présence de ThyX et Fola .....	73
Chapitre II. Thymidylate synthases et dynamique des génomes procaryotes .....	74
I. Problématique.....	74
II. Article II : « Flavin Dependent Thymidylate Synthase X Limits Chromosomal DNA Replication » .....	76
III. Discussion .....	88
A. La thymidylate synthase ThyX limite la vitesse de réplication d’ <i>Escherichia coli</i> , un organisme utilisant naturellement ThyA.....	88
B. Les organismes procaryotes utilisant ThyX répliquent lentement leur ADN et ont une croissance lente.....	89

C.	La thymidylate synthase ThyA est nécessaire à la réplication de grands génomes	90
D.	La cinétique de synthèse du thymidylate limite l'expansion des génomes procaryotes	91
Chapitre III. Etude de deux étapes clés du métabolisme des pyrimidines chez <i>H. pylori</i> . 93		
I.	La synthèse du dTMP.....	93
A.	Introduction .....	93
B.	Matériels et Méthodes.....	94
a.	Souches bactériennes et conditions de croissances .....	94
b.	Techniques moléculaires.....	95
c.	« Tandem Affinity Purification » de protéines interagissant avec ThyX .....	96
C.	Résultats et Discussion.....	96
a.	Tentative d'inactivation conditionnelle du gène thyX .....	96
b.	Recherche de protéines interagissant avec la thymidylate synthase X.....	99
1.	« Tandem Affinity Purification ».....	99
2.	« Pull-down ».....	103
II.	La synthèse de l'UMP .....	104
A.	Introduction .....	104
B.	Matériels et Méthodes.....	105
a.	Souches bactériennes et conditions de croissance.....	105
b.	Techniques moléculaires.....	106
c.	Séquences et recherche de motifs.....	107
d.	Purification de Hp-OPRTase et mesure de l'activité enzymatique .....	108
C.	Résultats.....	109
a.	<i>Helicobacter pylori</i> est sensible à de fortes concentrations de 5-fluorouracile	109
b.	Le mécanisme de transport de l'uracile doit être passif et non facilité .....	110
c.	Recherche bioinformatique de protéines contenant un motif « PRTase »....	112
d.	Hp-OPRTase compense l'absence d'UPRTase chez <i>E. coli</i> .....	112
e.	L'expression de Hp-OPRTase induit une sensibilité au 5FU chez <i>E. coli</i> .....	113

f. L'activité secondaire UPRTase de HP-OPRTase n'est pas suffisante pour permettre la croissance de <i>H. pylori</i> en absence de voie de novo .....	114
g. Test de la spécificité de substrat de Hp-OPRTase in vitro.....	115
h. Cristallogenèse de Hp-OPRTase .....	116
D. Discussion.....	118
<b>DISCUSSION GENERALE ET PERSPECTIVES .....</b>	<b>123</b>
Thymidylate synthases.....	124
I. Distribution des deux thymidylate synthases dans le monde procaryote.....	124
II. Simplement des corrélations ?.....	126
III. Thymidylate synthase et eucaryotes ?.....	128
Le métabolisme des pyrimidines chez <i>H. pylori</i> : l'orotate phosphoribosyltransférase....	130
I. La thymidylate synthase ThyX : une protéine essentielle et solitaire ? .....	130
II. Un double rôle pour les OPRTases bactériennes ?.....	131
III. Quelques perspectives.....	131
A. génétiques.....	131
B. biochimiques .....	132
<b>REFERENCES BIBLIOGRAPHIQUES.....</b>	<b>133</b>

## LISTE DES ABBREVIATIONS

A :	Adénosine
Å :	Angström
Ade :	Adénine
ADN :	Acide désoxyribonucléique
ARN :	Acide ribonucléique
C :	Cytidine
CH <sub>2</sub> H <sub>4</sub> folate :	Méthylène tétrahydrofolate
Cyt :	Cytosine
d :	Désoxy-
DHF :	Dihydrofolate
DHFR :	Dihydrofolate réductase
DP :	Diphosphate
dUMP :	désoxyuridine 5'-monophosphate
dTMP :	désoxythymidine 5'-monophosphate ou thymidylate
5FU :	5-fluorouracile
5-FdUMP :	5-fluorodésoxyuridine 5'-monophosphate
FAD :	Flavine adénine dinucléotide
G :	Guanosine
Gua :	Guanine
H <sub>2</sub> folate :	Dihydrofolate
H <sub>4</sub> folate :	Tétrahydrofolate
HPLC :	Chromatographie liquide à haute pression
kb :	k=Kilo base
kDa :	Kilo dalton
Mb :	Mega base
MP :	Monophosphate
MTHF :	Méthylène tétrahydrofolate
NADH :	Nicotinamide-adénine-dinucléotide
NADPH :	Nicotinamide-adénine-dinucléotide-phosphate
ORF :	Open reading frame, ou phase ouverte de lecture
pb :	Paires de bases
PRPP :	5-phosphoribosyl-1-pyrophosphate
PRTase :	Phosphoribosyltransférase
T :	Thymidine ou ribothymidine
Thy :	Thymine
<i>thyA</i> :	Gène codant la thymidylate synthase A
ThyA :	Thymidylate synthase A
<i>thyX</i> :	Gène codant la thymidylate synthase X
ThyX :	Thymidylate synthase X
TP :	Triphosphate
UMP :	Uridine monophosphate
Ura :	Uracile
Uri :	Uridine



## TABLE DES ILLUSTRATIONS

Figure 1 : Voies métaboliques de biosynthèse et de récupération des pyrimidines chez les procaryotes. ....	16
Figure 2 : Structure tridimensionnelle de la thymidylate synthase ThyA d' <i>E. coli</i> .....	19
Figure 3 : Mécanisme réactionnel de la thymidylate synthase ThyA.....	20
Figure 4 : Structure tridimensionnelle de la thymidylate synthase ThyX de PBCV-1.....	21
Figure 5 : Mécanisme réactionnel de la thymidylate synthase ThyX.....	23
Figure 6 : Réaction catalysée par l'UPRTase. ....	33
Figure 7 : Relation entre la composition relative en GC (GC%) et la taille du génome.....	42
Figure 8 : Distribution de la taille des génomes en fonction du groupe phylogénétique.....	44
Figure 9 : Relation entre le nombre de gènes et la taille du génome chez les procaryotes .....	45
Figure 10 : Structure du tétrahydrofolate.....	50
Figure 11 : Modification de la morphologie de <i>R. capsulatus</i> en limitation en dTMP.....	69
Figure 12 : Distribution de la DHFR Fola et des thymidylate synthases ThyA et ThyX chez les bactéries .....	71
Figure 13 : Culture d' <i>E. coli</i> BL21 ( <i>thyA</i> <sup>+</sup> , <i>folA</i> <sup>+</sup> ) portant (courbes violettes) ou non (lignes bleues) le plasmide pVEX ( <i>thyX</i> <sup>+</sup> ) .....	72
Figure 14 : Distributions phylogénétiques de <i>thyX</i> et <i>thyA</i> dans la plupart des groupes procaryotes. ....	75
Figure 15 : Effet de l'induction de <i>thyX</i> sur la croissance de la souche d' <i>E. coli</i> FE013 ( $\Delta$ <i>thyA</i> ) .....	88
Figure 16 : Deux scénarii possibles expliquant la corrélation positive entre la taille du génome et la quantité de bases. ....	91
Figure 17 : Différentes étapes de construction du fragment de délétion du gène <i>thyX</i> . ....	98
Figure 18 : Représentation schématique des différentes étapes de la purification de complexe protéique par la technique du TAP-tag.....	100
Figure 19 : Différentes étapes de construction du plasmide portant la fusion du gène <i>thyX</i> et de l'étiquette TAP.....	101

Figure 20 : Immuno-blot sur des extraits cellulaires des souches 26695 et FE008 .....	102
Figure 21 : Migration sur gel SDS-PAGE de l'éluat protéique issu de la purification TAP .....	103
Figure 22 : Structure chimique de l'uracile, du 5-fluorouracile et de l'orotate. ....	109
Figure 23 : La souche <i>H. pylori</i> 26695 cultivée en absence ou en présence de 5-fluorouracile .....	110
Figure 24 : Schéma de la délétion du gène <i>hp1175</i> . ....	111
Figure 25 : A- Analyse SDS-PAGE de la protéine recombinante Hp-OPRTase. B- Cinétique de la catalyse de Hp-OPRTase pour son substrat naturel : l'orotate. ....	115
Figure 26 : Cristaux de Hp-OPRTase .....	117
Figure 27 : Cristaux de Hp-OPRTase. ....	118
Figure 28 : Conséquence de l'expansion du génome sur le temps de réplication. ....	127
Tableau 1 : Exemples de bactéries pathogènes possédant le gène <i>thyX</i> et maladies associées .....	28
Tableau 2 : Exemple de bactéries dont des mutants auxotrophes pour la thymidine ont été décrits	29
Tableau 3 : Modulation de la TLD par mutation dans des gènes de la réparation de l'ADN .....	31
Tableau 4 : Exemples de bactéries possédant plusieurs réplicons .....	40
Tableau 5 : Groupes mono-carbonés portés par le THF .....	51
Tableau 6 : Listes des souches et plasmides utilisées dans ce travail. ....	95
Tableau 7 : Listes des amorces utilisées dans ce travail. ....	96
Tableau 8 : Liste des plasmides et souches bactériennes utilisées. ....	106
Tableau 9 : Liste des amorces utilisées. ....	107
Tableau 10 : Phénotypes de croissance des différentes souches d' <i>E. coli</i> utilisées en réponse à 20 $\mu\text{g.ml}^{-1}$ de 5FU. ....	113

## AVANT PROPOS

Le thème de recherche concernant la thymidylate synthase ThyX a été initié dans les deux laboratoires au sein desquels j'ai effectué les travaux présentés dans cette thèse. La découverte conjointe, par l'équipe d'Hannu Myllykallio de l'Institut de Génétique et Microbiologie à Orsay et l'équipe d'Ursula Liebl du Laboratoire d'Optique et Biosciences (LOB) à l'Ecole Polytechnique, de cette nouvelle famille d'enzymes en 2002 a ouvert le champ à de nombreuses études biochimiques couplées à l'étude des structures tridimensionnelles. Ces travaux, ainsi que ceux d'autres équipes à travers le monde, ont permis de caractériser le site catalytique, les sites de fixation des différents substrats et d'élaborer un modèle du mécanisme réactionnel.

A mon arrivée en 2004, c'est un autre aspect de la thématique ThyX qui m'a été proposé. Les analyses génomiques de la distribution des thymidylate synthases chez les procaryotes avaient mis en évidence que les organismes microaerophiles utilisent préférentiellement la thymidylate synthase ThyX pour répliquer leur génome. Les premières études biochimiques de ThyX avaient elles montré que son activité est inhibée par l'oxygène. La mise en parallèle de ces deux observations semblait indiquer que l'oxygène avait un rôle important dans l'activité de cette enzyme. D'une façon plus large, mon but était d'étudier ThyX *in vivo* chez les bactéries. Pour ce faire, il fallait un modèle expérimental et l'organisme retenu pour étudier cette question a été la bactérie pathogène humaine *Helicobacter pylori*. Différentes souches de *Helicobacter pylori* ainsi que les outils génétiques permettant de les manipuler nous ont été aimablement fournis par Hilde de Reuse de l'Unité de Pathogénie Bactérienne des Muqueuses (Institut Pasteur). Ma première tâche a été de mettre en place les conditions nécessaires à la culture de la bactérie au LOB, grâce aux conseils de Hilde de Reuse et à l'expertise de Stéphane Skouloubris, alors chercheur post-doctorant dans l'équipe d'Hannu Myllykallio. Cette thèse devait être originellement une étude *in vivo* de ThyX chez *H. pylori*, par des approches de biologie moléculaire et de génétique. Trois années plus tard, c'est une étude

bien plus globale des conséquences de l'utilisation de ThyX sur la physiologie des procaryotes que j'ai le plaisir de vous présenter.

J'ai organisé la présentation de cette thèse sous la forme de trois chapitres correspondants à trois projets que j'ai menés durant cette thèse. Le premier, initialement débuté par Damien Leduc lorsqu'il était doctorant sous la direction de Hannu Myllykallio, traite de l'influence de la thymidylate synthase X sur le métabolisme des folates. Le deuxième projet est une étude des conséquences des différences d'activités catalytiques des deux familles de thymidylate synthases sur la réplication de l'ADN et plus généralement sur la dynamique des génomes. Le dernier chapitre reporte un travail qui a pour but de mieux comprendre deux des étapes du métabolisme des pyrimidines chez *H. pylori*

Ainsi, les différents travaux réalisés durant ma thèse, et présentés dans ce manuscrit, ont pour point commun le métabolisme des bases pyrimidiques chez les organismes procaryotes. En conséquence, je vais débiter ce manuscrit par une introduction présentant les voies de biosynthèse et de récupération des pyrimidines, en détaillant plus particulièrement la synthèse du thymidylate. Afin de permettre une compréhension des travaux présentés dans le chapitre II, je compléterai cette introduction par une présentation des connaissances actuelles concernant la nature et l'évolution des génomes procaryotes. Enfin, j'exposerai la problématique de ma thèse.



## **INTRODUCTION**

## Chapitre I. Le métabolisme des précurseurs des acides nucléiques

### I. Les acides nucléiques

#### A. *Un rapide historique de la découverte de l'ADN*

L'histoire de l'étude des acides nucléiques débuta durant la seconde partie du XIX<sup>ème</sup> siècle en Suisse. C'est en 1869 que Friedrich Miescher isola pour la première fois, à partir de cellules de pus, une substance contenant du phosphate, qu'il nomma alors nucléine et que Richard Altmann renommera acide nucléique en 1889 [pour revue (Dahm, 2005)]. Près de soixante ans après la première purification d'ADN, Phoebus Levene identifia les briques constitutives de l'ADN : les quatre bases adénine (A), cytosine (C), guanine (G), et thymine (T).

Le rôle de cette macromolécule était alors inconnu. Le lien entre l'ADN et l'hérédité a été démontré pour la première fois en 1944 grâce aux travaux d'Oswald T. Avery, Colin McLeod et Maclyn McCarty (Avery *et coll.*, 1944). Ces travaux étaient basés sur ceux de Frederick Griffith qui avait montré que des souches non virulentes de pneumocoques mises en contact avec des pneumocoques virulents inactivés par la chaleur pouvaient être « transformées » en bactéries virulentes (Griffith, 1928). Il avait postulé qu'un « principe transformant » permettait le transfert de propriétés d'un type de bactérie à un autre. Avery et collaborateurs purifièrent le principe actif qui se révéla être de l'acide désoxyribonucléique. Le lien entre l'ADN et la génétique venait d'être trouvé. Ce lien sera confirmé en 1952 par les travaux d'Alfred Hershey et Martha Chase qui ont montré que seul l'ADN du bactériophage T2 pénétrait dans la cellule et non les protéines virales (Chase et Hershey, 1952).

Au début des années 1950, Erwin Chargaff montra que la composition en bases de l'ADN variait entre les espèces, mais que les bases de l'ADN au sein d'une espèce sont présentes à un taux fixe : il y a autant de A que de T, et autant de G que de C (Chargaff, 1950). Cette information fut cruciale à James Watson et Francis Crick pour proposer une structure moléculaire de l'ADN en 1953

(Watson et Crick, 1953). En se basant sur les clichés aux rayons X de Rosalind Franklin et Maurice Wilkins, ils proposèrent que l'ADN devait se présenter sous la forme d'une double hélice dans laquelle A s'apparie toujours avec T, et C toujours avec G. Cette structure même de l'ADN permit aux auteurs de proposer que « [...] l'appariement spécifique des bases [...] suggère immédiatement un possible mécanisme du recopiage du matériel génétique ». La réplication semi conservative ne sera démontrée que cinq ans plus tard par Matthew Meselson et Franklin Stahl (Meselson et Stahl, 1958). Francis Crick proposa dès 1958 que l'ADN agirait comme support moléculaire d'une information qui s'exprime à travers les protéines, ceci en accord avec la démonstration de George Beadle et Edward Tatum que chaque gène est responsable de la production d'une enzyme (Beadle et Tatum, 1941). En 1961, les travaux de François Jacob et Jacques Monod montrèrent que l'intermédiaire entre ADN et protéine était un acide ribonucléique, l'ARN messager (ARNm) (Jacob et Monod, 1961). Les ARN de transfert, participant à la traduction du message porté par les ARNm en protéines, furent eux identifiés pour la première fois en 1958 par Hoagland (Hoagland *et coll.*, 1958). Durant la deuxième partie des années 60, le code génétique fut déchiffré par Har G. Khorana, Marshall W. Nirenberg et leurs collaborateurs (Khorana, 1965; Nirenberg *et coll.*, 1965) rendant concrète l'intuition de Francis Crick proposant que trois nucléotides doivent correspondre à un acide aminé d'une protéine.

En 1977, Frederick Sanger tourna une nouvelle page dans l'histoire de la biologie moderne en mettant au point une méthode de séquençage de l'ADN (Sanger *et coll.*, 1977). La première séquence complète du génome d'un microorganisme procaryote, la bactérie *Haemophilus influenzae*, fut publiée en 1995 (Fleischmann *et coll.*, 1995). L'année suivante, il en fut de même pour le premier eucaryote, la levure *Saccharomyces cerevisiae* l'année suivante (Goffeau *et coll.*, 1996). Les deux projets de séquençage du génome humain se verront publiés en 2001 (Lander *et coll.*, 2001; Venter *et coll.*, 2001). A ce jour, le National Center for Biotechnology Information (NCBI) indique que 567 génomes procaryotes sont complètement séquencés et que 841 sont en cours. Ces dernières années, une quantité sans précédent d'informations a ainsi été rendue disponible permettant l'essor de nouvelles disciplines telles que la génomique et la post-génomique.



Rétrospectivement, il est intéressant de noter que la simplicité de la molécule d'ADN, comparée à la composition et à la structure complexe des protéines, a longtemps masqué son rôle central dans la cellule. L'ADN a deux grandes fonctions dans le support de l'hérédité. Il permet de stocker l'information déterminant le développement et le fonctionnement de l'organisme dans son environnement, et de transmettre cette information à la descendance en se répliquant. Cette information est stockée sous forme de mots composés de lettres provenant d'un alphabet réduit composé de quatre lettres.

### ***B. La nature des constituants des acides nucléiques***

Quelle est donc la nature chimique de ces « lettres » ? Les acides nucléiques, ARN ou ADN, sont des macromolécules composées d'une succession de nucléotides, et pourront être qualifiés de polymères de nucléotides ou encore de polynucléotides. Un nucléotide résulte de la liaison d'un phosphate sur le pentose d'un nucléoside. Le nucléoside résulte de la liaison N-glycosidique d'un ose à une base azotée (ou nucléobase). Il existe cinq nucléobases différentes classées en deux familles : les pyrimidines comprenant la thymine (Thy), l'uracile (Ura) et la cytosine (Cyt), et les purines que sont l'adénine (Ade) et la guanine (Gua). La nature du pentose détermine la nature du nucléoside et le type de polynucléotides qu'il intégrera (ARN pour ribose, et ADN pour désoxyribose). S'il s'agit d'un ribose les nucléosides résultant seront l'adénosine (A), la guanosine (G), la ribothymidine (T), l'uridine (U) et la cytidine (C); sinon la désoxyadénosine (dA), désoxyguanosine (dG), la désoxythymidine (dT), la désoxyuridine (dU) et la désoxycytidine (dC) dans le cas d'un désoxyribose.

La liaison, par estérification en 5'-OH, d'un acide phosphorique mono-, di- ou triphosphate conduit à la formation des nucléotides. C'est sous la forme triphosphate que les nucléotides peuvent être utilisés pour synthétiser les molécules d'ADN et d'ARN. Ces molécules sont constituées d'un assemblage de nucléotides reliés respectivement par un groupement phosphate en 5' et un groupement OH en 3'.

Les acides nucléiques, ADN ou ARN, ne contiennent chacun que quatre types de nucléobases. Les deux bases à noyaux puriques sont communes à l'ADN et à l'ARN, mais seule la base pyrimidique cytosine est présente dans les deux polynucléotides. La quatrième nucléobase pyrimidique est l'uracile pour l'ARN et son analogue méthylé, la thymine, pour l'ADN.

## II. Le métabolisme des bases pyrimidiques

Il existe deux types de stratégies pour synthétiser un nucléotide donné. Nous distinguerons la synthèse dite *de novo*, des voies dites de récupération (ou de recyclage). La synthèse *de novo* représente un coût énergétique important pour la cellule, ainsi de nombreux organismes disposent de voies alternatives permettant, soit de recycler les groupes puriques ou pyrimidiques issus de la dégradation des acides nucléiques (recyclage), soit de récupérer certains précurseurs directement à partir du milieu extérieur (récupération). Compte tenu de la nature des travaux présentés dans cette thèse, nous nous focaliserons sur les voies métaboliques conduisant à la synthèse des pyrimidines. Une vision d'ensemble des voies *de novo* et de récupération sont schématisé sur la Figure 1.



## A. La synthèse *de novo*

### a. Biosynthèse des ribonucléotides UTP et CTP

La première réaction de cette voie métabolique est la formation de carbamoylphosphate à partir de glutamine, de bicarbonate et d'ATP. Cette étape est catalysée par la carbamoylphosphate synthase codée par l'opéron *pyrA* (nommé *carAB* chez *Escherichia coli*). Le carbamoylphosphate est à l'origine des voies de synthèse *de novo* des pyrimidines, mais également de l'arginine. Le groupe amine de l'aspartate est alors greffé au carbamoylphosphate par l'aspartate carbamoyltransférase (codé par *pyrB*). La dihydroorotase (*pyrC*) et la dihydroorotate oxydase (*pyrD*), en deux étapes de cyclisation et oxydation, vont conduire à la formation de l'acide orotique (ou orotate). L'étape suivante fait intervenir le 5-phosphoribosyl-1-pyrophosphate (PRPP) qui est aussi utilisé comme substrat par au moins dix enzymes dans des processus telles la synthèse des purines, de l'histidine, du tryptophane ou des cofacteurs amido-nicotiniques. La synthèse du PRPP à partir d'ATP et de ribose 5-phosphate est catalysée par la phosphoribosylpyrophosphate synthétase. Le transfert du ribose 5-phosphate du PRPP à l'orotate est catalysé par l'orotate phosphoribosyltransférase, ou OPRTase (*pyrE*). L'acide orotidylique (OMP) ainsi formé va être décarboxylé par l'OMP décarboxylase (*pyrF*) conduisant à la formation d'uridine monophosphate (UMP). Il est à noter que chez l'homme, ces deux étapes sont catalysées par une seule enzyme. Deux étapes successives de phosphorylation par l'UMP kinase (*pyrG*) et la nucléoside diphosphokinase (*pyrH*) vont conduire à la formation du premier nucléotide pyrimidique incorporable dans l'ARN, l'UTP. Le second, le CTP, sera issu de la déamination de l'UTP par la CTP synthase.

### b. Biosynthèse du dCTP

La synthèse des désoxribonucléotides débute par la conversion des ribonucléoside diphosphates en 2'-désoxyribonucléoside diphosphates par la ribonucléoside diphosphate réductase. Comme nous l'avons vu, et contrairement à l'UDP, le CDP n'a pas d'intermédiaire lors de la formation de l'UTP. Il sera donc synthétisé à partir de celui-ci par la nucléoside diphosphokinase. Le dCTP,

comme le dATP, le dGTP et le dUTP, dérive du désoxyribonucléoside diphosphate correspondant. Cette réaction est catalysée par la ribonucléoside diphosphate réductase (*ndrAB*).

### c. Biosynthèse du dTTP: les thymidylate synthases

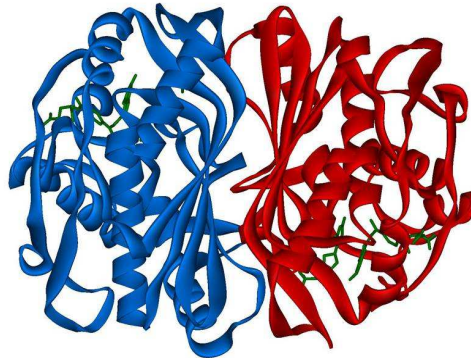
Il est intéressant de noter une exception à cette voie : les désoxyribonucléotides de la thymine n'ont pas d'équivalent ribonucléotidique libre. La ribothymidine présente dans les ARN de transfert (ARN-t) est synthétisée à partir d'une uridine incorporée dans le précurseur de l'ARN-t par une tRNA (uracil-5-)-methyltransférase. Le dTMP (ou thymidylate), précurseur du dTTP, doit donc être synthétisé à partir du désoxyuridine 5'-monophosphate (dUMP). Le dUMP peut être produit par l'hydrolyse du dUTP catalysée par la dUTPase ou par la déamination de dCMP catalysée par dCMP déaminase. La conversion du dUMP en dTMP est catalysée par la thymidylate synthase. Il existe deux familles indépendantes de thymidylate synthase, ThyA et ThyX. Ces deux enzymes ne sont pas orthologues, elles ne présentent aucune similarité de séquences, ni de structure et les mécanismes réactionnels sont différents (Myllykallio *et coll.*, 2002). Il est intéressant de noter que les organismes eucaryotes, à l'exception de *Physarum polycephalum* et *Dictyostelium discoideum* (Dynes et Firtel, 1989), possèdent seulement l'enzyme ThyA. Chez les procaryotes, la distribution phylogénétique des deux thymidylate synthases est complexe et quasi-exclusive (Leduc *et coll.*, 2004b; Myllykallio *et coll.*, 2002). Ce dernier point a été le sujet d'une partie de travaux présentés dans cette thèse et sera développé dans le chapitre II. Je présenterai d'abord une vue générale des connaissances biochimiques des deux familles enzymatiques. J'exposerai ensuite les mécanismes de régulation des thymidylate synthases. Enfin, je montrerai les applications biomédicales de l'inhibition des thymidylate synthases ainsi que les mécanismes moléculaires sous-jacents.

#### (i) Caractéristiques biochimiques

##### 1. La thymidylate synthase « classique », ThyA

La thymidylate synthase ThyA, mise en évidence pour la première fois dans les années 1950 (Friedkin et Roberts, 1956), est une enzyme homo-dimérique, formée de deux sous-unités de 30 à

35 kDa (Figure 2) qui catalyse la conversion du dUMP et méthylène tétrahydrofolate ( $\text{CH}_2\text{THF}$ ) en thymidylate (dTMP) et dihydrofolate (DHF) [revue dans (Carreras et Santi, 1995)]. Le DHF produit par ThyA est ensuite réduit en tétrahydrofolate (THF) par la dihydrofolate réductase (DHFR codée par le gène *folA*). Le THF sera ensuite recyclé en  $\text{CH}_2\text{THF}$  par la sérine hydrométhyltransférase (SHMT). Il apparaît donc qu'un cycle en trois étapes est nécessaire au recyclage du substrat folique.



**Figure 2 : Structure tridimensionnelle de la thymidylate synthase ThyA d'*E. coli* (Montfort *et coll.*, 1990).  
Chaque monomère est représenté dans une couleur différente.**

L'intérêt de ThyA en tant que cible anticancéreuse, sur lequel nous reviendrons par la suite, a dirigé les recherches concernant cette enzyme vers une dissection minutieuse de son mécanisme réactionnel. De plus, les très nombreuses structures tridimensionnelles de la protéine - plus de trente à ce jour - complexée avec ses différents substrats, ont permis de localiser précisément les résidus impliqués dans la fixation de ces derniers.

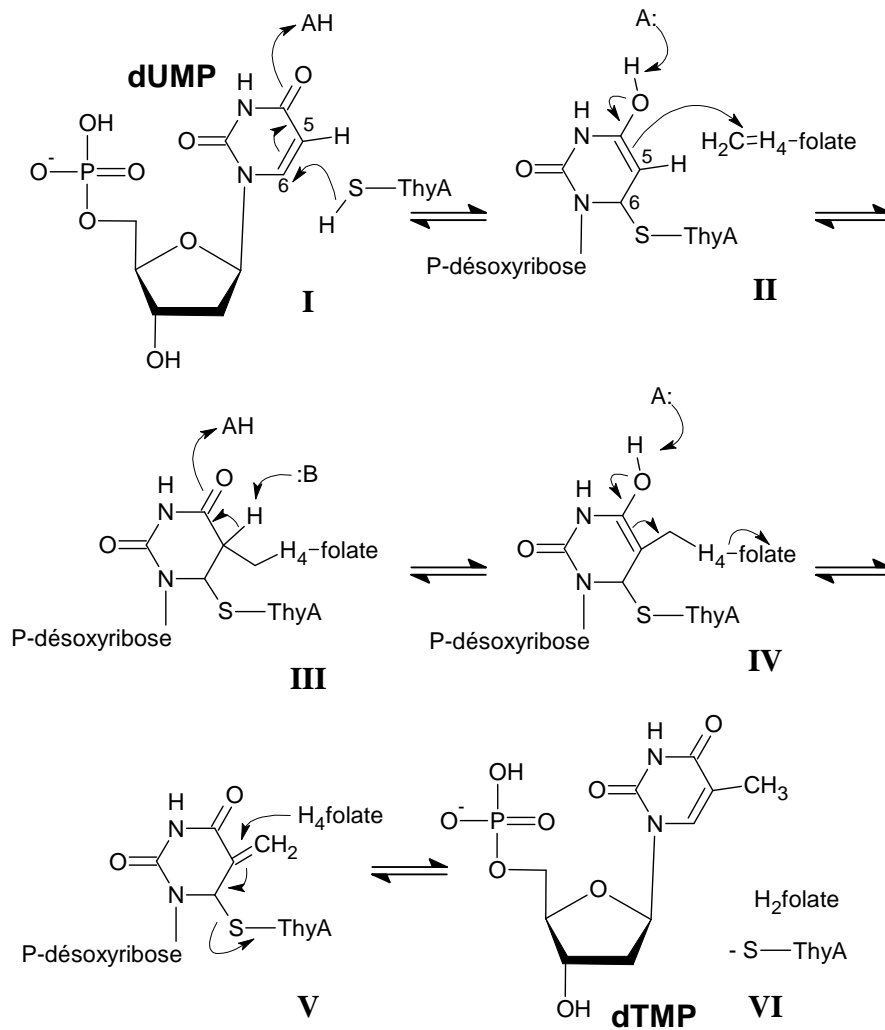


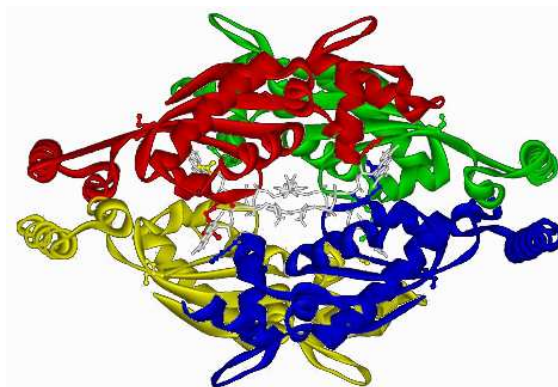
Figure 3 : Mécanisme réactionnel de la thymidylate synthase ThyA [d'après (Finer-Moore *et coll.*, 2003)]

La réaction catalysée par ThyA est de type séquentiel ordonné et peut être décomposée en six étapes. En premier lieu, l'enzyme forme un complexe ternaire réversible avec le dUMP, puis la catalyse débute par l'attaque nucléophile (étape I) du groupement thiol SH d'une cystéine de l'enzyme sur le carbone en position 6 du cycle pyrimidine du dUMP. L'intermédiaire enzyme/dUMP ainsi formé est instable. Un transfert d'électron impliquant le CH<sub>2</sub>THF se produit, conduisant à la formation d'un intermédiaire énolate. Le CH<sub>2</sub>THF est ainsi converti en ion iminium (N<sup>+</sup><sub>5</sub>) provoquant son activation (étape II). Ce dernier va se fixer par une liaison covalente au niveau du carbone 6 de l'intermédiaire énolate (étape II). L'hydrogène du carbone 5 de l'énolate exocyclique est capté par une base indéterminée (étape III) ce qui entraîne une β-élimination du THF pour former un intermédiaire méthylène exocyclique (étape IV). Le THF agit ensuite comme réducteur et permet la

réduction du groupe méthylène transféré en groupement méthyle. En résulte la coupure de la liaison covalente entre la cystéine de l'enzyme et du dTMP néoformé, provoquant la  $\beta$ -élimination de l'enzyme (étape V). Les produits de la réaction, dTMP et DHF, peuvent être libérés (étape VI). Le  $\text{CH}_2\text{THF}$  a un double rôle dans cette réaction, il est à la fois donneur de carbone et réducteur du groupement méthylène.

## 2. La thymidylate synthase flavine dépendante, ThyX

La récente découverte de la thymidylate synthase ThyX en 2002 (Myllykallio *et coll.*, 2002), explique que les données biochimiques concernant les enzymes de la famille ThyX sont relativement peu nombreuses en comparaison à celles concernant ThyA. Les protéines ThyX présentent des grandes différences avec les protéines ThyA, à plusieurs niveaux. Contrairement à ThyA qui se présente dans sa forme active sous une conformation homo-dimérique, ThyX se présente sous une configuration homo-tétramérique, formée de sous-unités de 26 à 30 kDa (Figure 4).



**Figure 4 : Structure tridimensionnelle de la thymidylate synthase ThyX de PBCV-1 (Graziani *et coll.*, 2006). Chaque monomère est représenté par une couleur différente.**

Comme ThyA, elle catalyse la conversion du dUMP en dTMP en utilisant le  $\text{CH}_2\text{THF}$  comme co-substrat, mais nécessite un troisième substrat : la forme réduite du nicotinamide adénine dinucléotide phosphate (NADPH). Une différence majeure avec la réaction catalysée par ThyA est que le produit final du dérivé folique n'est pas le DHF mais le THF. Ainsi, la réaction catalysée par ThyX n'a théoriquement pas besoin d'être couplée à la réaction catalysée par la DHFR. Une autre caractéristique importante de ThyX, qui la différencie de ThyA, est qu'elle contient un cofacteur



dérivé de la riboflavine. La flavoprotéine ThyX nécessite pour sa catalyse une flavine sous sa forme réduite, FADH<sub>2</sub> (Graziani *et coll.*, 2004). Contrairement aux protéines ThyA, les séquences protéiques de ThyX sont peu conservées et un motif spécifique a été défini : RHRX<sub>7-8</sub>S. Les récents travaux de Jon Ulmer et collaborateurs concernant ThyX de *Mycobacterium tuberculosis* ont permis d'élargir ce motif à la séquence suivante (Y/W)X<sub>19-40</sub>HX<sub>25-28</sub>RHRX<sub>7-8</sub>SXR(Y/F)X<sub>68-114</sub>R, où le résidu histidine souligné agit en tant que résidu catalytique et le résidu sérine souligné correspond au nucléophile (Ulmer *et coll.*, 2008). Les études biochimiques par mutagenèse dirigée de ThyX de *Helicobacter pylori*, couplé à l'analyse des données cristallographiques de ThyX de *Thermotoga maritima* (Kuhn *et coll.*, 2002), ont permis de mettre en évidence que chaque site actif est localisé à l'interface de 3 sous unités (Leduc *et coll.*, 2004a). Cette caractéristique est rare et, à ce jour, n'a été décrite que pour la famille des adénylosuccinate lyases (Brosius et Colman, 2002). Les détails du mécanisme catalytique de ThyX sont sujets à discussion, notamment l'ordre de fixation des substrats (Agrawal *et coll.*, 2004; Chernyshev *et coll.*, 2007; Graziani *et coll.*, 2004; Graziani *et coll.*, 2006; Griffin *et coll.*, 2005; Ulmer *et coll.*, 2008). Par souci de clarté, je ne présenterai qu'un seul schéma réactionnel. De façon arbitraire, j'exposerai ici le mécanisme proposé au laboratoire par Graziani et collaborateurs concernant ThyX du *Paramecium bursaria* chlorella virus (PBCV-1), virus à ADN double brin infectant une algue unicellulaire, *Chlorella* NC64A (Graziani *et coll.*, 2004; Graziani *et coll.*, 2006).

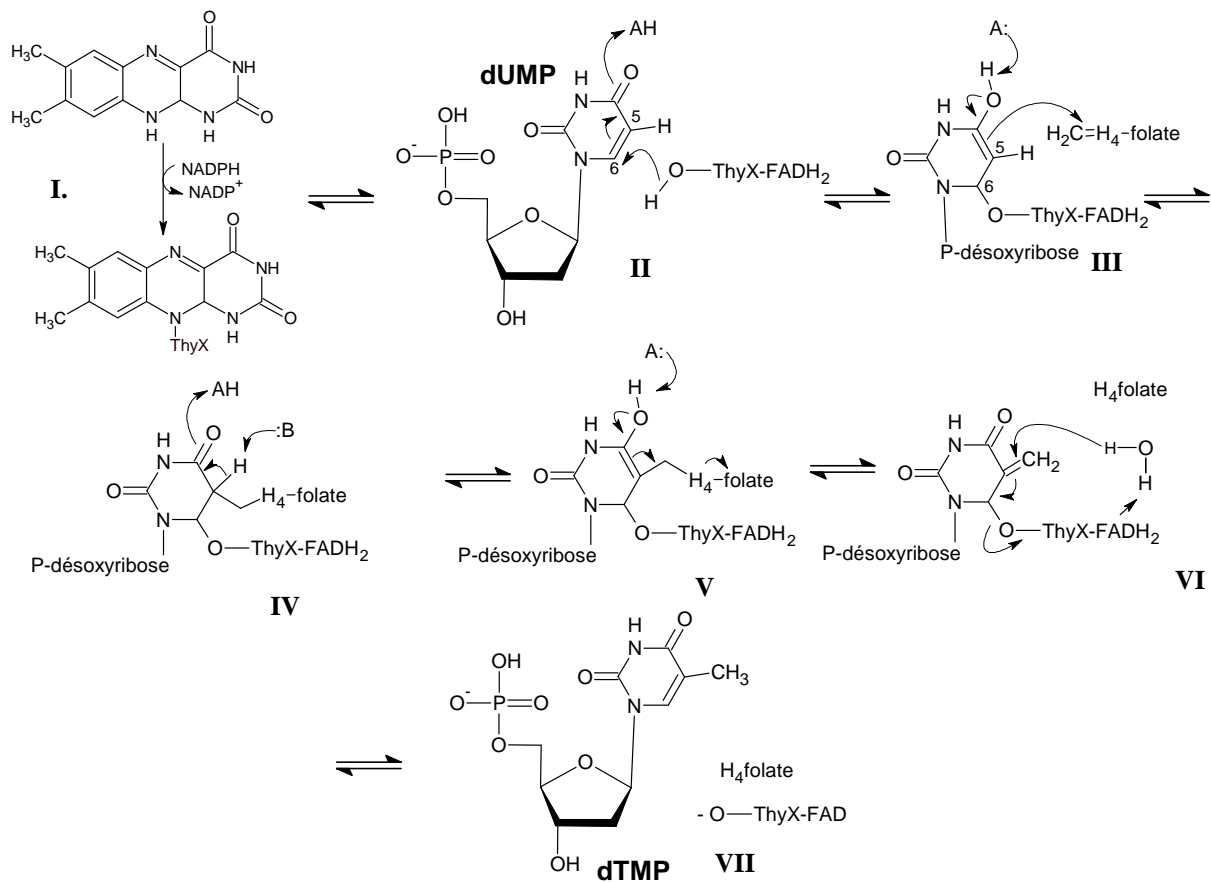


Figure 5 : Mécanisme réactionnel de la thymidylate synthase ThyX [d'après (Graziani *et coll.*, 2006)].

La réaction catalysée par ThyX est de type séquentiel et peut être décomposée en sept étapes (Figure 5). La réaction débute par l'oxydation du NADPH qui va réduire le cofacteur FAD en  $\text{FADH}_2$  (étape I). L'attaque nucléophile du carbone  $\text{C}_6$  du noyau pyrimidine du dUMP est ensuite catalysée par une sérine, contrairement à ThyA où le résidu responsable est une cystéine (étape II). Le  $\text{CH}_2\text{THF}$  agit ensuite comme donneur de groupement méthylène permettant la méthylation du dUMP conduisant à la formation d'un intermédiaire  $\text{CH}_2\text{-dUMP}$  (étape III). Les étapes III, IV et V sont supposées identiques à celles de la réaction catalysée par ThyA incluant la déprotonation du  $\text{H}^+$  en position 5 du noyau pyrimidine du dUMP. La réduction du groupement méthylène transféré sera réalisée par le  $\text{FADH}_2$  et non par le THF comme pour ThyA (étape VI) (Agrawal *et coll.*, 2004; Graziani *et coll.*, 2004). Les produits de la réaction, THF et dTMP, sont enfin libérés de l'enzyme (étape VII).

## (ii) Thymidylate synthases et régulation

### a. Autorégulation traductionnelle de ThyA

Chez les mammifères, la thymidylate synthase est régulée à différents niveaux et de façon complexe, elle est notamment capable de fixer son propre ARN messenger (ARNm) et d'agir ainsi comme un auto-répresseur traductionnel (Liu *et coll.*, 2002). Elle serait également impliquée dans la régulation traductionnelle négative d'autres processus cellulaires en interagissant avec les ARNm du gène codant le suppresseur de tumeur p53 ou les facteurs de transcription de la famille *myc* (Ju *et coll.*, 1999). Une autre régulation négative de la thymidylate synthase a été attribuée à l'expression d'un ARN anti-sens naturel (rTS $\alpha$ ) (Chu et Dolnick, 2002). Les auteurs ont pu provoquer un clivage spécifique de l'ARNm de ThyA en induisant la production d'ARN rTS $\alpha$  dans des cellules humaines.

L'expression de la thymidylate synthase chez l'Homme peut, à l'inverse, être régulée positivement. L'augmentation du niveau d'expression de l'ARNm et de l'enzyme lors de la transition de la phase G1 à la phase S (synthèse de l'ADN) a été récemment montré comme étant dépendante du facteur de transcription LSF (Powell *et coll.*, 2000). Les auteurs ont montré que la fixation spécifique de LSF sur des sites au niveau du promoteur et de régions introniques était indispensable à l'induction de la thymidylate synthase. L'utilisation d'un allèle muté de LSF négatif dominant a permis de montrer qu'en absence de régulation positive par LSF, le niveau basal de thymidylate synthase n'est pas suffisant pour permettre une correcte répllication de l'ADN et conduit à un phénomène d'apoptose.

Les régulations présentées ci-dessus correspondent à des variations physiologiques du niveau de thymidylate synthase ; mais les cellules de mammifères peuvent également adapter ce niveau en réponse à des drogues ciblant ThyA ou des dommages à l'ADN. Il est clairement établi, qu'en réponse à des traitements au 5 fluorouracile (mécanismes d'action présentés ci-après), certaines cellules induisent initialement une diminution de niveau de thymidylate synthase, suivi d'une induction [pour revue (Peters *et coll.*, 2002)]. Une autre preuve de régulation de l'activité

thymidylate synthase a été présenté par Batra et collaborateurs (Batra *et coll.*, 2004) lors d'expériences d'irradiation de souris aux rayons gamma. Lors d'une irradiation aux rayons gamma, des dommages importants sont subis par l'ADN. Des mécanismes variés de réparation de l'ADN sont alors mise en place et vont nécessiter un apport supplémentaire en nucléobases et folates. Les auteurs ont montré que durant les 96 heures suivant une telle irradiation, on observe une augmentation significative, dose dépendante, des activités thymidylate synthase et dihydrofolate réductase. Ces deux exemples montrent que la cellule eucaryote est capable de réguler sa capacité à synthétiser du thymidylate en fonction de ses besoins.

Alors que le nombre de travaux chez les mammifères est relativement important, il est surprenant de noter la quasi-inexistence de travaux concernant la régulation de ThyA chez les bactéries. L'interrogation des bases de données regroupant les résultats de diverses études transcriptomiques, M3D et PRODORIC (Faith *et coll.*, 2008; Munch *et coll.*, 2003), ne permettent pas de mettre en évidence de régulation selon les phases de croissance comme chez les eucaryotes. On notera seulement que chez *E. coli*, il a été proposé qu'il existerait un couplage traductionnel entre la phosphatidylglycerol:prolipoprotein diacylglycerol transférase (*lgt*) et la thymidylate synthase (*thyA*) lié au chevauchement entre la phase ouverte de lecture du gène *lgt* et le site de fixation du ribosome du gène *thyA* (Gan *et coll.*, 1995). De plus, une étude bioinformatique systématique semble indiquer que les gènes *thyA* de certains organismes (comme *Listeria monocytogenes* ou *Salmonella typhi*) pourraient être régulés par atténuation transcriptionnelle (Merino et Yanofsky, 2005). Ces prédictions restent à être validés expérimentalement.

#### b. ThyX est-elle aussi son propre régulateur ?

La littérature concernant la régulation de la thymidylate synthase ThyX, comme celle de son analogue bactérien ThyA, est quasi-inexistante. Le seul article publié à ce jour montre que la protéine ThyX de l'archée *Pyrococcus furiosus* peut se lier à l'ARN *in vitro*, et plus particulièrement à la structure tige-boucle présente au point de démarrage de la traduction de son propre ARNm (Kanai *et*

*coll.*, 2006). Les auteurs concluent que la capacité à fixer l'ARN pourrait être une propriété commune aux deux familles de thymidylate synthases. Cette conclusion peut sembler hâtive pour plusieurs raisons : 1) cette propriété n'a été décrite que pour les thymidylate synthases ThyA eucaryotes et non bactériennes ; 2) le travail présenté a été réalisé exclusivement *in vitro* ; 3) la structure tige-boucle décrite n'est pas retrouvée pour les gènes *thyX* des autres organismes (données non montrées). Comme pour les gènes *thyA*, il est prédit que les gènes *thyX* de certains organismes comme *Chlamydia trachomatis* ou encore *Geobacter sulfurreducens* pourraient être régulés par atténuation transcriptionnelle.

### (iii) Importance biomédicale des thymidylate synthases

Le domaine biomédical porte un intérêt particulier au métabolisme des pyrimidines, en particulier aux thymidylate synthases. Les cellules qui prolifèrent rapidement nécessitent une synthèse *de novo* importante des désoxyribonucléotides comparativement aux cellules qui prolifèrent plus lentement. Il apparaît alors que l'inhibition d'une étape clé de ce métabolisme, telle que celle de la synthèse du thymidylate, va entraîner une altération du cycle cellulaire et de la division cellulaire conduisant à la mort de la cellule. Deux domaines d'application biomédicale tirent partie de ce phénomène : la recherche de molécules anticancéreuses et celle de composés antibactériens.

#### a. Inhibiteurs directs et indirects de ThyA

Parmi les composés utilisés en chimiothérapie anticancéreuse, on retrouvera le 5-fluorouracile (5FU), dérivé des pyrimidines. Le 5FU est utilisé pour le traitement d'un nombre important de tumeurs solides depuis les années 1950 (Heidelberger *et coll.*, 1957). Le 5FU agit à différents niveaux, principalement par l'inhibition de la thymidylate synthase, et également par l'incorporation de ses métabolites dans l'ARN et de l'ADN. Ce n'est pas le 5FU qui inhibe directement la thymidylate synthase mais son métabolite, le 5-fluoro-2'-deoxyuridine 5'-monophosphate (5FdUMP). Le 5FdUMP va se fixer sur le même site et avec la même affinité que le dUMP, mais

contrairement à l'hydrogène, l'atome de fluorine en position 5' ne sera pas déplacé. Ainsi le 5FdUMP et le folate réduit vont être lié de façon covalente à l'enzyme, inhibant la réaction. Le 5FU est un inhibiteur indirect de la thymidylate synthase dont l'efficacité va dépendre des voies de récupération de son analogue naturel, l'uracile (détaillées dans le paragraphe traitant des voies de récupération et de recyclage des pyrimidines). Récemment des inhibiteurs directs de la thymidylate synthase ont été développé tel que le FdUMP[10], un décamère de 5FdUMP (Liao *et coll.*, 2005).

Une autre façon indirecte d'inhiber la réaction catalysée par ThyA est d'inhiber la DHFR. La synthèse de dTMP par ThyA est couplée à la conversion de MTHF en DHF, et ce folate réduit doit être recyclé en THF par la DHFR. En bloquant cette étape de recyclage, le folate va se retrouver « piégé » sous la forme de DHF (« DHF trap »), ce qui aura pour conséquence de progressivement faire disparaître le stock de MTHF disponible. La réaction catalysée par la ThyA sera alors inhibée par l'absence d'un de ses substrats. Les composés inhibiteurs de la DHFR font partis de la famille des anti-folates ; on notera en particulier le méthotrexate utilisé en chimiothérapie et le triméthoprime en antibiothérapie [pour revue (Kompis *et coll.*, 2005)].

#### b. ThyX, une nouvelle cible antibactérienne

La récente découverte d'une nouvelle famille de thymidylate synthases, ThyX, a éveillé un intérêt certain concernant l'origine de l'ADN, mais a également permis de mettre en lumière une cible antibactérienne inédite (Myllykallio *et coll.*, 2002). Pourquoi ThyX est-elle si intéressante ? Nous pouvons résumer la réponse en 3 points : 1) la synthèse du dTMP chez l'Homme est indépendante de ThyX ; 2) ThyA et ThyX ne partagent aucune homologie de séquence et de structure tridimensionnelle ; 3) ThyX possède un mécanisme catalytique différent de celui de ThyA. Compte tenu de la présence unique de ThyX dans de nombreux microorganismes pathogènes pour l'Homme (voir Tableau 1), l'idée que l'inhibition spécifique de cette enzyme conduirait à l'arrêt de la prolifération de l'agent pathogène apparaît séduisante.

Tableau 1 : Exemples de bactéries pathogènes possédant le gène *thyX* et maladies associées

Organismes	Maladies associées	Hôte
<i>Anaplasma marginale</i>	Anaplasmose	Bovin
<i>Bacillus anthracis</i>	Anthrax	Humain
<i>Borrelia burgdorferi</i>	Maladie de Lyme	Humain
<i>Campylobacter coli</i>	Gastro-entérites, septicémies, avortements	Humain
<i>Campylobacter jejuni</i>	Intoxications alimentaires	Humain
<i>Campylobacter lari</i>	Gastro-entérites, septicémies	Humain
<i>Campylobacter upsaliensis</i>	Gastro-entérites, septicémies, abcès, avortements	Humain
<i>Chlamydia pneumoniae</i>	Pneumonie	Humain
<i>Chlamydia trachomatis</i>	Trachomes	Humain
<i>Clostridium botulinum</i>	Botulisme	Humain
<i>Corynebacterium diphtheriae</i>	Diphthérie	Humain
<i>Ehrlichia canis</i>	Ehrlichiose	Canin, Humain
<i>Helicobacter hepaticus</i>	Troubles hépatiques	Humain
<i>Helicobacter pylori</i>	Ulcères à l'estomac	Humain
<i>Leptospira interrogans</i>	Leptospirose, maladie de Weil	Humain
<i>Mycobacterium leprae</i>	Lèpre	Humain
<i>Mycobacterium tuberculosis</i>	Tuberculose	Humain
<i>Rickettsia prowazekii</i>	Typhus	Humain
<i>Treponema pallidum</i>	Syphilis	Humain
<i>Tropheryma whipplei</i>	Maladie de Whipple (lipodystrophie intestinale)	Humain

En parallèle de mes travaux de thèse, Yap Boum du laboratoire d'Hannu Myllykallio a optimisé les tests d'activité de ThyX sur microplaque 96 puits. Le criblage de trois banques de petites molécules a été réalisé sur la thymidylate synthase ThyX de PBCV-1 et sur celles de trois bactéries pathogènes : *Helicobacter pylori*, *Mycobacterium tuberculosis* et *Chlamydia trachomatis*. Les banques de molécules testées présentent une grande diversité chimique et sont constituées de molécules naturelles et synthétiques, dont une faible proportion d'analogues de nucléotides. Parmi la dizaine d'inhibiteurs nanomolaires spécifiques de ThyX identifiés lors de ce criblage, une molécule a montré une activité inhibitrice sur la croissance de l'agent pathogène *Helicobacter pylori* avec une Concentration Minimale d'Inhibition inférieure à 1 µg/ml (communication personnelle de Yap Boum). Cette molécule a servi de base à la construction d'une nouvelle banque de 240 molécules ayant en commun leur structure et/ou leur pharmacophore. Le criblage de cette nouvelle banque a permis d'identifier deux nouveaux inhibiteurs spécifiques de ThyX ayant tous un effet inhibiteur sur la

croissance d'*Helicobacter pylori*. Afin d'évaluer la sélectivité de ces hits, un criblage a été réalisé *in vitro* sur la thymidylate synthase humaine, et aucune molécule ne présente d'effet sur l'activité de ThyA. L'ensemble de ces résultats est très encourageant pour le développement de nouveaux antibiotiques ciblant spécifiquement l'étape de synthèse du thymidylate dans les bactéries pathogènes utilisant ThyX (déclaration d'invention est en cours de rédaction).

### c. Le phénomène de « thymineless death »

Nous avons souligné l'efficacité des drogues qui ciblent l'étape de synthèse du dTMP, mais que se passe-t-il au niveau moléculaire dans la cellule ? Depuis plus de cinq décennies, il est connu que la privation en thymidine chez des microorganismes auxotrophes pour ce métabolite induit une mort cellulaire dénommée « mort par absence de thymine » ou « thymineless death » (TLD) (Cohen et Barner, 1954). Ce phénomène a été observé chez de nombreux organismes bactériens à Gram négatif et positif (voir Tableau 2), mais également chez la levure *Saccharomyces cerevisiae* et les cellules de mammifères (Ahmad *et coll.*, 1998).

**Tableau 2 : Exemple de bactéries dont des mutants auxotrophes pour la thymidine ont été décrits [d'après (Ahmad *et coll.*, 1998)]**

Bactéries à Gram négatif	Bactéries à Gram positif
<i>Aerobacter aerogenes</i>	<i>Bacillus anthracis</i>
<i>Escherichia coli</i>	<i>Bacillus megatarium</i>
<i>Klebsiella aerogenes</i>	<i>Bacillus subtilis</i>
<i>Mycoplasma laidlawii</i>	<i>Bacillus thuringiensis</i>
<i>Pseudomonas acidovorans</i>	<i>Diplococcus pneumoniae</i>
<i>Pseudomonas stutzeri</i>	<i>Enterococcus</i>
<i>Proteus mirabilis</i>	<i>Lactobacillus acidophilus</i>
<i>Salmonella typhimurium</i>	<i>Micrococcus radiodurans</i>
	<i>Pneumococcus</i>
	<i>Staphylococcus aureus</i>
	<i>Streptococcus faecalis</i>

Malgré l'abondance de publications sur le sujet, les mécanismes moléculaires de la TLD ne sont toujours pas totalement élucidés, même dans un organisme aussi intensivement étudié qu'*E. coli*. La diversité de résultats obtenus concernant la TLD au sein d'une même espèce rend la



proposition d'un modèle moléculaire simple délicate. Les différences de résultats peuvent être la conséquence de différence dans le contexte génétique des souches employées, des voies de réparation de l'ADN présentes et des conditions physiologiques et de culture employées. Dans leur revue, Ahmad, Kirk et Eisenstark proposent tout de même un modèle en se basant sur les points communs relevés dans les études concernant la TLD (Ahmad *et coll.*, 1998). Si on considère le cas de cellules se reproduisant activement qui sont ensuite placées dans un milieu dépourvu en thymidylate, deux cas vont se présenter. Les cellules ayant déjà terminé la réplication de leur ADN vont entrer en phase de division et seront protégées de la TLD à condition de ne pas commencer un nouveau cycle de réplication. Les cellules qui initient ou sont en cours de synthèse d'ADN continueront en utilisant le thymidylate restant et ceci jusqu'à épuisement du stock de ce dernier. A mesure que le stock de dTMP va diminuer, la probabilité de misincorporation du dUTP dans l'ADN augmente. L'uracile inséré dans l'ADN est le substrat de l'uracile-N-glycosylase (UNG), qui va produire des sites apuriniques (AP). Une AP endonucléase va reconnaître les sites AP comme substrat et générer des lésions (« nicks ») dans l'ADN. De plus, lors de la réplication, l'ADN polymérase va laisser un interstice (gap) face aux résidus adénine ou incorporer une base incorrecte. Les lésions ainsi générées, normalement réparées dans la cellule en présence de dTTP, vont ici s'accumuler. Dans cette situation, les auteurs proposent que des cassures double brin de l'ADN doivent être la cause majeure de la TLD, expliquant pourquoi des mutations dans certains gènes codant de protéines de réparation modulent la réponse à la TLD (voir tableau 3). Les auteurs proposent ainsi que les voies de réparation par recombinaison doivent être principalement impliquées.

Tableau 3 : Modulation de la TLD par mutation dans des gènes de la réparation de l'ADN [d'après (Ahmad *et coll.*, 1998)]

Gène	Fonction ou effet
Mutants résistants à la TLD	
<i>dnaA</i>	Initiation de la réplication de l'ADN
<i>dnaB</i>	élongation de la chaîne d'ADN
<i>groE</i>	Protéine de stress
<i>infA</i>	Protein chain initiation factor 1
<i>psiB</i>	Inhibiteur de l'induction SOS
<i>recA</i>	Recombinaison et réparation de l'ADN
<i>recF</i>	Recombinaison et réparation de l'ADN
<i>recJ</i>	Recombinaison et réparation de l'ADN
<i>recO</i>	Recombinaison et réparation de l'ADN
<i>recQ</i>	Recombinaison et réparation de l'ADN
<i>ung</i> de <i>B. subtilis</i>	Uracile DNA glycosylase
Mutants sensibles à la TLD	
<i>recBC</i>	Exonucléase V
<i>lig</i>	DNA ligase
<i>polA</i>	DNA polymerase I
<i>rep</i>	Hélicase
<i>uvrA</i>	UV endonuclease
<i>uvrD</i>	DNA-dependent ATPase et hélicase II
<i>dnaG</i>	Primase
<i>dnaC/D</i>	Initiation and chain elongation
Mutants ne présentant pas de changement à la TLD	
<i>lexA</i>	Répresseur SOS
<i>recN</i>	Recombinaison et réparation de l'ADN
<i>ung</i>	Uracile DNA glycosylase

Ce modèle s'appuie sur des travaux qui ont été effectués, pour la plupart, avant les années 1990. Parmi les rares travaux récents concernant la TLD chez les bactéries, les résultats de deux études indépendantes viennent compléter le modèle proposé ci-dessus. La première étude menée par Morganroth et Hanawalt avait pour but d'observer l'influence de la réplication sur la TLD en inhibant la synthèse d'ADN par de l'hydroxyurée (HU) (Morganroth et Hanawalt, 2006). Cette drogue utilisée à faible dose bloque la synthèse des quatre désoxyribonucléotides en inhibant spécifiquement la ribonucléoside diphosphate réductase. Les auteurs ont observé une diminution de la viabilité en condition de limitation en thymidine même en présence de HU. Ils ont pu ainsi montrer que la synthèse active d'ADN n'est pas obligatoirement requise pour provoquer la TLD. Guarino et

collaborateurs ont eux mesurés la présence de cassures double brin de l'ADN dans trois conditions limitantes en désoxyribonucléotides : inhibition de la ribonucléoside diphosphate réductase par l'HU, thermosensibilité de la ribonucléoside diphosphate réductase et par privation en thymine (Guarino *et coll.*, 2007). Ils ont ainsi pu montrer qu'en présence de HU des cassures double brin dépendantes de la résolvasse RuvABC sont formées indiquant que le redémarrage des fourches de réplication se fait par le phénomène de « réversion de fourche de réplication ». De façon plus surprenante, en cas de privation en thymine, des cassures double brin sont également induites mais indépendamment du système RuvABC, traduisant l'absence de réversion de fourche. Les auteurs proposent que dans le cas d'une privation en thymine le problème majeur n'est pas l'arrêt de la réplication mais la nature et le devenir des fourches de réplication arrêtées. Un autre modèle a été proposé par Sat et collaborateurs en se basant sur leur étude chez *E. coli* (Sat *et coll.*, 2003). Ils proposent que la TLD puisse être dépendante du système toxine/antitoxine *mazEF* et induirait ainsi un système programmé de mort cellulaire. Les auteurs ont montré que le système *mazEF* est induit suite à des modifications transcriptionnelles consécutives à la limitation en thymine.

Comme nous l'avons vu de nombreuses voies semblent impliquées dans la TLD et il n'est pas possible à ce jour de proposer un modèle moléculaire précis et surtout commun à l'ensemble des organismes. La généralisation du modèle *mazEF* nécessiterait des études complémentaires afin de rechercher si des systèmes toxine/antitoxine analogues sont également induits dans d'autres organismes lors d'une privation en thymine. Il est intéressant de noter que l'ensemble des travaux réalisés à ce jour concernant la TLD a été effectué chez des organismes utilisant l'enzyme ThyA pour synthétiser le dTMP. Ceci soulève une interrogation concernant ce phénomène chez les organismes utilisant ThyX.

### ***B. Les voies de récupération et de recyclage des pyrimidines***

Chez de nombreux microorganismes, les nucléobases et nucléosides peuvent être récupérés au lieu d'être synthétisés *de novo*. Ces voies de récupération remplissent trois fonctions majeures

dans la cellule : permettre la synthèse de l'ARN et de l'ADN ; fournir une source de carbone, d'azote et d'énergie ; ainsi que recycler les nucléobases et nucléosides issus du turnover des nucléotides. La récupération de nucléobases et nucléosides exogènes représente un bénéfice pour la cellule. En présence de sources exogènes, la cellule va favoriser la voie de récupération, en réprimant la synthèse des enzymes de la voie *de novo* et en augmentant son rétrocontrôle négatif. Les différentes pyrimidines récupérables sont l'uracile, l'uridine, la cytosine, le désoxyuridine, le désoxycytidine, la thymine et le désoxythymidine. Je détaillerai ici la voie de récupération de l'uracile qui a été l'objet de travaux présentés dans le chapitre III de cette thèse.

L'uracile peut être converti en UMP, soit via un intermédiaire uridine par l'action concertée de l'uridine phosphorylase et de l'uridine kinase, soit directement par l'uracile phosphoribosyltransférase (UPRTase). L'UPRTase catalyse le transfert du résidu ribosyl-phosphate d'une molécule de PRPP sur la nucléobase uracile. L'équation de la réaction catalysée est présentée dans la Figure 6.

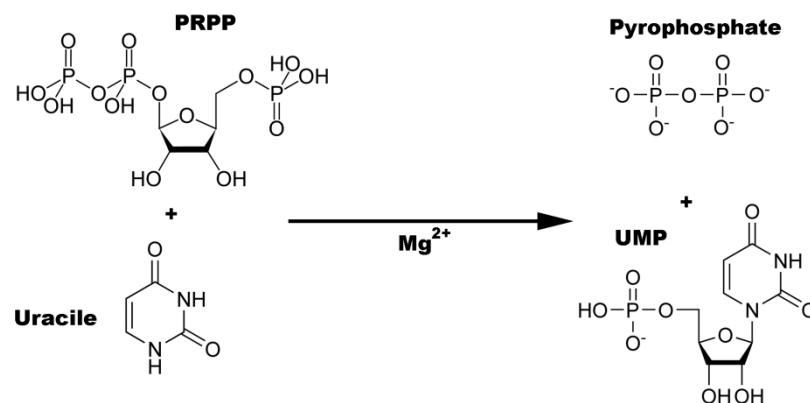


Figure 6 : Réaction catalysée par l'UPRTase.

Cette enzyme a été identifiée chez les bactéries (*Escherichia coli*, *Lactococcus lactis*, *Bacillus subtilis*), les archées (*Sulfolobus shibatae*) et les organismes eucaryotes dits inférieurs (*Saccharomyces cerevisiae*, *Toxoplasma gondii*). Les structures tridimensionnelles d'UPRTases ont permis de mettre en évidence que la reconnaissance du substrat uracile est dépendante d'un

arrangement stéréochimique spécifique empêchant la fixation de substrats portant un groupe exocyclique important en position 5. Bien que les plantes dites supérieures comme *Arabidopsis thaliana* possèdent l'UPRTase, aucune UPRTase n'a été décrite chez les mammifères. Une protéine présentant un domaine putatif UPRTase a été identifiée et partiellement caractérisée chez l'Homme, mais aucune activité UPRTase n'a pu être mesurée *in vitro*. En l'absence de toute preuve expérimentale, il n'est pas raisonnable de conclure si cette enzyme, abusivement dénommée « human UPRTase », est responsable de la faible activité UPRTase mesurée dans les cellules humaines. Chez l'Homme, cette activité proviendrait plus vraisemblablement d'une enzyme de la voie *de novo* de formation de l'UMP, l'OPRTase.

### **C. Le métabolisme des pyrimidines chez *Helicobacter pylori***

#### **a. Pourquoi *Helicobacter pylori* ?**

Un des modèles expérimentaux que j'ai été amené à manipuler durant ces travaux de thèse est la bactérie pathogène humaine *Helicobacter pylori*. Cette bactérie à Gram négatif de la famille des  $\epsilon$ -protéobactéries infecte environ la moitié de la population mondiale. Sa transmission se fait préférentiellement suivant la voie orale-orale, et fécale-orale à cause du manque d'hygiène et à la mauvaise qualité de l'eau (surtout dans les pays en voie de développement) (Megraud, 2003). La transmission peut être aussi génitale-orale en liaison avec certaines pratiques sexuelles (Eslick, 2000). La bactérie *H. pylori* est l'agent responsable de gastrites chroniques qui peuvent évoluer en ulcération gastroduodénale. La présence de ce pathogène est associée avec le développement de cancers gastriques et de lymphome de MALT. L'infection par *H. pylori* peut être traitée par une antibiothérapie, mais l'émergence de souches résistantes justifie la recherche de nouvelles cibles thérapeutiques ciblées. La virulence des bactéries pathogènes est dépendante non seulement de facteurs de virulence spécifiques, requis pour provoquer une infection, mais aussi d'une multitude de fonctions métaboliques nécessaire à la survie de la bactérie dans son microenvironnement spécifique comme la synthèse *de novo* des nucléotides. Comme nous allons le voir dans le paragraphe suivant,

la synthèse des nucléotides chez *H. pylori*, et plus particulièrement celle des pyrimidines, est non redondante et fortement dépendante des voies de synthèse *de novo*. Ainsi, le blocage d'une seule étape par l'inhibition d'une seule enzyme devrait suffire à bloquer la voie entière et conduire à une mort cellulaire.

#### b. Voies de novo et de récupération des pyrimidines

Le nombre de travaux concernant le métabolisme des nucléotides chez *H. pylori*, et celui des pyrimidines en particulier, est limité comparativement aux études concernant le pouvoir pathogène de la bactérie. Les bases des connaissances expérimentales concernant ces voies métaboliques chez *H. pylori* sont issues des travaux de Georges L. Mendz et de ses collaborateurs au milieu des années 1990 (Menz *et coll.*, 1994). L'analyse des séquences génomiques disponibles quelques années plus tard ont permis de comprendre la majorité des résultats obtenus précédemment (Marais *et coll.*, 1999; Tomb *et coll.*, 1997).

En se basant seulement sur les informations que nous fournit l'analyse des différents génomes de *H. pylori*, nous pouvons prédire que les voies *de novo* de synthèse des nucléotides et des déoxyribonucléotides pyrimidiques sont complètes. Il a été montré à plusieurs reprises que *H. pylori* est capable de croître sur un milieu défini dépourvu de pyrimidines préformées (Albertson *et coll.*, 1998; Testerman *et coll.*, 2001). Ceci confirme bien la précédente prédiction et confirme que la bactérie peut répliquer son génome en utilisant seulement les nucléotides pyrimidiques synthétisés *de novo*. Les activités enzymatiques ATCase, DHOase, OPRCase, ODCase, CTP synthétase, correspondantes à différentes étapes de cette voie *de novo*, ont pu également être mesurées *in situ* (Menz *et coll.*, 1994). La protéine responsable de la synthèse *de novo* du thymidylate dans cette bactérie (codée par le gène *hp1533*) n'a été découverte que récemment lors de la mise en évidence de la seconde famille de thymidylate synthases ThyX (Myllykallio *et coll.*, 2002).

Si la voie *de novo* apparaît complète et fonctionnelle comme nous venons de le voir, la situation concernant les voies de récupération des pyrimidines est moins claire. L'analyse génomique

semble indiquer que ces voies sont absentes. En effet, on ne retrouve pas de gène homologue à ceux codant les enzymes nécessaire à la récupération de l'uracile (UPRTase), l'uridine (uridine phosphorylase et uridine kinase), la thymidine (thymidine phosphorylase et thymidine kinase). Les données expérimentales basées sur l'incorporation de pyrimidines radio-marquées et sur la mesure des activités enzymatiques *in situ* sont globalement en accord avec la prédiction théorique. Cependant, deux résultats expérimentaux sont rétrospectivement inattendus. En premier lieu, bien qu'aucune incorporation significative de [6-<sup>3</sup>H]thymidine et de [2-<sup>14</sup>C]thymine n'ai pu être mesurée, une faible activité thymidine kinase a été mesurée. L'explication la plus vraisemblable proposée par les auteurs est que l'activité mesurée correspondrait à une spécificité de substrat relaxée d'une déoxypurine nucleoside kinase. Ainsi l'activité thymidine kinase ne serait donc pas physiologiquement significative. Le deuxième point concernant la récupération d'uracile et sa conversion en UMP est plus intrigant. En effet, les expériences d'incorporation de [2-<sup>14</sup>C]uracile ont montré une linéarité d'import du composé durant les 10 premières heures. Ceci semble indiquer que l'uracile est utilisé intact. De plus, une faible activité UPRTase (10 pmol d'UMP formé par minute et par mg de protéine) a été mesurée à partir de lysat cellulaire de *H. pylori*. Ainsi les résultats expérimentaux semblent indiquer que l'uracile pénètre la cellule, puis est métabolisé en UMP. A ce jour, le mécanisme responsable de l'incorporation d'uracile exogène est inconnu. Ces données expérimentales semblent ainsi en désaccord avec les données génomiques qui indiquent l'absence de gènes homologues à *uraA* (codant le transporteur de l'uracile chez *E. coli*) et *upp* (codant l'UPRTase). L'analyse de ce paradoxe fera l'objet de travaux exposés dans le chapitre III de cette thèse.

c. Les enzymes du métabolisme des pyrimidines chez *H. pylori*, de bonnes cibles antimicrobiennes ?

L'ensemble des données présentées ci-dessus indique que la voie *de novo* est suffisante et nécessaire à la croissance et à la survie de la bactérie pathogène *H. pylori*. Il apparaît évident qu'en inhibant spécifiquement l'activité d'une des enzymes impliquées dans ce métabolisme, on devrait

arrêter la synthèse des précurseurs des acides ribo- et désoxyribonucléiques, bloquer la transcription et la réplication de l'ADN. Trois travaux indépendants ont pu montrer l'essentialité de la voie *de novo* de synthèse des pyrimidines, et ce à différentes étapes de la voie.

La première étude présentée montre l'importance de la deuxième étape de cette voie métabolique. Burns et collaborateurs ont tenté d'inactiver le gène *pyrB* codant l'aspartate carbamoyltransferase (ACTase) par remplacement allélique avec une cassette de résistance au chloramphenicol (Cm) (Burns *et coll.*, 2000). Malgré plusieurs essais, aucun transformant résistant au Cm n'a pu être obtenu. Ce résultat semble indiquer l'essentialité de *pyrB*. Les auteurs ont également tenté de contourner l'étape catalysée par l'ACTase en ajoutant au milieu de culture soit de l'orotate soit de l'uracile. En effet, des travaux précédents avaient montré que ces deux composés semblaient être convertis en UMP *in vivo* (Mendz *et coll.*, 1994). Toutefois, même en présence d'uracile ou d'orotate, aucune bactérie portant une délétion du gène *pyrB* n'a pu être isolée. Ce résultat peut être interprété d'au moins deux manières. La première interprétation – pessimiste – conduit à l'impossibilité de conclure sur l'essentialité de *pyrB*. En étant plus optimiste, il est raisonnable de penser que la voie de récupération de l'uracile et de l'orotate, en incluant leur transport, pourrait ne pas être suffisamment active pour permettre à elles seules la survie de la bactérie.

La quatrième étape de la voie *de novo* est la synthèse d'orotate par la dihydroorotate dehydrogenase (DHODase). Copeland et collaborateurs ont développé des composés dérivés de pyrazole inhibiteurs de DHODase de *H. pylori* ( $K_i = 26 \pm 12$  nM), mais n'inhibant pas les enzymes homologues bactériennes Gram-positives (*Enterococcus faecalis*), ni la DHODase humaine ( $K_i > 10^5$  nM) (Copeland *et coll.*, 2000). Les auteurs ont testé l'effet de leurs molécules sur la croissance de 8 souches de *H. pylori* (dont 7 isolats cliniques); ils ont pu ainsi déterminer une concentration minimale d'inhibition (MIC) variant de 1 à 16 µg/ml. L'hypothèse posée par les auteurs est que la DHODase est la cible préférentielle de leurs composés *in vivo*. Les auteurs ont ajouté des concentrations croissantes d'orotate (produit de la réaction catalysée par la DHODase) au milieu de



culture. Ils ont ainsi montré que la MIC augmente en fonction de l'augmentation de la concentration en orotate dans le milieu, validant ainsi leur hypothèse. Cette étude montre clairement que la DHODase est une bonne cible antimicrobienne.

La troisième étude illustrant l'importance de la voie de synthèse *de novo* des pyrimidines visait l'identification systématique des gènes essentiels chez *H. pylori* (Chalker *et coll.*, 2001). Durant ce travail, Chalker et collaborateurs ont utilisé la même stratégie que celle décrite pour Burns et collaborateurs plus haut. Les auteurs ont ainsi tenté d'inactiver le gène *pyrE* codant l'orotate phosphorybosyltransferase (OPRTase) en présence ou absence d'uracile, OMP ou UMP. Dans toutes les conditions testées, aucune bactérie mutante n'a été sélectionnée. Les auteurs concluent ainsi à l'essentialité du gène *pyrE* et conséquemment de l'enzyme OPRTase.

Considérant ces trois études, les résultats combinés semblent démontrer que la voie de synthèse *de novo* des pyrimidines est essentielle à la survie de la bactérie *H. pylori*. Il est intéressant de noter que dans les trois cas, le rôle essentiel des enzymes est conditionné à la faible efficacité des voies de récupération conduisant à la formation de l'UMP, et en particulier celle de l'uracile. Nous pouvons donc nous interroger sur les risques liés à l'utilisation de composés inhibant les premières étapes de cette voie (avant la formation de l'UMP). Dans ces conditions, une dérégulation de la voie de récupération de l'uracile pourrait permettre une synthèse d'UMP suffisante à la survie de la bactérie, et ainsi d'éviter l'étape inhibée, ce qui pourrait conduire à l'apparition de souches résistantes. En absence de voie de récupération de la thymidine, il pourrait apparaître plus judicieux d'inhiber les étapes tardives de la voie métabolique, en bloquant par exemple la synthèse du thymidylate.

## **Chapitre II. Les génomes procaryotes**

Dans la première partie de cette introduction présentant les voies de synthèse des désoxyribonucléotides, j'ai insisté sur les différences qui peuvent exister entre les deux familles de thymidylate synthases. Les travaux présentés dans le chapitre II montreront comment la différence

d'efficacité catalytique influence indirectement la dynamique des génomes procaryotes (Article II). Afin de replacer ce travail dans son contexte, je présenterai ici une vision d'ensemble des différents facteurs décrits dans la littérature pouvant influencer la composition et la taille des génomes procaryotes. Avant cela, je vais rappeler brièvement les caractéristiques générales des génomes procaryotes, les informations concernant les génomes eucaryotes n'étant données qu'à titre de comparaison.

## I. Structure

La vision traditionnelle du génome procaryote est un génome petit, composé d'un unique chromosome circulaire d'ADN double-brin et qui possède une origine de réplication diamétralement opposé à une région terminus. Chaque cellule ne contiendrait qu'une seule copie du chromosome mais pourrait contenir d'autres réplicons sous la forme d'un ou de plusieurs plasmides. Cette description peut être considérée comme acceptable dans de nombreux cas, tel que pour la bactérie à Gram négatif *E. coli*, cependant il existe de plusieurs exceptions à ces règles. Par exemple, les chromosomes de certaines bactéries se présentent sous une forme linéaire et non circulaire comme chez *Streptomyces coelicolor* (Kieser *et coll.*, 1992), *Borrelia burgdorferi* (Ferdows et Barbour, 1989) ou *Agrobacterium tumefaciens* (Allardet-Servent *et coll.*, 1993). Outre la forme, le nombre de chromosomes peut aussi varier. Dans certains cas, la taille des plasmides peut être semblable à celle du chromosome principal, on parlera alors de méga-plasmides ou de deuxième chromosome. Nous nous heurtons ici à un problème de terminologie : comment distinguer un chromosome d'un méga-plasmide ? La plupart des auteurs considère qu'un chromosome porte des fonctions essentielles à la survie de la cellule, cette notion est bien évidemment subjective et dépend des conditions testées. Malgré tout, nous pouvons considérer qu'un réplicon est un chromosome s'il porte des gènes essentiels (comme les gènes codant les protéines de ménage), s'il n'est pas curable de la cellule ou bien s'il porte des opérons codant des ARN ribosomiques.

Tableau 4 : Exemples de bactéries possédant plusieurs réplicons [d'après (Ochman, 2002)]

Espèces	Appellation	Taille (kb)	Forme	nb d'ADNr
<i>Agrobacterium tumefaciens</i>	Chromosome	2843	Circulaire	3
	Chromosome	2057	Linéaire	2
	Plasmide	543	Circulaire	0
	Plasmide	214	Circulaire	0
<i>Borrelia burgdorferi</i>	Chromosome	911	Linéaire	1
	Plasmides (n=11)	9-54	Circulaire/Linéaire	0
<i>Brucella melitensis</i>	Chromosome	2117	Circulaire	3
	Chromosome	1178	Circulaire	1
<i>Clostridium acetobutylicum</i>	Chromosome	3941	Circulaire	11
	Mégaplasmide	192	Circulaire	0
<i>Deinococcus radiodurans</i>	Chromosome	2649	Circulaire	3
	Chromosome	412	Circulaire	0
	Mégaplasmide	177	Circulaire	0
	Plasmide	46	Circulaire	0
<i>Ralstonia solanacearum</i>	Chromosome	3716	Circulaire	3
	Mégaplasmide	2095	Circulaire	1
<i>Salmonella typhi</i>	Chromosome	4809	Circulaire	7
	Plasmide	218	Circulaire	0
	Plasmide	107	Circulaire	0
<i>Sinorhizobium meliloti</i>	Chromosome	3654	Circulaire	3
	Mégaplasmide	1683	Circulaire	0
	Mégaplasmide	1334	Circulaire	0
<i>Vibrio cholerae</i>	Chromosome	2941	Circulaire	8
	Chromosome	1072	Circulaire	0
<i>Yersinia pestis</i>	Chromosome	4654	Circulaire	6
	Plasmides (n=3)	10-96	Circulaire	0

## II. La composition en bases

Le contenu en bases guanine (G) et cytosine (C) ou « GC content » est la mesure la plus courante de la composition d'un génome. C'est en fait, le contenu - relatif - en bases G+C qui est utilisé ou  $GC\% = \frac{(G+C)}{(G+C)+(A+T)}$ . On dira d'un génome qu'il est riche en GC ou au contraire riche en AT (et non pauvre en GC). A ce jour, les valeurs extrêmes de GC% décrites chez les procaryotes sont 16,5% [*Carsonella ruddii* (Nakabachi et coll., 2006)] et 75% [*Anaeromyxobacter dehalogenans* 2CP-C (Sanford et coll., 2002)]. Bien que le GC% est connu pour être variable d'un groupe bactérien à un autre depuis longtemps (Sueoka, 1962), la disponibilité d'un nombre croissant de séquences

génomiques montre qu'il peut se révéler très variable au sein d'un même groupe phylogénique (Ochman, 2002). On pourra noter en particulier, le groupe des  $\alpha$ -protéobactéries dans lequel le GC% varie de 27,5% [*Ehrlichia ruminantium*] à 69% [*Rhodobacter sphaeroides*].

De nombreux auteurs ont tenté et cherchent toujours à identifier les pressions de sélection qui s'exercent sur les génomes pour les rendre plus ou moins riches en GC. La plupart des hypothèses émises sont fondées sur des causes environnementales. Dès 1970, Singer et Ames proposaient qu'un haut pourcentage en GC pouvait, dans certains cas, être une réponse à une exposition au rayonnement ultraviolet (Singer et Ames, 1970). McEwan et collaborateurs ont montré qu'il pouvait exister une relation entre le GC% et la capacité à fixer l'azote au sein d'un même groupe de bactéries (McEwan *et coll.*, 1998). Plus récemment, Musto et collaborateurs ont décrit une corrélation positive entre la température optimale de croissance et le GC% chez certains groupes de procaryotes (Musto *et coll.*, 2004). Cependant la validité de cette relation fait actuellement l'objet d'un débat (Marashi et Ghalanbor, 2004; Mitchell, 2007; Musto *et coll.*, 2005; Wang *et coll.*, 2006). Une critique qui pourrait être apportée aux précédentes études est qu'elles semblent refléter des pressions de sélection spécifiques et ne peuvent pas servir de base pour une compréhension générale de la variation du GC%. Ce n'est que récemment que Naya et collaborateurs ont proposé une explication plus générale à la variation du GC% chez les procaryotes (Naya *et coll.*, 2002). Les auteurs ont en effet montré que les organismes aérobies possédaient un haut GC% comparativement aux organismes anaérobies. Les auteurs proposent que l'augmentation du GC% puisse servir de protection contre les dommages liés à l'oxygène 1) en limitant la fréquence des acides aminés sensibles à l'oxydation en particulier la tyrosine et la phénylalanine (codé exclusivement par des A et des T aux sites non-synonymes) ; 2) en augmentant la thermo-stabilité de l'ADN au niveau des lésions de l'ADN ; 3) en offrant "en sacrifice" un G en position synonyme des codons. La relation entre la composition en bases et l'environnement mise en évidence par Naya et collaborateurs a été confirmée à l'échelle de communautés microbienne par une approche métagénomique par Foerstner et collaborateurs (Foerstner *et coll.*, 2005). Les auteurs ont montré que les microorganismes dans un environnement donné (eau de surface

océanique, sol de ferme, carcasse de baleine en eaux profondes et biofilm acidophile de drainage de mine) présentent des compositions semblables, indépendamment des distances phylogéniques. L'ensemble de ces travaux semble indiquer que la composition en bases nucléotidiques peut être influencée par la nature de la niche écologique. Récemment, Zhao et collaborateurs ont affirmé que le GC% était gouverné par la nature de la sous-unité dimérique alpha de la DNA polymérase III (DnaE1 ; DnaE2/DnaE1 ; PolC/DnaE3) (Zhao *et coll.*, 2007). Ce travail au titre provocateur est, à mon avis, à traiter avec prudence tant la nature même des groupements utilisés par les auteurs est parfois peu explicite. De plus aucune hypothèse plausible n'est avancée pour expliquer les corrélations décrites.

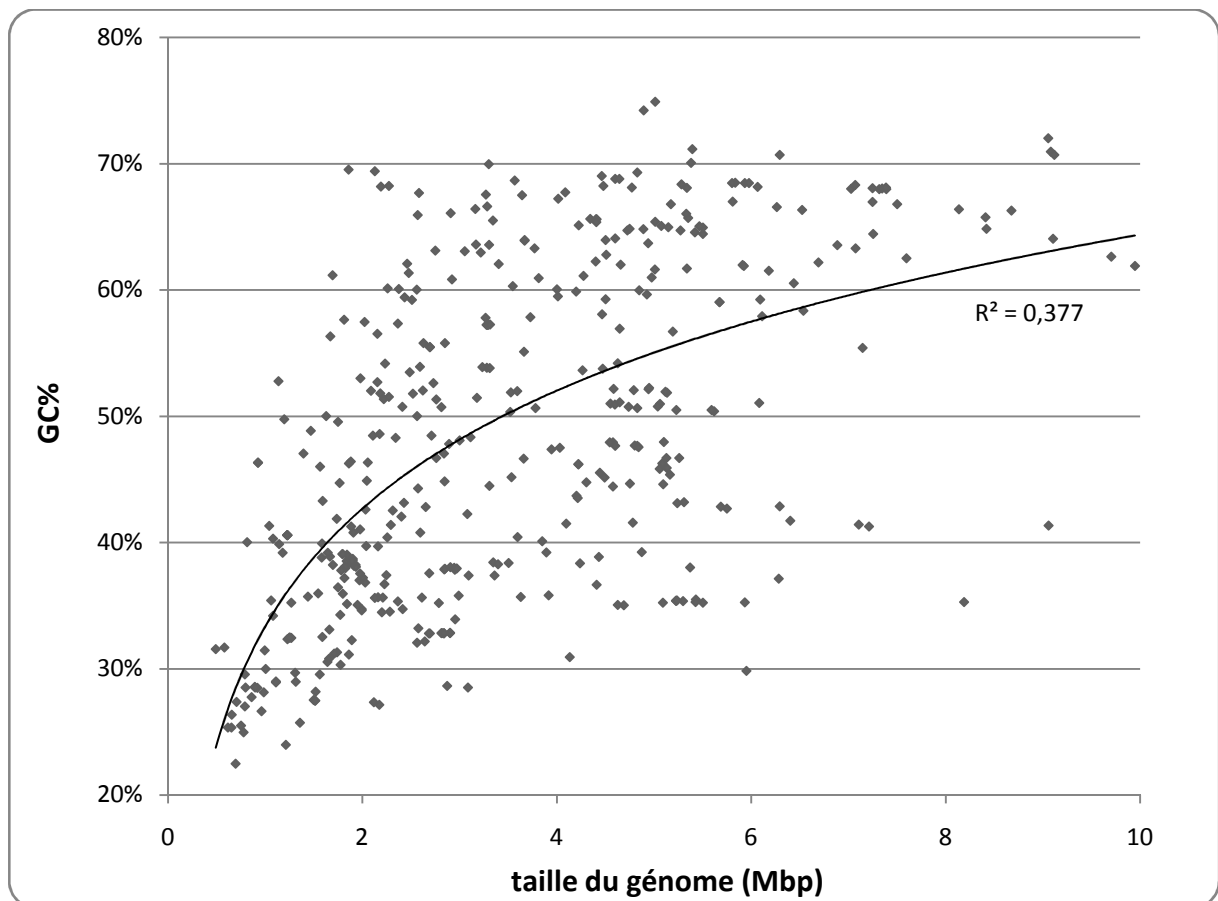


Figure 7 : Relation entre la composition relative en GC (GC%) et la taille du génome (données issues de l'article II).

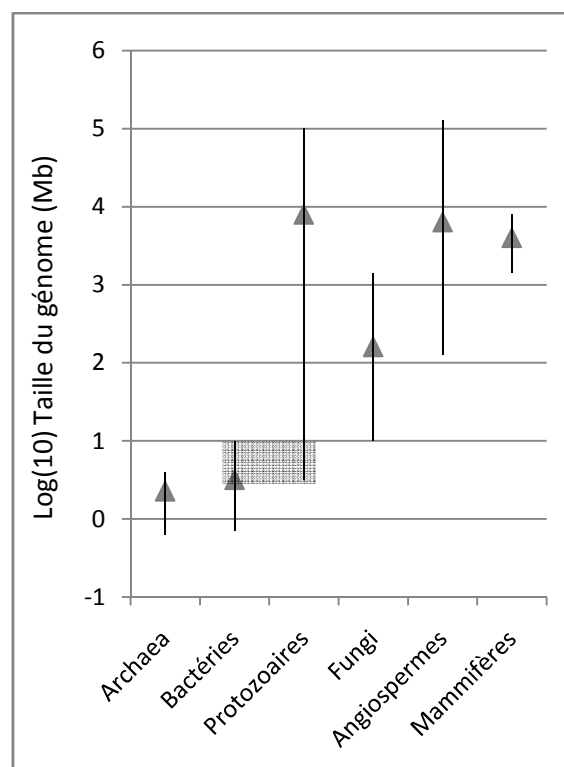
Ce que n'explique pas les théories exposées ci-dessus, c'est l'apparente corrélation positive entre la composition en GC et la taille du génome (Figure 7 ; coefficient de Spearman  $r = 0.59$  ;

$P < 0,0001$ ) (Bentley et Parkhill, 2004; Foerstner *et coll.*, 2005; Moran, 2002). Plusieurs théories sont proposées pour expliquer la richesse en AT des petits génomes des bactéries intracellulaires. La première implique l'absence ou la très faible activité de protéines impliquées dans les mécanismes de réparation de l'ADN et particulier l'UNG (Glass *et coll.*, 2000). Dans ces conditions, l'incorporation d'uracile dans l'ADN, par mutation ou par déamination de la cytosine (C→U), conduira lors de la réplication suivante à l'appariement de l'uracile à l'adénine résultant en une transition d'une paire G-C en une paire A-T. Si ce mécanisme semble probable pour *Ureaplasma* et *Mycoplasma* (Glass *et coll.*, 2000), aucune preuve expérimentale ne permet de l'étendre aux autres bactéries à petit génome riche en AT. Rocha et Danchin proposent une explication alternative qui lie la composition en GC à la disponibilité en nucléotides (Rocha et Danchin, 2002). Leur hypothèse repose sur le fait que les nucléotides GTP et CTP représentent un coût énergétique supérieur à l'ATP et l'UTP, et que l'ATP est abondamment présent compte tenu de son rôle majeur dans la cellule. Si les ressources venaient à être restreintes, la disponibilité en A et T serait plus grande que celle en C et G induisant un biais vers l'enrichissement en AT.

### III. La taille des génomes procaryotes

Le terme de taille du génome (ou « genome size ») désigne la quantité d'ADN contenue dans une copie complète du génome. Elle est mesurée soit en masse d'ADN, exprimée en picogrammes (pg), soit en nombre de paires de bases nucléotidiques, exprimée en million de paires de bases ou megabases (Mb ou Mbp). Il est possible de convertir une unité en l'autre en utilisant la relation  $1 \text{ pg} = 978 \text{ Mbp}$ . Les travaux concernant la taille de génomes procaryotes utiliseront exclusivement l'expression en paires de bases. Préalablement à la généralisation du séquençage, diverses techniques ont permis de mesurer, ou du moins d'estimer, la taille des génomes. Nous compterons principalement parmi les diverses techniques employées, l'électrophorèse en champ pulsé qui permet la migration de larges fragments d'ADN en utilisant l'alternance spatio-temporelle des champs électriques.

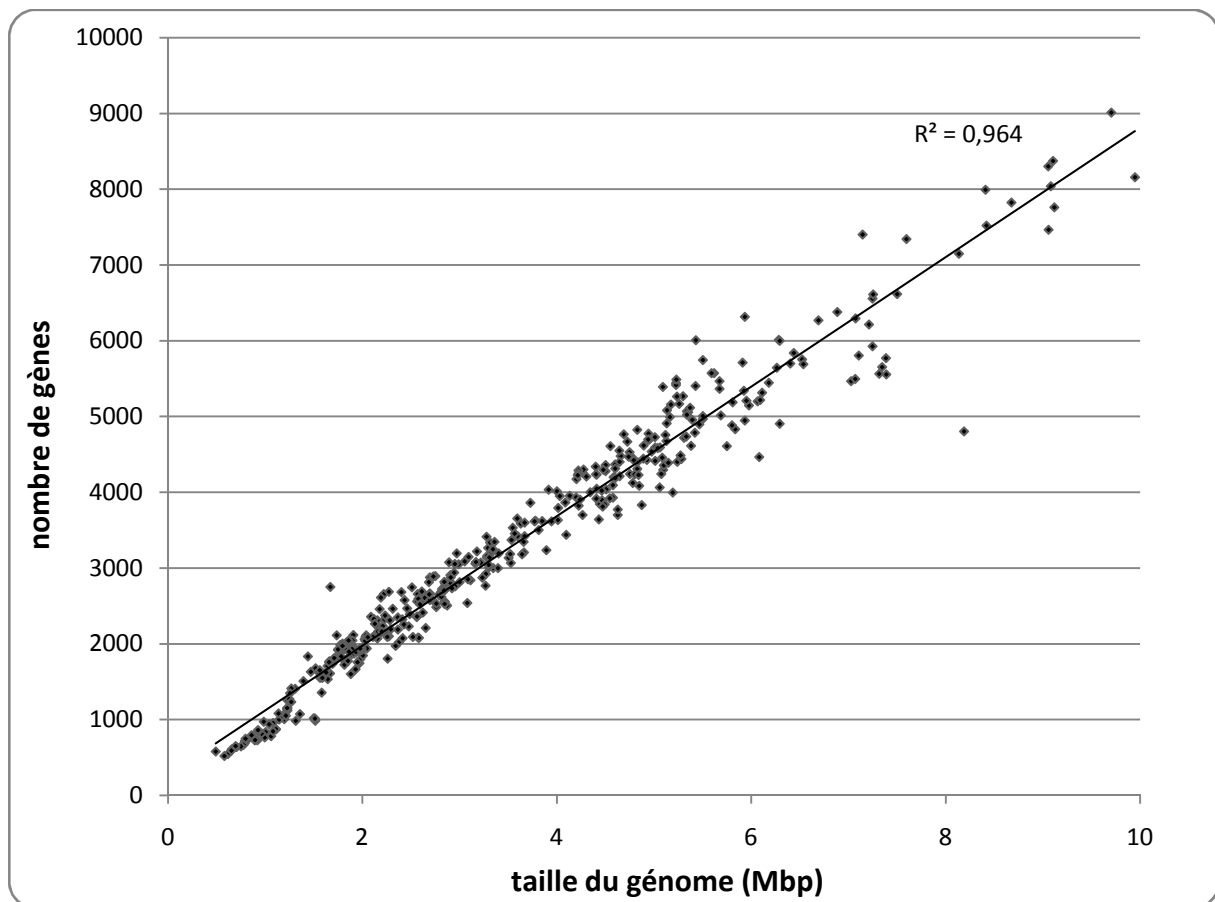
Comme nous l'avons vu dans l'introduction de cette partie, les génomes procaryotes sont décrits comme étant « petits », leur taille variant de 0,16 Mpb [*Carsonella ruddii* (Nakabachi *et coll.*, 2006)] à 9,9 Mpb [*Solibacter usitatus* (NCBI NC\_008536)]. Par petits, on entend petits comparativement aux génomes eucaryotes. Si comparativement aux génomes de mammifères, les génomes des procaryotes sont en effet beaucoup plus petits, il n'en est pas de même concernant les génomes des protistes. En effet, il existe un chevauchement entre les tailles des plus grands génomes procaryotes et celles des plus petits génomes protistes (Figure 8).



**Figure 8 : Distribution de la taille des génomes en fonction du groupe phylogénétique [modifié d'après (Gregory, 2005)].** Chaque élément représente les tailles minimales et maximales observées, ainsi que la taille moyenne (triangle plein). La zone grisée illustre la région de chevauchement de taille entre les génomes bactériens et protozoaires.

La variation de taille des génomes procaryotes varie d'un facteur d'environ 60, ce qui est très peu comparativement aux eucaryotes (facteur >45000) (Vellai et Vida, 1999). La grande variabilité de tailles des génomes eucaryotes a été, et est toujours, une source d'interrogations pour les scientifiques étudiant ce domaine. En effet, dès les années 1950, on s'est aperçu qu'il n'y avait pas de corrélation entre la complexité d'un organisme et la taille de son génome (ou valeur C), ce

phénomène fut appelé alors le paradoxe de la valeur C. Ce paradoxe fut rapidement résolu dans les années 1970 par la découverte de large quantité d'ADN non-codant présente dans les génomes eucaryotes (Cavalier-Smith, 1978). Ainsi il n'est plus question de paradoxe, mais une énigme persiste autour de la valeur C concernant la nature, l'origine, le rôle et la proportion de cet ADN non-codant. La situation peut paraître plus simple chez les procaryotes. Il n'existe pas d'énigme de la valeur C, car il existe une relation linéaire entre le nombre de gènes et la taille du génome (Figure 9). Ce phénomène se comprend par la faible proportion, environ 14%, d'ADN non codant présent dans les génomes procaryotes.



**Figure 9 : Relation entre le nombre de gènes et la taille du génome chez les procaryotes (données en annexe)**

Une question se pose : pourquoi les organismes procaryotes ont-ils des génomes si petits ? L'hypothèse la plus simple pourrait être que les génomes procaryotes ne peuvent pas s'agrandir car ils sont incapables d'intégrer de l'ADN exogène. Cette hypothèse peut être réfutée immédiatement



car les voies d'acquisition d'ADN, tel le transfert horizontal de gène et la duplication, sont présentes et variées chez les procaryotes [pour revue (Abby et Daubin, 2007)]. Leurs génomes sont dans un système dynamique de gain et de perte d'ADN sous l'influence de forces de sélection (Mira *et coll.*, 2001; Moran et Plague, 2004). Le gain d'ADN sera sélectionné par gain de fonction, alors que la perte résultera d'un biais délétère ou d'une dérive génétique. Dans le cas où une bactérie passerait d'une vie libre à une vie parasitaire dans un hôte, la probabilité d'acquérir de l'ADN exogène diminuant l'équilibre va se décaler vers la perte de matériel génétique. A l'opposé, la théorie nous indique qu'une bactérie libre à grand génome aura acquis des nouveaux gènes qui auront été sélectionnés pour leur avantage sélectif, comme la multiplication des voies métaboliques (Mira *et coll.*, 2001). Si ce modèle convient parfaitement pour expliquer la réduction des génomes des bactéries intracellulaires, il ne propose pas d'explication satisfaisante à l'apparente limitation en taille des génomes bactériens à environ 10 Mpb. Une hypothèse originale a été proposée en 2005 par Ranea et collaborateurs basée sur la comparaison entre la cellule bactérienne et une usine (Ranea *et coll.*, 2005). L'augmentation du nombre de fonctions métaboliques qu'une cellule est capable de réaliser représente un avantage reproductif en lui permettant de croître dans des milieux variés et peut être ainsi définie comme le « revenu de l'entreprise bactérienne ». Les gènes de régulation associés à ces voies métaboliques représentent eux le « coût logistique » nécessaire pour permettre la bonne marche de l'usine. Les auteurs ont remarqué deux grandes tendances : il existe une relation linéaire entre la taille du génome et le nombre de gènes dans les superfamilles de gènes correspondant aux fonctions métaboliques dans la cellule ; et une relation exponentielle entre la taille du génome et le nombre de gènes dans les superfamilles de gènes correspondant aux fonctions de régulation. Dit autrement, le coût logistique augmente plus rapidement que le revenu de l'entreprise. Ainsi, en appliquant des principes de base de la microéconomie qui définissent la taille optimale d'une usine, les auteurs ont déterminé que la taille optimale de « l'usine bactérienne » était d'environ 4800 gènes. Ils ont également pu prédire qu'au-delà de 10500 gènes la complexité liée à la régulation va

excéder la complexité métabolique, ainsi cette valeur représenterait la taille limite pour une bactérie. Cette limite est de façon étonnante très proche de la réalité biologique.

## Problématique de la thèse

La synthèse du dTMP est une étape cruciale dans le métabolisme des pyrimidines mais également pour la survie de tout organisme cellulaire. L'existence de deux voies – indépendantes – pour catalyser cette étape est intrigante. En effet, il est difficile d'imaginer une raison simple qui expliquerait pourquoi ces deux familles d'enzymes coexistent dans le monde procaryote sans pour autant suivre la phylogénie des organismes. Une grande partie des travaux présentés dans cette thèse a eu pour but de répondre à cette question. De façon plus générale, la problématique de cette thèse était de mieux comprendre quelles pouvaient être les conséquences de l'utilisation de la thymidylate synthase ThyX sur la physiologie bactérienne.

Les travaux réalisés durant cette thèse vont être présentés sous la forme de trois chapitres. Dans un premier temps, j'exposerai un travail concernant les implications de l'utilisation de ThyX sur le métabolisme des folates. En combinant des approches génétiques et une modélisation mathématique nous avons montré en quoi le métabolisme des folates diffère entre les bactéries suivant si elles utilisent ThyA, ThyX ou bien les deux enzymes. Le deuxième chapitre de cette thèse sera consacré à l'étude de la relation entre le type de thymidylate synthase utilisée et la vitesse de réplication de l'ADN chez les procaryotes. Ce travail qui combine des approches génomiques et génétiques, tout en intégrant les données biochimiques préexistantes, m'a permis de proposer un modèle concernant l'influence des thymidylate synthases dans le modelage des génomes procaryotes. Dans le troisième chapitre, je présenterai les résultats que j'ai obtenus concernant l'étude du métabolisme des pyrimidines chez la bactérie *H. pylori*. J'exposerai en premier lieu les travaux que j'ai effectué sur la thymidylate synthase ThyX, et ensuite mes travaux concernant l'orotate phosphorybosyltransferase. Pour conclure, j'exposerai les implications des travaux présentés dans cette thèse et les perspectives envisagées.



## **RESULTATS ET DISCUSSION**

## Chapitre I. Thymidylate synthases et métabolisme des folates

### I. Introduction

Comme nous l'avons vu dans l'introduction (voir p 18), bien que le méthylène-tétrahydrofolate (MTHF ou CH<sub>2</sub>THF) soit le substrat des deux familles de thymidylate synthases, sa fonction au cours de la catalyse diffère. Lors de la réaction catalysée par ThyX, le MTHF sert uniquement de donneur de groupement méthylène conduisant à la formation de tétrahydrofolate (THF). Dans la réaction catalysée par ThyA, en plus de son rôle de donneur de méthylène, il réalise la réduction du méthylène en méthyle, conduisant à la formation de dihydrofolate (DHF). Ainsi il apparaît évident que la nature de la thymidylate synthase présente dans un organisme va directement influencer la nature des folates présents dans la cellule. Les résultats présentés dans l'article I nous ont permis de mieux comprendre la nature de ces influences et les répercussions sur la physiologie de la bactérie. Afin de mieux appréhender ce travail, je vais introduire brièvement la nature et le rôle des dérivés foliques impliqués dans cette voie.

#### A. Rôle clé du tétrahydrofolate et de ses dérivés

Le THF est composé de trois groupes : un groupement ptéridine, un *p*-aminobenzoate et une chaîne mono-ou-polyglutamée (Figure 10).

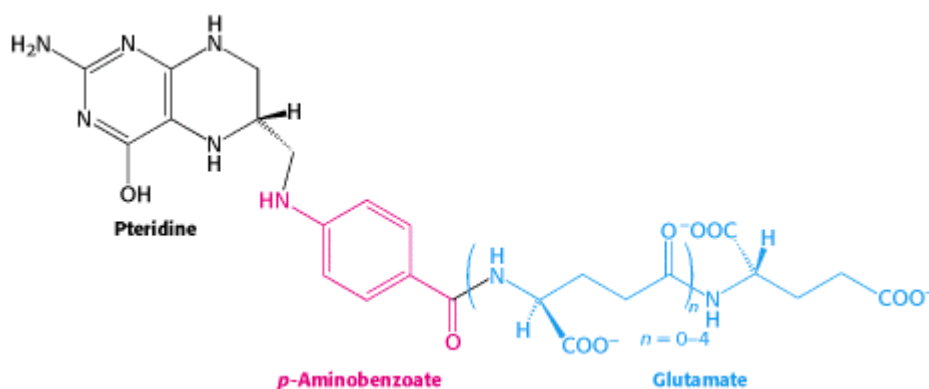


Figure 10 : Structure du tétrahydrofolate.

Le THF agit comme transporteur -et donc donneur- d'unités mono-carbonées qui sont liés aux atomes azotés N-5 et/ou N-10 ( $N^5$  et  $N^{10}$ ). Ces unités diffèrent selon leur degré d'oxydation ; trois

groupes peuvent être définis du plus réduit au plus oxydé: méthyle ; méthylène ; formyle, formimino ou méthényle (Tableau 5).

Tableau 5 : Groupes mono-carbonés portés par le THF

Degré d'oxydation	Groupement	
	Formule	Nom
Le plus réduit (=méthanol)	-CH <sub>3</sub>	Méthyle
Intermédiaire (=formaldéhyde)	-CH <sub>2</sub>	Méthylène
	-CHO	Formyle
Le plus oxydé (=acide formique)	-CHNH	Formimino
	-CH=	Méthényle

Les dérivés du THF servent de donneurs d'unité mono-carbonée dans différentes voies de biosynthèse (voir figure 1 de l'article I). Nous pourrions citer par exemple la synthèse de méthionine à partir d'homocystéine par transfert d'un groupement méthyle issu du *N*<sup>5</sup>-méthyltétrahydrofolate ; ou encore la synthèse des purines qui nécessite l'acquisition d'unité mono-carbonée à partir du *N*<sup>10</sup>-formyltétrahydrofolate. Bien évidemment la voie de biosynthèse du dTMP est dépendante du transfert d'un groupement méthyle à partir du MTHF. On peut noter que dans le cas de la réaction catalysée par ThyA ce dernier agira également comme réducteur du groupement méthylène. Les folates ne sont pas métaboliquement actif dans la cellule sous forme oxydée, ainsi le DHF produit lors de la réaction catalysée par ThyA doit être réduit en THF par la dihydrofolate réductase (DHFR). La DHFR a donc pour rôle principal de recycler les folates oxydés lors de la réaction catalysée par ThyA.

### B. Problématique

Avant le début de ma thèse, les preuves expérimentales de l'activité thymidylate synthase des protéines de la famille ThyX étaient de deux natures. D'une part, il avait été montré que la surexpression du gène *thyX*, provenant des divers organismes, est capable de restaurer la viabilité d'une souche d'*E. coli* dont le gène *thyA* est inactivé ; d'autre part l'activité thymidylate synthase de ThyX avait été démontrée *in vitro* (Mylykallio *et coll.*, 2002). Ainsi aucun des travaux publiés alors ne prouvait directement que ThyX soit bien responsable de la synthèse du dTMP *in vivo*. Le but initial

des travaux présentés dans l'article I était de démontrer expérimentalement l'activité thymidylate synthase de ThyX *in vivo*. Pour cela, il fallait trouver un organisme génétiquement modifiable, possédant le gène *thyX* mais ne portant pas de gène *thyA* et *folA*. Afin de pouvoir inactiver *thyX*, il fallait apporter à la cellule une source de dTMP pour permettre une synthèse correcte d'ADN. Ainsi, la construction d'une souche  $\Delta thyX$  ne pouvait se faire que dans un organisme possédant le gène *tdk* codant la thymidine kinase indispensable à la récupération de la thymidine exogène. C'est avec ces contraintes (20% des organismes ThyX possède *tdk* contre 58% pour les organismes ThyA), que Damien Leduc, alors doctorant dans l'équipe de Hannu Myllykallio, initia une étude génétique pour montrer le rôle *in vivo* de ThyX en utilisant l'alpha-protéobactérie *Rhodobacter capsulatus* (*thyX*<sup>+</sup>, *tdk*<sup>+</sup>, *thyA*<sup>-</sup>, *folA*<sup>-</sup>). Il lui a été possible, en présence de thymidine exogène, d'inactiver le gène *thyX* en remplaçant la copie chromosomique par une autre interrompue par une cassette de résistance à la spectinomycine. Il a ainsi pu montrer que ThyX est indispensable à la survie de la bactérie en absence de thymidine. Il a également montré, par une approche de complémentation fonctionnelle, que l'expression conjointe de ThyA et Fola (de *Rhodobacter sphaeroides*) permettait de restaurer la synthèse de dTMP et la croissance de la souche  $\Delta thyX$ .

C'est à cette étape que j'ai commencé à travailler sur ce projet. J'ai en particulier voulu mieux comprendre quelles étaient les conséquences de l'utilisation exclusive de l'une ou l'autre des thymidylate synthases sur le métabolisme des folates. En effet, les travaux génétiques présentés dans l'article I montraient que la complémentation de l'absence de *thyX* par le gène *thyA* seul (*thyX*<sup>-</sup>, *folA*<sup>-</sup>) conduit à une synthèse réduite de dTMP et une forte altération de la croissance de la bactérie. Ce résultat montrait que le DHF formé lors de la réaction catalysée par ThyA est inefficacement réduit en THF en absence de Fola, indiquant qu'une faible activité DHFR doit être présente chez *R. capsulatus* mais qu'elle n'est pas suffisante pour permettre une croissance rapide. J'ai émis l'hypothèse, qu'en présence de *thyA* seul, les folates devaient s'accumuler, au moins temporairement, sous forme de DHF et que la cellule devait être rapidement carencée en THF expliquant les phénotypes décrits plus haut. J'ai ainsi testé l'effet de l'addition de différents folates,

en particulier du THF et du MTHF, sur la croissance des différentes souches construites lors de cette étude. Comme nous le verrons par la suite, ces expériences ont abouti à un résultat inattendu.

En 2005, un modèle mathématique du cycle du folate chez l'Homme basé sur les données biochimiques disponibles a été publié (Nijhout *et coll.*, 2004). Afin de transposer ce modèle au système bactérien, nous avons décidé de simuler les expériences présentées dans l'article I. Nous avons établi une collaboration avec Frederik Nijhout et Michael Reed, auteurs du modèle humain. L'adaptation du modèle original au système bactérien a nécessité la recherche des voies métaboliques spécifiquement humaine à retirer, et celles présentes uniquement chez les bactéries à ajouter. J'ai ensuite réalisé une recherche bibliographique afin de trouver les constantes catalytiques des différentes enzymes bactériennes impliquées dans ces voies métaboliques. Enfin, nos collaborateurs ont simulé différentes situations, comme l'absence d'activité DHFR dans un organisme ThyA ou dans un organisme possédant les deux thymidylate synthases. Les résultats de ces simulations seront discutés dans le dernier paragraphe de l'article I.



**II. Article I : “Flavin-dependent thymidylate synthase ThyX activity : Implications for the folate cycle in Bacteria”**

**Damien Leduc\*, Frédéric Escartin\*, H. Frederik Nijhout, Michael C. Reed, Ursula Liebl,**

**Stéphane Skouloubris, Hannu Myllykallio**

*J Bacteriol.* 2007 Dec;**189**(23):8537-45.

\* les deux auteurs ont participé de façon égale à ce travail

## Flavin-Dependent Thymidylate Synthase ThyX Activity: Implications for the Folate Cycle in Bacteria<sup>∇†</sup>

Damien Leduc,<sup>1,2‡§</sup> Frédéric Escartin,<sup>3,4‡</sup> H. Frederik Nijhout,<sup>5</sup> Michael C. Reed,<sup>6</sup> Ursula Liebl,<sup>3,4</sup> Stéphane Skouloubris,<sup>1,2</sup> and Hannu Myllykallio<sup>1,2\*</sup>

INSERM Avenir group, Institut de Génétique et de Microbiologie, CNRS UMR8621, F-91405 Orsay, France<sup>1</sup>; Université Paris-Sud, F-91405 Orsay, France<sup>2</sup>; Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS UMR7645, F-91128 Palaiseau, France<sup>3</sup>; INSERM U696, F-91128 Palaiseau, France<sup>4</sup>; Department of Biology, Duke University, Durham, North Carolina 27708<sup>5</sup>; and Department of Mathematics, Duke University, Durham, North Carolina 27708<sup>6</sup>

Received 24 August 2007/Accepted 10 September 2007

**Although flavin-dependent ThyX proteins show thymidylate synthase activity in vitro and functionally complement *thyA* defects in heterologous systems, direct proof of their cellular functions is missing. Using insertional mutagenesis of *Rhodobacter capsulatus thyX*, we constructed the first defined *thyX* inactivation mutant. Phenotypic analyses of the obtained mutant strain confirmed that *R. capsulatus* ThyX is required for de novo thymidylate synthesis. Full complementation of the *R. capsulatus thyX::spec* strain to thymidine prototrophy required not only the canonical thymidylate synthase ThyA but also the dihydrofolate reductase FolaA. Strikingly, we also found that addition of exogenous methylenetetrahydrofolate transiently inhibited the growth of the different *Rhodobacter* strains used in this work. To rationalize these experimental results, we used a mathematical model of bacterial folate metabolism. This model suggests that a very low dihydrofolate reductase activity is enough to rescue significant thymidylate synthesis in the presence of ThyX proteins and is in agreement with the notion that intracellular accumulation of folates results in growth inhibition. In addition, our observations suggest that the presence of flavin-dependent thymidylate synthase X provides growth benefits under conditions in which the level of reduced folate derivatives is compromised.**

The folate cycle plays a central role in cell metabolism. Folate-dependent enzymes are required for methionine synthesis, numerous methylation reactions, and synthesis of purine and pyrimidine nucleotides. As the different loops of the folate cycle are interconnected, a mathematical model of this cycle has been described recently for eukaryotic cells (21). This basic model has the qualitative behavior seen in a variety of experimental studies on folate homeostasis in the cytosol of human cells. Moreover, it predicts that the activities of folate-dependent enzymes depend on the size of the total folate pool in a nonlinear fashion (23). For instance, actively dividing cells require large quantities of the DNA precursor thymidylate (dTMP). In human cells, the thymidylate synthase ThyA (EC 2.1.1.45) catalyzes the reductive methylation of dUMP to dTMP, using 5,10-methylenetetrahydrofolate (CH<sub>2</sub>THF) as a donor of one-carbon units and as a reductant (2). The resulting 7,8-dihydrofolate (DHF) is reduced to tetrahydrofolate (THF) by the dihydrofolate reductase (DHFR) FolaA (Fig. 1). As formation of dTMP is rate limiting for DNA replication, in human cells the *thyA* gene is up-regulated by the transcription

factor E2-F (5). This greatly enhances dTMP synthesis, whereas other branches of folate metabolism are scarcely affected (21).

In contrast to the human thymidylate synthase ThyA, the members of the novel ThyX family of thymidylate synthases (EC 2.1.1.148) are NAD(P)H oxidases that use flavin adenine nucleotide to mediate hydride transfer (1, 10, 11, 20). Therefore, although both ThyA and ThyX catalyze the formation of thymidylate in vitro, their reductive mechanisms are dramatically different (Fig. 1). ThyX catalysis results in the formation of THF, not DHF, as the product of the methylation reaction (12, 19), but virtually nothing is known to date about how the activity of the flavin-dependent thymidylate synthase ThyX influences the different folate-dependent branches of bacterial metabolism. This is of particular interest as, e.g., *Mycobacterium* and *Corynebacterium* species contain *thyA* and *thyX* genes, but to date why both genes are maintained in these organisms is poorly understood.

Previous genetic studies on *thyX* either were performed using poorly defined genetic backgrounds (7) or were based upon multicopy heterologous complementation systems using either bacterial (9, 12, 20, 29) or viral (10, 11) *thyX* genes.

In order to obtain direct evidence for the in vivo role of ThyX enzymes, we inactivated the *thyX* gene from the purple bacterium *Rhodobacter capsulatus* by insertion of an antibiotic cartridge. The *thyX::Spec<sup>r</sup>* mutant obtained was viable only when thymidine was provided in the growth medium, proving that ThyX proteins indeed function in de novo thymidylate synthesis. We also showed that functional complementation of the *thyX::Spec<sup>r</sup>* mutant requires not only *thyA* but also *folA*, whereas earlier studies showed that *thyX* alone complements a *thyA* deletion strain. This growth

\* Corresponding author. Mailing address: INSERM Avenir group, Institut de Génétique et de Microbiologie, CNRS UMR8621, F-91405 Orsay, France. Phone: 33 169 158 170. Fax: 33 169 157 808. E-mail: hannu.myllykallio@igmors.u-psud.fr.

† Supplemental material for this article may be found at <http://jb.asm.org/>.

‡ D.L. and F.E. contributed equally to this work.

§ Present address: Unité de Pathogénie Bactérienne des Muqueuses, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France.

∇ Published ahead of print on 21 September 2007.

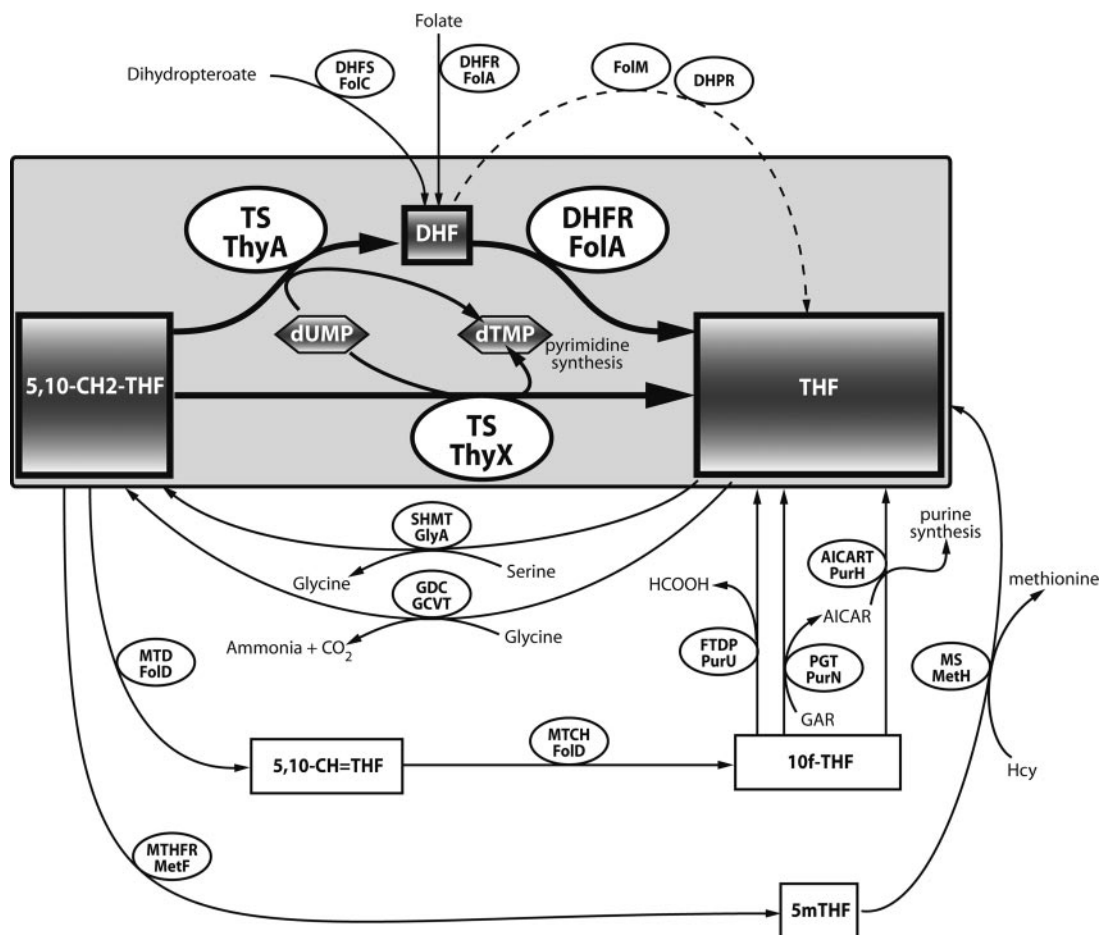


FIG. 1. Schematic representation of different folate metabolic pathways identified in *Bacteria*. The ThyX and ThyA pathways for de novo thymidylate synthesis are highlighted by the large gray box. Although both thymidylate synthases produce dTMP, they use different reductive mechanisms. Genetic inactivation of FolA or inhibition of its activity by trimethoprim leads to the accumulation of DHF in ThyA-containing organisms, thus resulting in growth inhibition (see Fig. 5). In contrast, ThyX proteins produce THF directly as a reaction product. GlyA (serine hydroxymethyltransferase) can be found in both ThyX- and ThyA-containing organisms. All the reactions were considered for simulations. Substrates are enclosed in rectangles, and enzymes are enclosed in ellipses. In several cases both the eukaryotic and bacterial abbreviations for the enzymes are shown (redox active compounds participating in many of the reactions are not shown). Abbreviations: 5mTHF, 5-methyltetrahydrofolate; 5,10-CH<sub>2</sub>-THF, 5,10-methylenetetrahydrofolate; 5,10-CH=THF, 5,10-methenyltetrahydrofolate; 10f-THF, 10-formyltetrahydrofolate; AICART, aminoimidazolecarboxamide ribotide transformylase; TS, thymidylate synthase; MTD, 5,10-methylenetetrahydrofolate dehydrogenase; MTCH, 5,10-methylenetetrahydrofolate cyclohydrolase; PGT, phosphoribosyl glycinamidetransformylase; SHMT, serine hydroxymethyltransferase; MTHFR, 5,10-methylenetetrahydrofolate reductase; MS, methionine synthase; GAR, glycinamide ribotide; AICAR, aminoimidazolecarboxamide ribotide; GDC, glycine decarboxylase; DHPR, dihydropteridine reductase; DHFS, dihydrofolate synthase; GCVT, glycine cleavage system aminomethyltransferase; FTDP, formyltetrahydrofolate deformylase; Hcy, homocysteine.

defect in the absence of FolA cannot be rescued by addition of exogenous folates, which unexpectedly results in growth inhibition. Mathematical modeling of folate metabolism for ThyX-containing bacteria in the presence and absence of ThyA indicated that ThyA is a critical determinant of reduced folates in actively dividing bacterial cells. In particular, in organisms with only the thymidylate synthase ThyX, a very low level of DHFR activity is sufficient to rescue significant thymidylate synthesis. These findings suggest that the presence of ThyX proteins is beneficial when the steady-state level of reduced folate molecules is compromised (for instance, upon exposure to antifolates).

#### MATERIALS AND METHODS

**Bacterial strains, media, and growth conditions.** The bacterial strains and plasmids used in this study and their relevant characteristics are listed in

Table 1. *R. capsulatus* strains were grown at 35°C under chemoheterotrophic or photoheterotrophic conditions using enriched MPYE medium (3 g peptone per liter, 3 g yeast extract per liter, 1.6 mM MgCl<sub>2</sub>, 1.0 mM CaCl<sub>2</sub>) or on Sistrom's minimal medium (27a). For photoheterotrophic growth on solid medium, anaerobic jars and H<sub>2</sub>/CO<sub>2</sub> generator envelopes from BBL GasPak were used. Respiratory doubling times under aerobic growth conditions in liquid medium were calculated from semilogarithmic growth curves obtained at 630 nm. CH<sub>2</sub>-THF and THF were kindly provided by R. Moser (Merck Eprova AG). Where necessary, antibiotics, nucleotides, and folate derivatives were added at the following final concentrations: tetracycline, 2 µg/ml; spectinomycin, 10 µg/ml; rifampin, 75 µg/ml; thymidine, 1 or 50 µg/ml; and folates, 50 or 100 µg/ml.

**Chromosomal inactivation of the *R. capsulatus* thyX gene by insertion mutagenesis.** All oligonucleotides used in this work were obtained from Sigma-Genosys. The plasmid constructs obtained were confirmed by restriction analyses and DNA sequencing. To construct pDL5 (Table 1), the chromosomal region carrying the *R. capsulatus* strain MT1131 thyX gene was amplified by PCR using primers ORF311-C and ORF311-E (Table 1). The amplified 1.7-kb DNA fragment was inserted into the pGEM-T Easy vector (Promega), and its correct

TABLE 1. Bacterial strains, plasmids, and oligonucleotides used in this work

Strain, plasmid, or oligonucleotide	Relevant characteristics	Source or reference
<i>R. capsulatus</i> strains		
MT1131	Parental strain	F. Daldal
DL1	<i>thyX</i> ::Spec <sup>r</sup>	This study
<i>R. sphaeroides</i> Ga		
	<i>crt</i>	F. Daldal
<i>E. coli</i> strains		
SURE	e14 <sup>-</sup> (McrA <sup>-</sup> ) Δ( <i>mcrCB-hsdSMR-mrr</i> )171 <i>endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC</i> ::Tn5 (Kan <sup>r</sup> ) <i>uvrC</i> [F <sup>-</sup> <i>proAB lacI<sup>q</sup>ZΔM15</i> Tn10 (Tet <sup>r</sup> )]	Stratagene
HB101	F <sup>-</sup> <i>proA2 hsdS20 recA13 ara-14 lacY xyl-5 galK2 rpsL20 supE44 rplsL20 supE44 proA2 mtl-1</i>	ATCC
S17-1	<i>thi pro ara recA</i> StrepRP4 2-Tet::Mu-Kan::Tn7	27
Plasmids		
pGEM-T Easy	Amp <sup>r</sup>	Promega
pBluescript II KS+	Amp <sup>r</sup>	Stratagene
pRK415	Tet <sup>r</sup> broad-host-range plasmid	13
pRK2013	Kan <sup>r</sup> helper	6
pHP45Ω	Spec <sup>r</sup> Amp <sup>r</sup>	22
pSUP202	Amp <sup>r</sup> Cm <sup>r</sup> Tet <sup>r</sup> , suicide plasmid	27
pDL5	<i>R. capsulatus thyX</i> in pGEM-T Easy	This study
pDL7	3.5-kb PstI fragment with <i>thyX</i> ::Spec <sup>r</sup> allele in pSUP202	This study
pDL9	1.6-kb NotI-BglII fragment with <i>thyX</i> gene in pBluescript II KS+	This study
pDL11	1.4-kb PstI fragment with <i>R. capsulatus thyX</i> gene in pRK415	This study
pDL12	1.7-kb SpeI-EcoRI fragment with <i>R. sphaeroides thyA</i> and <i>folA</i> in pRK415	This study
pDL14	1.4-kb XbaI-EcoRI fragment with <i>R. sphaeroides thyA</i> in pRK415	This study
Oligonucleotides		
ORF311-A	5'-ATGTCGCTTACCCAGACC	This study
ORF311-B	5'-ACCCACTCCTCCACAAACT	This study
ORF311-C	5'-CATGCGGTAGGCGGTATAG	This study
ORF311-E	5'-AATCTCCAAGCGTTTGCTTC	This study
OLI-SPEC	5'-CTCGTTTTCTGGAAGGCGA	This study
THYA-DR-A	5'-CACAACCTCTGGGGATCAAG	This study
THYA-DR-B	5'-GACCTGCATGTTGTCCCTTC	This study

orientation was verified by restriction analyses. The construct was digested with NotI and BglII, and a 1.6-kb fragment containing the *R. capsulatus thyX* gene was ligated into the NotI and BamHI sites of pBluescript II KS+ (Stratagene) to generate pDL9 (Table 1). A 2-kb BamHI fragment containing an Sm<sup>r</sup>-Spec<sup>r</sup> resistance cartridge from pHP45Ω (22) was inserted into the unique BamHI site of pDL9 in order to interrupt the *R. capsulatus thyX* gene. A 3.5-kb PstI fragment of the plasmid obtained was subsequently ligated into the suicide vector pSUP202, yielding pDL7.

pDL7, carrying the *R. capsulatus thyX*::Spec<sup>r</sup> allele, is not capable of replication in *Rhodobacter* strains. This plasmid was transferred into *R. capsulatus* strain MT1131 (Rif<sup>r</sup> Tet<sup>r</sup> Spec<sup>r</sup>) from *Escherichia coli* S17-1 (Rif<sup>r</sup>) by diparental conjugation as described previously (4). As inactivation of *thyX* was expected to result in thymidine auxotrophy, transconjugants were selected for spectinomycin resistance on enriched MPYE solid medium containing thymidine and rifampin added as described above. A total of 235 Spec<sup>r</sup> colonies were screened for tetracycline resistance (single recombination) or sensitivity (double recombination). Fourteen tetracycline-sensitive clones, corresponding to ~6% of the total number of transconjugants, were maintained. The expected chromosomal location of the spectinomycin resistance cartridge and the absence of the *thyX* wild-type gene in strain DL1 were confirmed by PCR using primers ORF311-A, ORF311-B, and OLI-SPEC (Fig. 2).

The pDL11 vector carrying the wild-type allele of *R. capsulatus thyX* (ORF311) (Fig. 2) was obtained by insertion of a 1.4-kb PstI fragment from pDL9 into the low-copy-number vector pRK415 at the unique PstI site (Table 1). In *Rhodobacter sphaeroides*, the *thyA* and *folA* genes form a gene cluster on the chromosome, which allows simultaneous cloning of them. The *R. sphaeroides thyA* and *folA* genes were amplified using primers THYA-DR-A and THYA-DR-B. The 1.7-kb DNA fragment obtained was inserted into pGEM-T Easy (Promega), and its orientation was confirmed by restriction analyses. The construct was digested with SpeI and EcoRI, and a 1.7-kb fragment containing both the *thyA* and *folA* genes was ligated into the XbaI and EcoRI sites of pRK415,

yielding pDL12. The pDL14 vector carrying only *R. sphaeroides thyA* was obtained using primers THYA-DR-A and THYA-DR-B. The pRK415 derivatives pDL11, pDL12, and pDL14 were transferred by triparental conjugation from *E. coli* HB101 (Rif<sup>r</sup>) into *R. capsulatus* DL1 (Rif<sup>r</sup>) with the helper plasmid pRK2013 carrying the RK2 transfer genes (6). Tet<sup>r</sup> transconjugants were selected on enriched MPYE medium plates in the presence of added thymidine and rifampin.

**In vivo labeling of chromosomal DNA with [5,6-<sup>3</sup>H]uracil.** Isotope labeling experiments with *R. capsulatus* cells were performed using MPYE liquid medium containing [5,6-<sup>3</sup>H]uracil (49.0 Ci/mmol; Amersham Biosciences). Radioactively labeled DNA was isolated from early-stationary-phase cultures using cesium chloride gradients. Prior to centrifugation, the refraction indices of the cesium chloride gradients were adjusted to 1.3980, corresponding to a density of 1.6850 g/ml. Uracil incorporated into RNA was not detected, as the gradients under the conditions described separate RNA and DNA.

Dialyzed DNA was hydrolyzed to free nucleobases by boiling it in 50 mM perchloric acid for 1 h. Samples were neutralized with potassium hydroxide, and the free nucleobases were separated by high-performance liquid chromatography (Beckman Gold system) using reverse-phase chromatography (Ultrasphere octyldecyl silane; column dimensions, 4.6 mm by 25 cm). Isocratic elution was performed using 10 mM potassium phosphate buffer (pH 4) at a flow rate of 1 ml/min. Isotope incorporation into deoxycytidine (dC) and deoxythymine (dT) was monitored by using an online radioactive flow detector (LabLogic β-RAM) with FlowLogic SafeScint scintillation liquid. The identities of the radioactive peaks were confirmed by simultaneous monitoring of known nucleobase standards. Data analyses were performed using Laura Light 3 software from LabLogic.

**Mathematical model of bacterial folate metabolism.** We have estimated that the intracellular concentration of the total folate pool in *Bacteria* is approximately 50 μM (17). The general model of Nijhout et al. (21) for mammalian hepatic folate metabolism was modified to make it more appropriate for bacte-

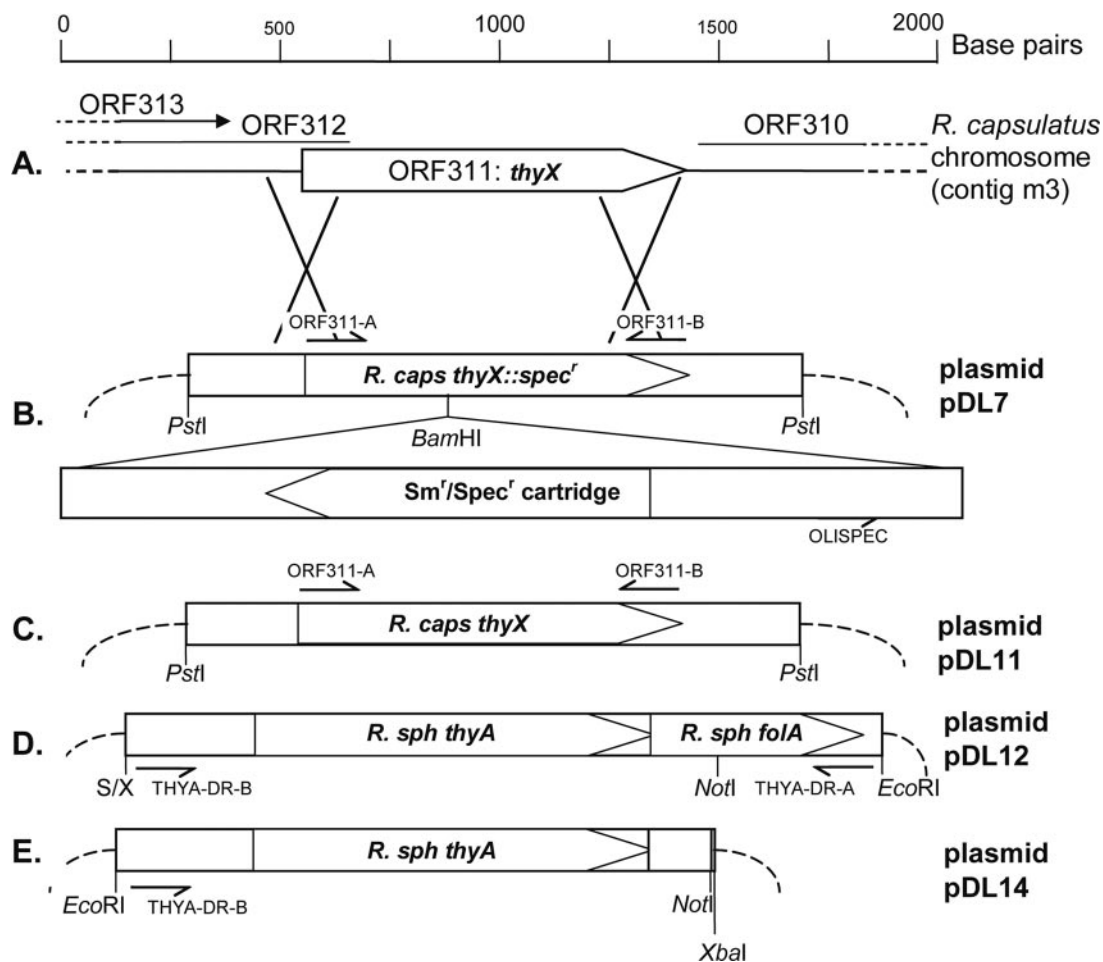


FIG. 2. Physical maps of the chromosomal regions carrying the *thyX* gene of *R. capsulatus*, as well as of the different plasmids used in this work. (A) ORF310, ORF312, and ORF313 are putative genes around *R. capsulatus thyX* (ORF311). (B) Disruption of *R. capsulatus thyX* by insertion of an  $\text{Sm}^r$ - $\text{Spec}^r$  cartridge. The resistance cartridge inserted at the unique *Bam*HI site is transcribed in the direction opposite that of *thyX*. (C to E) Schematic representation of plasmids used in the complementation studies, including the approximate locations of the primers used in cloning (see Table 1). Note that *R. sphaeroides thyA* and *folA* form a gene cluster and are likely translationally coupled (not shown). *R. sph*, *R. sphaeroides*; *R. caps*, *R. capsulatus*; S/X, *SpeI/XbaI* hybrid site.

ria. This involved adding several reactions and modifying the kinetic parameters of many of the enzymes involved. The values used for simulations are included in the supplemental material. The reaction scheme that we modeled is shown in Fig. 1 (which shows the features relevant for thymidylate metabolism). We have added a pathway, catalyzed by ThyX, that converts  $\text{CH}_2\text{THF}$  to THF. We have also added an additional pathway from DHF to THF catalyzed by dihydropteridine reductase and/or FolM. It was demonstrated that these enzymes have been shown to possess DHFR activity in *Thermus thermophilus* (28) and *E. coli* (8), respectively, and are believed to be an alternative pathway from DHF to THF in other bacteria as well. The available data indicate that the catalytic efficiencies of these alternative enzymes with DHFR activity are relatively low (8, 28), as their reported  $V_{\max}$  values appear to correspond to approximately 10 to 25% of the  $V_{\max}$  values measured for a number of bacterial FolA proteins (see the supplemental material). Our earlier in vitro measurements suggested that the dTMP-forming activity of ThyX proteins is significantly less than that of ThyA proteins. Also shown in Fig. 1 is the glycine cleavage system (modeled as the glycine decarboxylation reaction), since it exists in some bacteria but not in the cytosol of eukaryotic cells. We removed the nonenzymatic conversion of THF to  $\text{CH}_2\text{THF}$  from the bacterial model of folate metabolism. Formylmethionyl tRNA<sup>Met</sup> is essential for initiation of protein synthesis in bacterial systems but not in eukaryotic systems and cannot be provided exogenously. Consequently, we added a new substrate, 5-formyltetrahydrofolate, and the reactions that interconvert it with 10-formyltetrahydrofolate and methylenetetrahydrofolate. These reactions are catalyzed by the multifunctional FolD protein in bacteria (3). All of

these enzymes are assumed to have Michaelis-Menten kinetics. Kinetic constants used in simulations were collected from the BRENDA database (<http://www.brenda.uni-koeln.de/>).

## RESULTS

**In silico analysis of the genome sequence of *R. capsulatus* SB1003.** Strains with *thyA* or *thyX* deleted are expected to be either nonviable or thymidine auxotrophic, depending on whether they also carry the *tdk* gene coding for thymidine kinase, which is required for salvage of extracellular thymidine. Consequently, studies on the in vivo functions of flavin-dependent thymidylate synthase X are facilitated by the use of strains that contain *tdk* in addition to *thyX*. Using manual similarity searches and the pedant database (<http://pedant.gsf.de/>) to access the genome sequence of *R. capsulatus* SB1003, we identified several bacterial and archaeal species that apparently contain *thyX* and *tdk* genes, and *R. capsulatus* is one of these species. Importantly, *thyA* and *folA* were not found in the *R. capsulatus* genome sequence. Genetic tools allowing efficient



TABLE 2. Growth phenotypes of *R. capsulatus* strains used in this work<sup>a</sup>

Strain	Genotype	Growth in:			Doubling time (min) with <sup>b</sup> :		
		MPYE medium + thymidine	Minimal medium MA	Minimal medium MA + thymidine	No CH <sub>2</sub> THF	50 µg/ml CH <sub>2</sub> THF	100 µg/ml CH <sub>2</sub> THF
MT1131	<i>thyX</i> <sup>+</sup> (wild type)	+++	+++	+++	122	180	225
DL1	<i>thyX::Spec</i> <sup>r</sup>	+++	–	+++	NM <sup>c</sup>	NM	NM
DL1/pDL7	<i>thyX</i> <sup>+</sup> / <i>thyX::Spec</i> <sup>r</sup>	+++	+++	+++	NM	NM	NM
DL1/pDL12	<i>thyA</i> <sup>+</sup> <i>folA</i> <sup>+</sup> / <i>thyX::Spec</i> <sup>r</sup>	+++	+++	+++	133	198	371
DL1/pDL14	<i>thyA</i> <sup>+</sup> / <i>thyX::Spec</i> <sup>r</sup>	+++	± <sup>d</sup>	+++	287	325	346

<sup>a</sup> Growth phenotypes of the strains were scored after 2 days (MPYE medium plates) or 3 days (minimal medium MA plates) of incubation at 35°C. Similar results were obtained under respiratory and photosynthetic growth conditions.

<sup>b</sup> Doubling times were determined using MPYE medium.

<sup>c</sup> NM, not measured.

<sup>d</sup> Some growth without formation of individual colonies was visible after extended incubation.

manipulation of *Rhodobacter* species have been developed (15, 18). Interestingly, in the closely related species *R. sphaeroides* *thyA* and *folA* are present, whereas *thyX* is absent. Therefore, *R. capsulatus* is an excellent organism for heterologous complementation studies designed to understand the in vivo role(s) of ThyX proteins.

**Inactivation of *R. capsulatus thyX* results in thymidine auxotrophy.** To inactivate *R. capsulatus thyX*, we inserted an Sm<sup>r</sup>-Spec<sup>r</sup> antibiotic resistance cartridge into the unique BamHI site of this gene as described in Materials and Methods (Fig. 2). The *R. capsulatus thyX::Spec*<sup>r</sup> allele obtained was transferred by diparental conjugation into *R. capsulatus* MT1131. Spectinomycin-resistant transconjugants were selected on enriched MPYE medium containing thymidine in the presence of rifampin to eliminate the *E. coli* donor strain. The absence of a wild-type *thyX* copy in the retained tetracycline-sensitive strain was confirmed by PCR. Phenotypic testing of the *thyX::Spec*<sup>r</sup> strain obtained, designated DL1, indicated that it grew on enriched MPYE or minimal medium plates supplemented with 1 µg/ml thymidine (Table 2), whereas in the absence of thymidine, growth was not detected. The observed thymidine auxotrophy did not result from polar effects due to insertion of the Sm<sup>r</sup>-Spec<sup>r</sup> cartridge into the chromosomal *thyX* gene, as the *thyX::Spec*<sup>r</sup> strain after complementation in *trans* by *R. capsulatus thyX* had a doubling time similar to that of the parental strain MT1131 (Table 2). These genetic data provide the first direct indication of the in vivo function of ThyX proteins in de novo synthesis of dTMP.

**Nonorthologous functional replacement of *thyX* by *R. sphaeroides thyA* or *thyA* and *folA* genes.** Heterologous complementation tests have shown that *E. coli thyA* can be functionally replaced by *thyX* genes from a variety of sources (9, 11, 20, 29). To investigate whether the reciprocal nonorthologous replacement is possible, we constructed transferable plasmids carrying either the *R. sphaeroides thyA* gene alone (pDL14) or the *R. sphaeroides thyA* and *folA* genes (pDL12) (Fig. 2). pDL12 and pDL14 were transferred independently into *R. capsulatus thyX::Spec*<sup>r</sup> strain DL1 by triparental conjugation. On solid medium, only the strain complemented by plasmid pDL12 (carrying *R. sphaeroides thyA* and *folA*) formed colonies in a thymidine-independent manner (Table 2), whereas no growth was observed in the presence of pDL14 carrying only the *thyA* gene. This observation indicates that *thyA* alone cannot efficiently replace *R. capsulatus thyX* and that both *thyA* and *folA*

are necessary for full functional complementation of a *thyX* defect under these conditions.

To quantitatively address the level of functional complementation, the doubling times for the strains described above were measured using liquid cultures in minimal medium MA and MA containing thymidine (50 µg/ml). In the presence of thymidine, the doubling times of all strains were essentially identical to that of the wild-type strain. In agreement with phenotypic observations on solid minimal medium, MT1131 and DL1/pDL12 had similar doubling times (103 and 108 min, respectively), whereas the negative control strain DL1/pRK415 was unable to grow under these conditions (Table 2). Strain DL1/pDL14, which grew only very poorly on solid minimal medium (Table 2), had a doubling time of 310 min in liquid minimal medium, a value approximately three times higher than the value measured for the wild-type strain. The slow growth in the presence of ThyA but in the absence of Folate presumably shows that reduced folate derivatives are exhausted during the exponential growth and that the oxidized forms produced, like DHF, are not efficiently regenerated. This results in a substantial decrease in the growth rate and cell yield (Fig. 3). During the later stages of growth, another activity could replenish the folate pool at a low rate in the latter strain. Altogether, these observations indicate that the absence

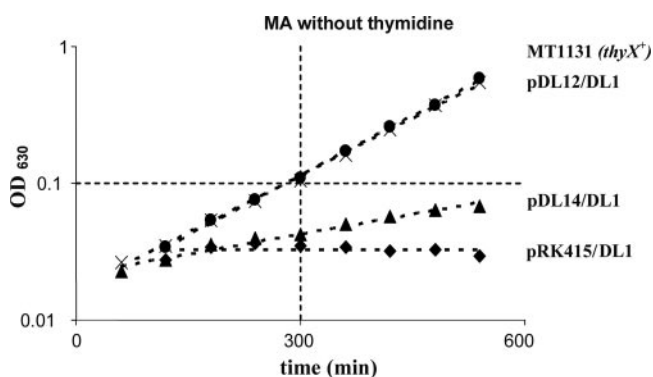


FIG. 3. Semilogarithmic growth curves for strains MT1131 (*thyX*<sup>+</sup>), DL1/pDL12 (*thyX/thyA*<sup>+</sup> *folA*<sup>+</sup>), DL1/pDL14 (*thyX/thyA*<sup>+</sup>), and DL1/pRK415 (*thyX*) on minimal medium MA without thymidine. OD<sub>630</sub>, optical density at 630 nm.

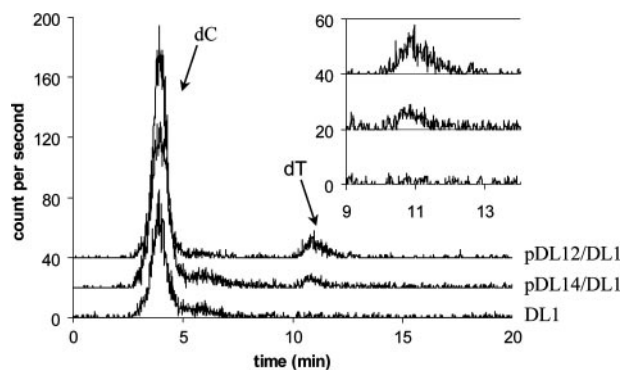


FIG. 4. Incorporation of tritium from [5,6-<sup>3</sup>H]uracil into dC and dT of DL1/pDL12 (*thyX::Spec<sup>r</sup>/thyA<sup>+</sup> folA<sup>+</sup>*), DL1/pDL14 (*thyX::Spec<sup>r</sup>/thyA<sup>+</sup>*), and DL1/pRK415/DL1 (*thyX*). Elution positions of dC and dT separated by isocratic high-performance liquid chromatography are indicated by arrows. Peak areas were quantified using Laura Light software and were normalized using the total DNA quantity.

of Fola limits cellular growth when ThyA is used for thymidylate synthesis.

**Fola and ThyA are functionally coupled during thymidylate synthesis.** To extend the results of the genetic complementation studies described above, we investigated whether *R. capsulatus* strain DL1 is indeed unable to synthesize thymidylate de novo. To this end, *R. capsulatus* chromosomal DNA was radioactively labeled with [5,6-<sup>3</sup>H]uracil and subsequently hydrolyzed by boiling in perchloric acid. The use of [5,6-<sup>3</sup>H]uracil results in labeling of dC at positions 5 and 6 of the pyrimidine ring and in labeling of dT at position 6 (the label from position 5 is lost during ThyX/ThyA catalysis). Free nucleobases were separated by high-pressure liquid chromatography using isocratic elution.

The negative control strain DL1/pRK415 produced only one peak corresponding to dC, whereas the same strain carrying pDL12 was able to synthesize de novo both dC (peak at 4 min) and dT (peak at 11 min) (Fig. 4). An intermediate result was

obtained for the DL1/pDL14 strain carrying only *R. sphaeroides thyA* on the plasmid. Although the latter strain was able to synthesize dT, its thymidylate synthesis capability was reduced by 70% compared to that of DL1/pDL12 (Fig. 4), as indicated by quantitative analyses that were normalized using the total DNA quantity.

**Extracellular THF and CH<sub>2</sub>THF inhibit the growth of *R. capsulatus*.** Some bacteria, like *Lactobacillus casei* (26) and cyanobacteria (14), actively transport folates, raising the possibility that the growth defect that we observed for *R. capsulatus* DL1/pDL14 (*thyA::Spec<sup>r</sup>/thyA<sup>+</sup>*) could be rescued by adding various folates to the growth medium. To test this hypothesis, we cultivated this strain in the presence of THF. We observed that, instead of restoring growth as we expected, addition of THF resulted in marked growth inhibition of this strain (Fig. 5). This phenotype could indicate either that strain DL1/pDL14 (*thyA::Spec<sup>r</sup>/thyA<sup>+</sup>*) is specifically inhibited by THF or, alternatively, that the observed phenotype is a characteristic feature of all *R. capsulatus* strains, including the wild type. Consequently, we cultivated the various strains used in this work in the presence and absence of various folates (Fig. 5). Strikingly, we found that addition of THF and CH<sub>2</sub>THF markedly inhibited, in a dose-dependent manner, the growth of all strains used. In this respect it noteworthy that in *Eukarya* extracellular folate compounds inhibit the folate-dependent enzymes through product inhibition (23).

Figure 5 also shows that the inhibitory effect of CH<sub>2</sub>THF on the growth of all *R. capsulatus* strains (including the wild type) was particularly pronounced, prompting us to measure the doubling times of these strains in the presence of 50 and 100 μg/ml of CH<sub>2</sub>THF (Table 2). The results obtained indicate that addition of CH<sub>2</sub>THF significantly increases the doubling times of the strains used. In addition to the product inhibition, increased CH<sub>2</sub>THF levels could specifically reverse the reaction catalyzed by serine hydroxymethyltransferase shown in Fig. 1 (25). The observed growth defect was transient, as prolonged cultivation even in the presence of 100 μg/ml of

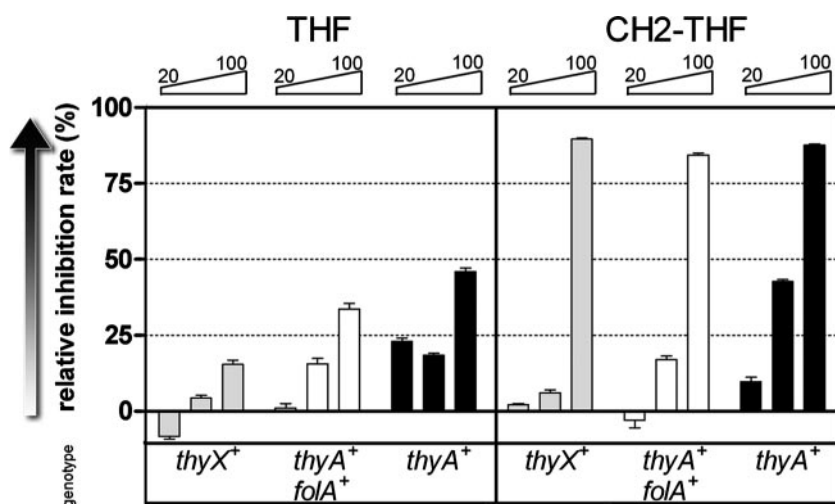


FIG. 5. Growth inhibition of *R. capsulatus* by folate derivatives. Growth was scored after 12 h of incubation in MPYE media to which various folate derivatives were added at concentrations of 20, 50, and 100 μg/ml. Gray bars, strain MT1131 (*thyX<sup>+</sup>*); open bars, strain DL1/pDL12 (*thyX::Spec<sup>r</sup>/thyA<sup>+</sup> folA<sup>+</sup>*); black bars, strain DL1/pDL14 (*thyX::Spec<sup>r</sup>/thyA<sup>+</sup>*).

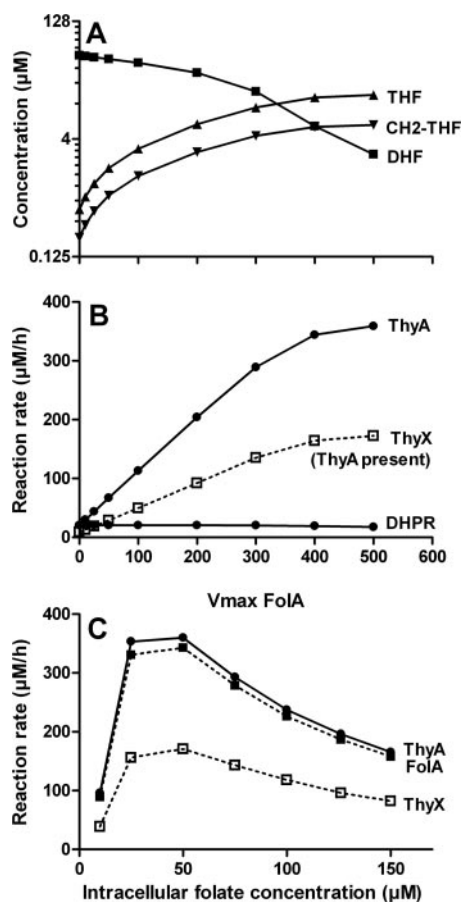


FIG. 6. (A) Effects of variation in the activity of Fola (DHFR) on selected steady-state concentrations of folate compounds in bacteria containing only ThyA. The model indicates that the decrease in the  $V_{\max}$  of Fola results in depletion of the THF pool in the cell. CH<sub>2</sub>-THF, methylenetetrahydrofolate. (B) Rate of thymidylate synthesis (for ThyA and ThyX) as a function of the Fola reaction rate. The graph also indicates that the dihydropteridine reductase (DHPR) rate is expected to be independent of the Fola rate. The simulations predicted that the relatively high ThyA activity renders the thymidylate synthesis activity dependent on Fola, even in the presence of ThyX. For these simulation results the model assumed that Fola and ThyA are upregulated, thus representing a rapidly dividing cell. (C) Effect of variation in the intracellular folate concentration on the activity of enzymes involved in folate metabolism. The model indicates that an increase in the folate pool results in inhibition of several enzymes, such as ThyA, Fola, and ThyX.

CH<sub>2</sub>-THF eventually resulted in growth similar to that observed in the absence of this compound (data not shown).

**Mathematical modeling of thymidylate metabolism in Bacteria.** The experimental results indicated that in thymidylate synthase ThyA-containing bacteria, which account for approximately two-thirds of the microbial species whose complete genome sequences are known, DHFR Fola is a key determinant of folate-dependent metabolic pathways (Fig. 1). Without efficient DHFR activity, ThyA catalysis would result in depletion of the THF pool in rapidly growing bacteria. Evidence for this can be seen in the simulation results shown in Fig. 6. For bacteria containing ThyA and Fola, the simulation shows that a decrease in the  $V_{\max}$  of Fola leads to a situation where the steady-state concentration of THF decreases progressively

(Fig. 6A). In agreement with this notion, recent isotope ratio-based profiling studies of *E. coli* folate pools in trimethoprim-treated cells have revealed decreased THF levels and increased DHF (or more oxidized compound) levels in comparison to the levels in nontreated cultures (17). Consequently, when the level of Fola activity is low, the intracellular concentration of the ThyA substrate CH<sub>2</sub>-THF declines rapidly to a very low level, resulting in a decreased rate of the ThyA reaction. This predicts that when Fola activity is inhibited, alternative ways to reduce DHF (8, 28) are not efficient enough to take on the DHF production in fast-growing ThyA-containing bacteria, explaining why these bacteria are highly sensitive to trimethoprim, a specific inhibitor of bacterial Fola. The mathematical simulation model also predicts that a  $V_{\max}$  of Fola of at least 120  $\mu\text{M/h}$  is necessary to support maximal thymidylate synthesis activity in bacteria containing ThyA and Fola (simulation not shown). The presence of ThyX and the absence of ThyA render the “DHF trap” nonfunctional, as ThyX proteins do not catalyze the net oxidation of folate molecules (Fig. 1). Low levels of DHFR activity could be enough to maintain the total folate pool sufficiently reduced to ensure correct ThyX-dependent DNA replication. Genome analyses also show that approximately 10% of bacterial species contain flavin-dependent thymidylate synthase ThyX and Fola. As our results indicate that thymidylate synthesis in ThyX-containing organisms is essentially Fola independent and cells are resistant to trimethoprim (data not shown), the predicted function of Fola in the latter subset of organisms is to increase the intracellular concentration of reduced folate derivatives for other folate-dependent pathways in RNA and protein metabolism.

Interestingly, the model also shows that in bacterial systems with ThyX, ThyA, and Fola (approximately 5% of bacterial systems, including those of *Mycobacterium* species), DNA synthesis is still dependent on Fola, and consequently, the cells are sensitive to Fola inhibitors (Fig. 6B). In this case, the “DHF trap” would reduce the intracellular concentration of the ThyX substrate, CH<sub>2</sub>-THF, to a very low level (Fig. 6B) that would be insufficient to sustain a high rate of ThyX-driven dTMP synthesis. In support of this prediction, we have observed that *E. coli* wild-type cells overexpressing ThyX proteins in *trans* are still sensitive to trimethoprim (data not shown). It is noteworthy that according to this model, folate concentrations cannot be maintained at a high level as this would result in growth inhibition (Table 2 and Fig. 6C).

In slowly growing bacteria, if Fola is absent (but ThyA and ThyX are both present), the “DHF trap” does not necessarily exist; in this case, alternative enzymes catalyzing THF formation, like *E. coli* FolM (8) or *T. thermophilus* dihydropteridine reductase (DH<sub>Ti</sub>) (28), could convert enough DHF to THF to keep the remainder of the folate cycle operating at near-normal rates (simulations not shown). This situation could naturally occur in some slowly growing bacteria, for instance, *Crocospaera watsonii* (doubling time, 35 h) and *Dehalococcoides* species (doubling time, 19 h), which contain ThyA and ThyA/ThyX, respectively, but appear to lack Fola and Tdk.

## DISCUSSION

In agreement with earlier biochemical and heterologous complementation studies, our studies have proven that flavin-



dependent ThyX proteins play an important cellular role in de novo thymidylate synthesis. We demonstrate here that the genetic complementation of an *R. capsulatus thyX::Spec<sup>r</sup>* strain requires not only *thyA* but also *folA*. We also show that the absence of *folA* in a *thyA*<sup>+</sup> strain results in a decrease in dTMP production, resulting in a DNA replication defect. This is presumably due to the fact that the DHF resulting from ThyA catalysis is not reduced at a sufficiently high rate by alternative folate reductases that must be present in ThyX-containing species, including *R. capsulatus* (19). We also show that the addition of THF and particularly CH<sub>2</sub>THF can markedly inhibit bacterial growth. To our knowledge, this growth inhibition, presumably resulting from product inhibition of the folate-dependent enzymes (Fig. 5 and 6C), has not been previously described.

Our mathematical model predicts that in *thyX*<sup>+</sup> strains even a low dihydrofolate-reducing activity is sufficient for thymidylate synthesis. This suggests that different oxidoreductases belonging to a large family of short-chain alcohol dehydrogenases that have been described as suppressors of *folA*, like FolM (8) and DH<sub>T<sub>r</sub></sub> (28), could act as promiscuous DHFRs. These enzymes possess trimethoprim-insensitive DHFR activity in vitro and can complement *folA* defects in *E. coli* when they are overexpressed from high-copy-number plasmids. We detected at least 19 sequences showing sequence similarity to genes encoding FolM and/or DH<sub>T<sub>r</sub></sub> in the *R. capsulatus* and *R. sphaeroides* genome sequences. Although we have not investigated the possible physiological roles of any of these dehydrogenases in bacterial folate metabolism, our experiments indicate that the alternative pathways for formation of reduced folates are not sufficient for the recycling of oxidized folates formed by *R. sphaeroides* ThyA in the absence of FolA activity during DNA replication. Note also that we cannot exclude the possibility that additional novel enzymes with THF-forming activities have not been discovered yet.

As ThyA proteins perform the only currently known cellular reaction catalyzing the net oxidation of THF, these enzymes function as a critical determinant of reduced folate levels. In this respect, mycobacteria that contain both *thyA* and *thyX* genes provide a particularly interesting case. Systematic transposon mutagenesis of *Mycobacterium bovis* BCG has indicated that *thyX* encodes essential functions even in the presence of *thyA*, whereas the accumulation of mutations in the *thyA* gene represents a mechanism of developing resistance to drugs targeting folate metabolism (24). Our simulations also predict that the presence of *thyX* in mycobacteria provides a molecular basis for resistance to antifolates targeting FolA. It is also noteworthy that the availability of distinct redox cofactors required by ThyA and ThyX proteins might differ during the different stages of DNA replication, suggesting additional differences in the coordination of cellular metabolism in *thyA*- and *thyX*-containing species.

In conclusion, our results demonstrated that ThyX proteins are required for de novo thymidylate synthesis. The thymidylate synthase ThyA and some enzymes involved in folate metabolism, like DHFRs and dihydropteroate synthases, are well-characterized therapeutic targets in anticancer and antimicrobial treatments (16). The essential role of the thymidylate synthase ThyX in bacterial survival and the presence of this protein in several pathogens therefore offer an attractive opportunity to design novel antibacterial drugs.

## ACKNOWLEDGMENTS

We thank F. Daldal for necessary *Rhodobacter* strains and plasmids that were required for this research.

This work was supported by a grant from the CNRS Programme Microbiologie Fondamentale (to U.L. and H.M.). H.M. also acknowledges financial support from the INSERM Bioavenir program and the Fondation Bettencourt Schuller. F.E. and D.L. received a fellowship from the French Ministry of Research. F.N. and M.R. acknowledge support from NIH grant 5RO1-CA105437 and NSF grants DMS-0109872 and DMS-061670.

## REFERENCES

- Agrawal, N., S. A. Lesley, P. Kuhn, and A. Kohen. 2004. Mechanistic studies of a flavin-dependent thymidylate synthase. *Biochemistry* **43**:10295–10301.
- Carreras, C. W., and D. V. Santi. 1995. The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.* **64**:721–762.
- D'Ari, L., and J. C. Rabinowitz. 1991. Purification, characterization, cloning, and amino acid sequence of the bifunctional enzyme 5,10-methylenetetrahydrofolate dehydrogenase/5,10-methylenetetrahydrofolate cyclohydrolase from *Escherichia coli*. *J. Biol. Chem.* **266**:23953–23958.
- Davis, J., T. J. Donohue, and S. Kaplan. 1988. Construction, characterization, and complementation of a Puf<sup>-</sup> mutant of *Rhodobacter sphaeroides*. *J. Bacteriol.* **170**:320–329.
- DeGregori, J., T. Kowalik, and J. R. Nevins. 1995. Cellular targets for activation by the E2F1 transcription factor include DNA synthesis- and G1/S-regulatory genes. *Mol. Cell. Biol.* **15**:4215–4224.
- Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* **77**:7347–7351.
- Dynes, J. L., and R. A. Firtel. 1989. Molecular complementation of a genetic marker in *Dictyostelium* using a genomic DNA library. *Proc. Natl. Acad. Sci. USA* **86**:7966–7970.
- Giladi, M., N. Altman-Price, I. Levin, L. Levy, and M. Mevarech. 2003. FolM, a new chromosomally encoded dihydrofolate reductase in *Escherichia coli*. *J. Bacteriol.* **185**:7015–7018.
- Giladi, M., G. Bitan-Banin, M. Mevarech, and R. Ortenberg. 2002. Genetic evidence for a novel thymidylate synthase in the halophilic archaeon *Halobacterium salinarum* and in *Campylobacter jejuni*. *FEMS Microbiol. Lett.* **216**:105–109.
- Graziani, S., J. Bernauer, S. Skouloubris, M. Graille, C. Z. Zhou, C. Marchand, P. Decottignies, H. van Tilbeurgh, H. Myllykallio, and U. Liebl. 2006. Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J. Biol. Chem.* **281**:24048–24057.
- Graziani, S., Y. Xia, J. R. Gurnon, J. L. Van Etten, D. Leduc, S. Skouloubris, H. Myllykallio, and U. Liebl. 2004. Functional analysis of FAD-dependent thymidylate synthase ThyX from *Paramecium bursaria* Chloroella virus-1. *J. Biol. Chem.* **279**:54340–54347.
- Griffin, J., C. Roshick, E. Iliffe-Lee, and G. McClarty. 2005. Catalytic mechanism of *Chlamydia trachomatis* flavin-dependent thymidylate synthase. *J. Biol. Chem.* **280**:5456–5467.
- Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* **70**:191–197.
- Klaus, S. M., E. R. Kunji, G. G. Bozzo, A. Noiriell, R. D. de la Garza, G. J. Basset, S. Ravel, F. Rebeille, J. F. Gregor III, and A. D. Hanson. 2005. Higher plant plastids and cyanobacteria have folate carriers related to those of trypanosomatids. *J. Biol. Chem.* **280**:38457–38463.
- Koch, H.-G., H. Myllykallio, F. Daldal, and M. Lee. 1998. Using genetics to explore cytochrome function and structure in *Rhodobacter*. *Methods Enzymol.* **297**:81–94.
- Kompis, I. M., K. Islam, and R. L. Then. 2005. DNA and RNA synthesis: antifolates. *Chem. Rev.* **105**:593–620.
- Lu, W., Y. K. Kwon, and J. D. Rabinowitz. 2007. Isotope ratio-based profiling of microbial folates. *J. Am. Soc. Mass Spectrom.* **18**:898–909.
- Marrs, B. 1974. Genetic recombination in *Rhodospseudomonas capsulata*. *Proc. Natl. Acad. Sci. USA* **71**:971–973.
- Myllykallio, H., D. Leduc, J. Filee, and U. Liebl. 2003. Life without dihydrofolate reductase FolA. *Trends Microbiol.* **11**:220–223.
- Myllykallio, H., G. Lipowski, D. Leduc, J. Filee, P. Forterre, and U. Liebl. 2002. An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**:105–107.
- Nijhout, H. F., M. C. Reed, P. Budu, and C. M. Ulrich. 2004. A mathematical model of the folate cycle: new insights into folate homeostasis. *J. Biol. Chem.* **279**:55008–55016.
- Prentki, P., and H. M. Krisch. 1984. In vitro insertional mutagenesis with a selectable DNA fragment. *Gene* **29**:303–313.
- Reed, M. C., H. F. Nijhout, M. L. Neuhouser, J. F. Gregory III, B. Shane, S. J. James, A. Boynton, and C. M. Ulrich. 2006. A mathematical model gives

- insights into nutritional and genetic aspects of folate-mediated one-carbon metabolism. *J. Nutr.* **136**:2653–2661.
24. **Rengarajan, J., C. M. Sassetti, V. Naroditskaya, A. Sloutsky, B. R. Bloom, and E. J. Rubin.** 2004. The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria. *Mol. Microbiol.* **53**:275–282.
25. **Schirch, V., S. Hopkins, E. Villar, and S. Angelaccio.** 1985. Serine hydroxymethyltransferase from *Escherichia coli*: purification and properties. *J. Bacteriol.* **163**:1–7.
26. **Shane, B., and E. L. Stokstad.** 1975. Transport and metabolism of folates by bacteria. *J. Biol. Chem.* **250**:2243–2253.
27. **Simon, R., U. Priefer, and A. Puhler.** 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* **1**:784–791.
- 27a. **Sistrom, W. R.** 1960. A requirement for sodium in the growth of *Rhodospseudomonas spheroides*. *J. Gen. Microbiol.* **22**:778–785.
28. **Wilquet, V., M. Van de Castele, D. Gigot, C. Legrain, and N. Glansdorff.** 2004. Dihydropteridine reductase as an alternative to dihydrofolate reductase for synthesis of tetrahydrofolate in *Thermus thermophilus*. *J. Bacteriol.* **186**:351–355.
29. **Zhong, J., S. Skouloubris, Q. Dai, H. Myllykallio, and A. G. Barbour.** 2006. Function and evolution of plasmid-borne genes for pyrimidine biosynthesis in *Borrelia* spp. *J. Bacteriol.* **188**:909–918.

## SUPPLEMENTAL DATA

*Kinetic parameter values used in the model*  
(*K<sub>m</sub> in μM, V<sub>max</sub> in μM/hr*)

Parameter	Literature	Model	Reference <sup>1</sup>	Organism
<b>DHFR</b>				
<i>K<sub>m</sub></i> ,DHF	0.97	1.0	(16)	<i>Escherichia coli</i>
<i>K<sub>m</sub></i> ,NADPH	2.52	2.5	(16)	<i>Escherichia coli</i>
<i>V<sub>max</sub></i>		500		
<b>ThyA</b>				
<i>K<sub>m</sub></i> ,dUMP	1.1-1.2	1.0	(4, 11, 13)	<i>Escherichia coli</i>
<i>K<sub>m</sub></i> ,5,10-CH <sub>2</sub> -THF	8.1-10	10	(11, 13)	<i>Escherichia coli</i>
<i>V<sub>max</sub></i>		2250		
<b>ThyX</b>				
K cat	0.35		(8)	<i>PBCV-1</i>
<i>K<sub>m</sub></i> ,dUMP	15	15	(8)	<i>PBCV-1</i>
<i>K<sub>m</sub></i> ,5,10-CH <sub>2</sub> -THF	20	20	(8)	<i>PBCV-1</i>
<i>V<sub>max</sub></i>		1200		
<b>K<sub>m</sub>NADPH</b>	40		(8)	<i>PBCV-1</i>
<b>DHFS – specific activity</b>	0.535			<i>Escherichia coli</i>
<b>K<sub>m</sub> dihydropteroate</b>	0.6			
<b>FoIM - specific activity (V<sub>max</sub>)</b>	0.083		(7)	<i>Escherichia coli</i>
<b>DHPR – specific activity</b>	5.2		(20)	<i>Thermus thermophilus</i>
<i>V<sub>max</sub></i>		225		
K <sub>m</sub> DHF		0.5		
K <sub>m</sub> NADPH		4.0		
<b>MTCH (positive direction is from 5,10-CH=THF to 10f-THF)</b>				
<i>K<sub>m</sub></i> ,5,10CH=THF	80	80	(14)	<i>Methylobacterium extorquens</i>
<i>Spec activity</i>	5,5		(2)	<i>Escherichia coli</i>
<i>V<sub>max</sub></i>		48000		
<b>PGT</b>				
<i>K<sub>m</sub></i> ,10f-THF	77.5 - 84.8	80	(10)	<i>Escherichia coli</i>
<i>K<sub>m</sub></i> ,GARP	8 - 12.2	10	(10)	<i>Escherichia coli</i>
<i>Spec activity</i>	11.6		(10)	<i>Escherichia coli</i>
<i>V<sub>max</sub></i>		4000		
<b>AICART</b>				
<i>K<sub>m</sub></i> ,10f-THF		5.9		
<i>K<sub>m</sub></i> ,AICARP		50		
<i>V<sub>max</sub></i>		10000		

<sup>1</sup> If the reference is not indicated, the values were collected from BRENDA database.

<b>FTDF</b>				
$K_{m10f-THF}$	49	50	(12)	<i>Escherichia coli</i>
<i>Spec Activity</i>	0.174		(12)	<i>Escherichia coli</i>
$V_{max}$		3000		
<b>SHMT (positive direction is from THF to 5,10-CH<sub>2</sub>-THF)</b>				
$K_{m,Ser}$	150-300	300	(5, 15)	<i>Escherichia coli</i>
$K_{m,THF}$	15-80	50	(5, 15)	<i>Escherichia coli</i>
<i>Spec Activity</i>	13.6		(15)	<i>Escherichia coli</i>
$V_{max}$		6000		
<b>MS</b>				
$K_{m,5mTHF}$	30-60	50	(17, 18)	<i>Escherichia coli</i>
$K_{m,Hcy}$	16-69	50	(18, 21)	<i>Escherichia coli</i>
<i>Spec Activity</i>	3.8		(6)	<i>Escherichia coli</i>
		5000		
<b>MTD (5,10-CH<sub>2</sub>-THF → 5,10-CH=THF)</b>				
<i>Spec Activity</i>	200		(2)	<i>Escherichia coli</i>
$V_{max}$		100000		
$K_{m,5,10-CH_2-THF}$	26-45	30	(1, 3)	<i>Streptococcus faecium</i> ; <i>Salmonella thyphimurium</i>
$K_{m,NADP+}$	21-92		(1, 3)	<i>Streptococcus faecium</i> ; <i>Salmonella thyphimurium</i>
<b>MTHFR (5,10-CH<sub>2</sub>-THF → 5mTHF)</b>				
$K_{m,5,10-CH_2-THF}$	0.5-3.9	50	(9, 19)	<i>Escherichia coli</i>
$K_{m,NADH}$	17-20	50	(9, 19)	<i>Escherichia coli</i>
$V_{max}$		15		

## References

1. **Albrecht, A. M., F. K. Pearce, and D. J. Hutchison.** 1968. Methylene tetrahydrofolate dehydrogenase of the amethopterin-resistant strain *Streptococcus faecium* var. *durans* A and its repressibility by serine. *J Bacteriol* **95**:1779-89.
2. **D'Ari, L., and J. C. Rabinowitz.** 1991. Purification, characterization, cloning, and amino acid sequence of the bifunctional enzyme 5,10-methylene tetrahydrofolate dehydrogenase/5,10-methenyl tetrahydrofolate cyclohydrolase from *Escherichia coli*. *J Biol Chem* **266**:23953-8.
3. **Dalal, F. R., and J. S. Gots.** 1967. Purification of 5,10-methylene tetrahydrofolate dehydrogenase from *Salmonella typhimurium* and its inhibition by purine nucleotides. *J Biol Chem* **242**:3636-40.
4. **Fantz, C., D. Shaw, W. Jennings, A. Forsthoefel, M. Kitchens, J. Phan, W. Minor, L. Lebioda, F. G. Berger, and H. T. Spencer.** 2000. Drug-resistant variants of *Escherichia coli* thymidylate synthase: effects of substitutions at Pro-254. *Mol Pharmacol* **57**:359-66.
5. **Fu, T. F., E. S. Boja, M. K. Safo, and V. Schirch.** 2003. Role of proline residues in the folding of serine hydroxymethyltransferase. *J Biol Chem* **278**:31088-94.
6. **Fujii, K., and F. M. Huennekens.** 1974. Activation of methionine synthetase by a reduced triphosphopyridine nucleotide-dependent flavoprotein system. *J Biol Chem* **249**:6745-53.

7. **Giladi, M., N. Altman-Price, I. Levin, L. Levy, and M. Mevarech.** 2003. FolM, a new chromosomally encoded dihydrofolate reductase in *Escherichia coli*. *J Bacteriol* **185**:7015-8.
8. **Graziani, S., Y. Xia, J. R. Gurnon, J. L. Van Etten, D. Leduc, S. Skouloubris, H. Myllykallio, and U. Liebl.** 2004. Functional analysis of FAD-dependent thymidylate synthase ThyX from *Paramecium bursaria* Chlorella virus-1. *J Biol Chem* **279**:54340-7.
9. **Guenther, B. D., C. A. Sheppard, P. Tran, R. Rozen, R. G. Matthews, and M. L. Ludwig.** 1999. The structure and properties of methylenetetrahydrofolate reductase from *Escherichia coli* suggest how folate ameliorates human hyperhomocysteinemia. *Nat Struct Biol* **6**:359-65.
10. **Inglese, J., D. L. Johnson, A. Shiau, J. M. Smith, and S. J. Benkovic.** 1990. Subcloning, characterization, and affinity labeling of *Escherichia coli* glycinamide ribonucleotide transformylase. *Biochemistry* **29**:1436-43.
11. **Mahdavian, E., H. T. Spencer, and R. B. Dunlap.** 1999. Kinetic studies on drug-resistant variants of *Escherichia coli* thymidylate synthase: functional effects of amino acid substitutions at residue 4. *Arch Biochem Biophys* **368**:257-64.
12. **Nagy, P. L., A. Marolewski, S. J. Benkovic, and H. Zalkin.** 1995. Formyltetrahydrofolate hydrolase, a regulatory enzyme that functions to balance pools of tetrahydrofolate and one-carbon tetrahydrofolate adducts in *Escherichia coli*. *J Bacteriol* **177**:1292-8.
13. **Phan, J., E. Mahdavian, M. C. Nivens, W. Minor, S. Berger, H. T. Spencer, R. B. Dunlap, and L. Lebioda.** 2000. Catalytic cysteine of thymidylate synthase is activated upon substrate binding. *Biochemistry* **39**:6969-78.

14. **Pomper, B. K., J. A. Vorholt, L. Chistoserdova, M. E. Lidstrom, and R. K. Thauer.** 1999. A methenyl tetrahydromethanopterin cyclohydrolase and a methenyl tetrahydrofolate cyclohydrolase in *Methylobacterium extorquens* AM1. *Eur J Biochem* **261**:475-80.
15. **Schirch, V., S. Hopkins, E. Villar, and S. Angelaccio.** 1985. Serine hydroxymethyltransferase from *Escherichia coli*: purification and properties. *J Bacteriol* **163**:1-7.
16. **Shaw, D., J. D. Odom, and R. B. Dunlap.** 1999. High expression and steady-state kinetic characterization of methionine site-directed mutants of *Escherichia coli* methionyl- and selenomethionyl-dihydrofolate reductase. *Biochim Biophys Acta* **1429**:401-10.
17. **Taylor, R., and H. Weissbach.** 1973. The Enzymes, 3rd Ed. (Boyer, P.D., ed.) **12**:121-165.
18. **Taylor, R., and H. Weissbach.** 1971. *Methods Enzymol* **17B**:379-388.
19. **Trimmer, E. E., D. P. Ballou, M. L. Ludwig, and R. G. Matthews.** 2001. Folate activation and catalysis in methylenetetrahydrofolate reductase from *Escherichia coli*: roles for aspartate 120 and glutamate 28. *Biochemistry* **40**:6216-26.
20. **Wilquet, V., M. Van de Castele, D. Gigot, C. Legrain, and N. Glansdorff.** 2004. Dihydropteridine reductase as an alternative to dihydrofolate reductase for synthesis of tetrahydrofolate in *Thermus thermophilus*. *J Bacteriol* **186**:351-5.
21. **Zhou, Z. S., A. E. Smith, and R. G. Matthews.** 2000. L-Selenohomocysteine: one-step synthesis from L-selenomethionine and kinetic analysis as substrate for methionine synthases. *Bioorg Med Chem Lett* **10**:2471-5.

### III. Discussion

#### A. *ThyX est une thymidylate synthase in vivo*

Les travaux présentés dans l'article I démontrent pour la première fois *in vivo* que ThyX est une thymidylate synthase dont l'activité est essentielle en absence de thymidine exogène. Conformément aux prédictions, l'activité de ThyX est indépendante de celle de FoliA. Le défaut de croissance de la souche pDL14/DL1 ( $\Delta thyX$ ,  $tdk^+$ ,  $thyA^+$ ,  $folA^-$ ) en absence de thymidine suggère que le phénomène de « thymineless death » semble également présent chez les organismes utilisant ThyX. Une information supplémentaire est apportée par l'observation microscopique de la morphologie des cellules de *R. capsulatus* présentée dans la Figure 11 où nous pouvons voir qu'en absence de thymidine cette souche présente un phénotype de filamentation caractéristique d'un défaut de réplication et division. La souche exprimant ThyA et FoliA présente un phénotype comparable à pDL7/DL1 ( $thyX^+/thyX::spec^R$ ), indiquant que le couple ThyA/FoliA remplace efficacement ThyX.

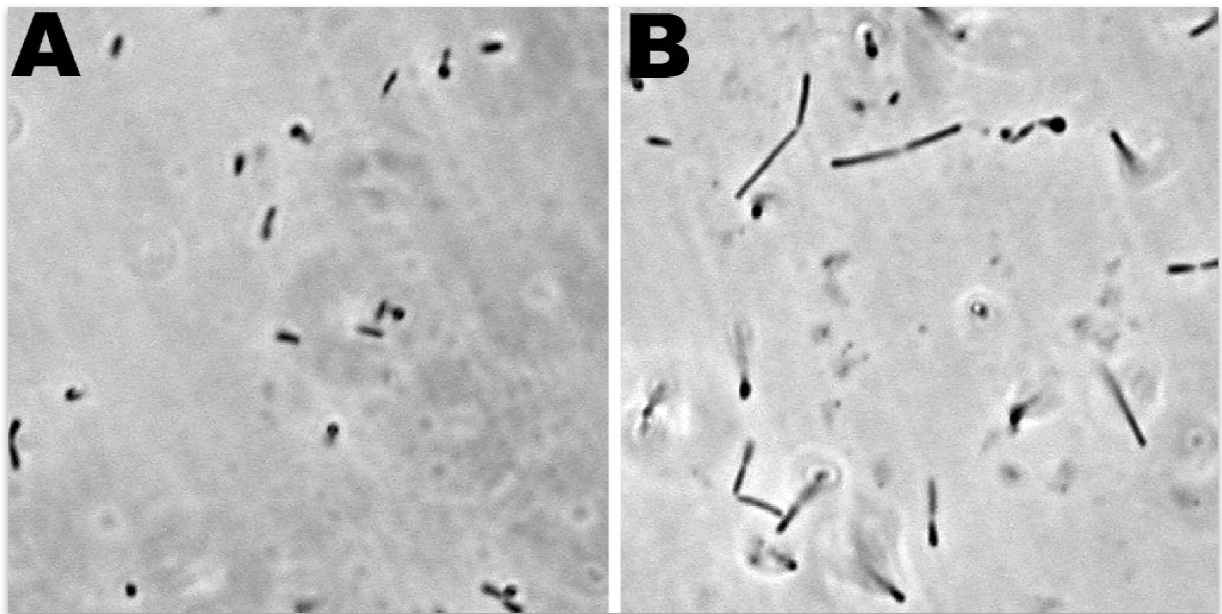


Figure 11 : Modification de la morphologie de *R. capsulatus* en limitation en dTMP. Les souches pDL7/DL1 ( $thyX^+/thyX::spec^R$ ) [A] et pDL14/DL1 ( $thyA^+/thyX::spec^R$ ) [B] ont été cultivées en milieu MPYE en absence de thymidine exogène.



### ***B. ThyA et FoaA sont fonctionnellement couplées***

L'expérience de complémentation avec le gène *thyA* seul nous a appris que l'absence de la DHFR FoaA conduit à une diminution de la synthèse de dTMP. Ce résultat nous indique que ThyA et FoaA sont des enzymes fonctionnellement couplées, ainsi la synthèse de thymidylate catalysée par ThyA est sous la directe dépendance de l'activité de recyclage du DHF de FoaA. Nous avons imaginé pouvoir compenser la perte de viabilité de la souche n'exprimant que ThyA en lui apportant une source de THF exogène. Au contraire, nous avons observé une inhibition de croissance à partir de 50µg/ml de THF. Il s'est avéré que l'ajout de MTHF a un effet encore plus drastique, et ceci indépendamment du contexte génétique. L'utilisation du modèle mathématique nous a permis de proposer que l'augmentation artificielle du pool de folate doit conduire à l'inhibition transitoire des enzymes impliquées dans ce métabolisme, expliquant le ralentissement de croissance.

### ***C. Les cas atypiques***

Les résultats de cette étude laissent penser que deux types d'organismes doivent être présents dans le monde bactérien : ceux utilisant le couple ThyA/FoaA (comme *E. coli* ou *Bacillus subtilis*) et ceux utilisant ThyX ne possédant pas de gène *foaA* (comme *R. capsulatus* ou *H. pylori*). En se basant sur l'étude de 395 génomes bactériens séquencés, nous obtenons la répartition présentée dans la Figure 12 : 60% des organismes étudiés utilisent le couple ThyA/FoaA contre un peu moins de 30% pour ThyX.

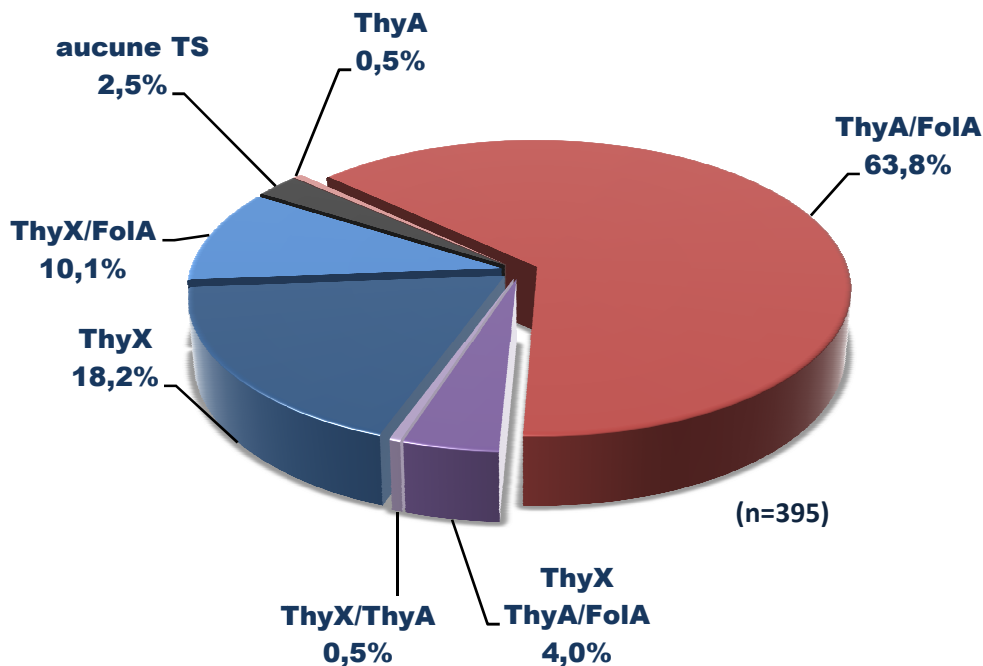


Figure 12 : Distribution de la DHFR FoIA et des thymidylate synthases ThyA et ThyX chez les bactéries. Les données sont extraites de l'étude présentée dans l'article II.

#### a. Absence de FoIA en présence de ThyA

Parmi les quelques pourcents restants, on observe : quatre organismes qui possèdent le gène *thyA* mais pas *folA* (Figure 12). Dans ce cas, la conversion du DHF en THF est sous la seule dépendance d'enzymes alternatives possédant une faible activité DHFR, comme FoIM d'*E. coli* (Giladi et coll., 2003) ou la dihydropteridine reductase DHTt de *Thermus thermophilus* (Wilquet et coll., 2004). Or nous avons vu que dans le cas de *R. capsulatus*, ces activités n'étaient pas suffisantes pour permettre la correcte réplication de l'ADN. En revanche, le modèle mathématique nous indique que si la vitesse de croissance est suffisamment lente, ces activités DHFR accessoires permettraient de maintenir le pool de folates à un niveau suffisant. Il est intéressant de noter que parmi ces organismes, on retrouve deux bactéries à croissance lente en conditions de laboratoire : *Crocospaera watsonii* et *Dehalococcoides sp.* qui ont des temps de doublement de 35 et 19 heures respectivement. *Dehalococcoides sp.* a également la particularité de posséder à la fois *thyA* et *thyX*.

### b. Présence de ThyX, ThyA et FoaA

Environ 4% des bactéries de l'échantillon possèdent *thyX*, *thyA*, et *foaA* comme les *Corynebacteria* et les *Mycobacteria* (Figure 12). Dans ce cas, le modèle prédit que la synthèse de thymidylate est sous la dépendance de FoaA malgré la présence de ThyX. En effet, si FoaA est inhibée, le DHF formé par ThyA restera « piégé » ce qui induira une diminution du THF et du CH<sub>2</sub>THF disponibles ; ceci aboutira finalement à l'inhibition des deux thymidylate synthases.

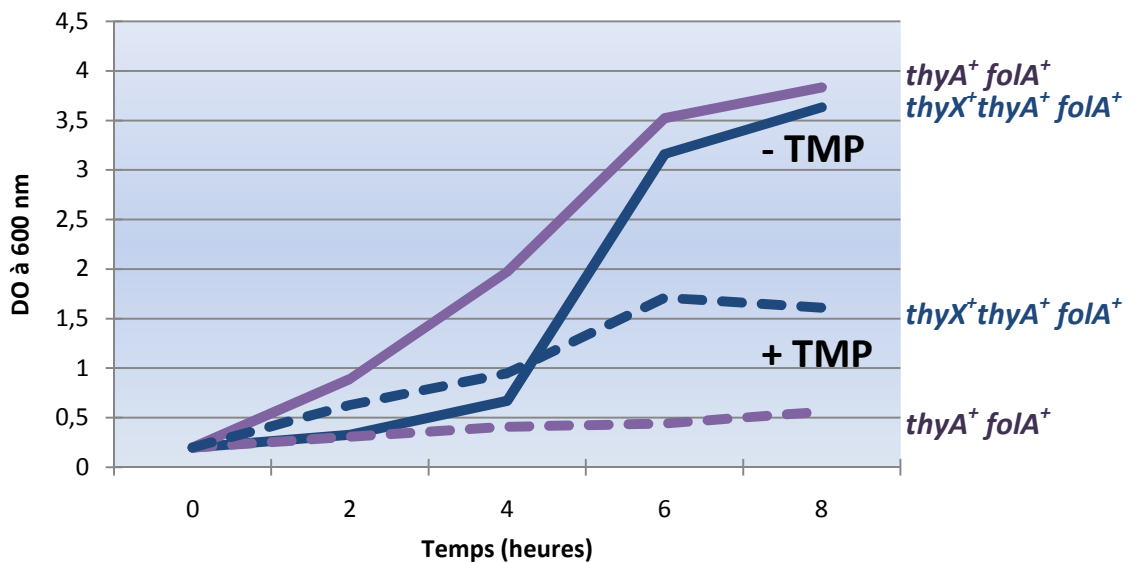


Figure 13 : Culture d'*E. coli* BL21 (*thyA*<sup>+</sup>, *foaA*<sup>+</sup>) portant (courbes violettes) ou non (lignes bleues) le plasmide pVEX (*thyX*<sup>+</sup>). Les souches ont été cultivées en absence (lignes pleines) et présence (lignes pointillées) de 10 µg/ml de triméthoprimine (TMP).

J'ai recréé ce système artificiellement en utilisant une souche d'*E. coli* BL21 (F- *dcm ompT hsdS(rB- mB-) gal thyA*<sup>+</sup> *foaA*<sup>+</sup>) exprimant ou non le gène *thyX* issu du virus PBCV-1 en *trans* à partir d'un promoteur *tac* inductible à l'IPTG d'un plasmide dérivé du pGEX-2T (Amersham Biosciences). On observe sur la Figure 13 que la croissance de la souche sauvage est complètement inhibée à une concentration de 10 µg/ml de triméthoprimine (TMP) [inhibiteur de FoaA]. L'expression de ThyX en présence de TMP confère une faible résistance, mais ne restaure pas la viabilité de la souche conformément à la prédiction. Dans quel cas alors la présence des deux systèmes de synthèse du thymidylate pourrait-elle conférer un avantage ? En cas d'inhibition de FoaA, ThyX pourrait servir de « voie de secours » de synthèse du dTMP si et seulement si l'activité de ThyA est diminuée. Cette

diminution peut résulter soit de l'accumulation de mutation(s) (Rengarajan *et coll.*, 2004) ; soit d'une régulation transcriptionnelle ou traductionnelle de *thyA* (répression) et/ou de *thyX* (induction). Des travaux concernant *Corynebacterium glutamicum* indiqueraient qu'en présence de TMP l'activité de ThyA soit diminuée et que l'expression du gène *thyX* soit positivement régulée (Cho *et coll.*, 2006). Il est également envisageable que ces deux gènes soient régulés différemment au cours de la vie de la bactérie en réponse par exemple à la disponibilité de cofacteurs comme le NADPH.

### c. Présence de ThyX et Fola

Un autre point surprenant est que près de 35% des organismes utilisant ThyX comme seule thymidylate synthase possèdent un gène *folA* (Figure 12). Les données expérimentales et les simulations mathématiques indiquent que la synthèse de dTMP par ThyX est indépendante de Fola. En aucun cas, la présence de Fola dans un organisme ThyX ne peut présenter un désavantage pour la cellule. On peut penser que la présence de Fola permet d'augmenter la concentration en dérivés réduits du THF pour les autres voies métaboliques dépendante des folates. Cependant, si on considère une force de sélection neutre, la présence de Fola peut être également la signature d'une acquisition récente de ce gène ou de la perte du gène *thyA*.

## Chapitre II. Thymidylate synthases et dynamique des génomes procaryotes

### I. Problématique

Nous avons vu dans chapitre précédent que la nature des différences de mécanismes réactionnels entre ThyA et ThyX avait des répercussions sur le métabolisme connexe des folates. Les travaux présentés dans l'article II montrent en quoi les différences d'efficacité au niveau de l'étape de réduction du groupement méthylène en méthyle entre les deux enzymes ont indirectement des répercussions sur la dynamique des génomes procaryotes.

Avant tout, je souhaite replacer ce travail dans son contexte original. Nous avons vu dans le chapitre I que la surexpression de ThyX dans une souche d'*E. coli*  $\Delta thyA$  conférait une résistance à l'anti-folate, TMP. Nous avons alors voulu tester si l'expression de *thyX* au niveau physiologique de *thyA* (sous le contrôle du promoteur de *thyA*) conduirait également à un phénotype de résistance au TMP. Dans ce but, j'ai construit une souche d'*E. coli* dans laquelle j'ai remplacé le gène *thyA* par l'ORF de *thyX* de PBCV-1 (codant la protéine ThyX la plus active décrite à ce jour), le remplacement a été marqué par une cassette de résistance à la kanamycine (voir Matériel et Méthodes de l'article II). J'ai alors réalisé des tests de sensibilité au TMP sur des cultures liquides en milieu minimum dépourvu de thymidine. Afin de prévenir tout artéfact dû à la présence résiduelle de thymine, j'ai décidé de réaliser la pré-culture en milieu minimum sans thymidine. Cependant aucune cultureensemencée à partir de cette pré-culture n'a permis la croissance bactérienne en présence ou absence de TMP. C'est l'observation de la morphologie des cellules au microscope qui a permis de comprendre ce phénomène. En effet, la majorité des cellules issues de la pré-culture se présentait sous la forme de longs filaments, ce qui semblait indiquer que la réplication et/ou la division cellulaire était altérée(s). Comme nous l'avons vu dans le chapitre I, le remplacement de *thyX* par le couple *thyA/foIA* avait restauré un phénotype sauvage ; alors qu'ici la réciproque conduisait vraisemblablement à une « thymineless death » (TLD). Nous avons alors mis en parallèle cette observation et les données biochimiques disponibles. Ces dernières indiquaient clairement que l'efficacité catalytique de ThyX

était limitée directement ou indirectement par l'étape de réduction de la flavine par le NADP(H). Ce phénomène s'illustre par la faible activité de ThyX [ $\sim 3\text{-}5\%$ ] *in vitro* comparativement à ThyA [100%] (voir table 1 de l'article II).

J'ai pu montrer, par des approches génétiques, que les différences mécanistiques entre les deux thymidylate synthases influencent, non seulement, l'efficacité catalytique de la réaction, mais aussi la physiologie de la cellule en limitant la vitesse de réplication de l'ADN. De plus, une analyse globale sur plus de 400 génomes procaryotes m'a permis de mettre en évidence que la synthèse de thymidylate est un facteur limitant de l'expansion des génomes procaryotes. L'ensemble de ce travail a également permis de porter un nouveau regard sur la distribution apparemment aléatoire de ThyA et de ThyX chez les Procaryotes (Figure 14).

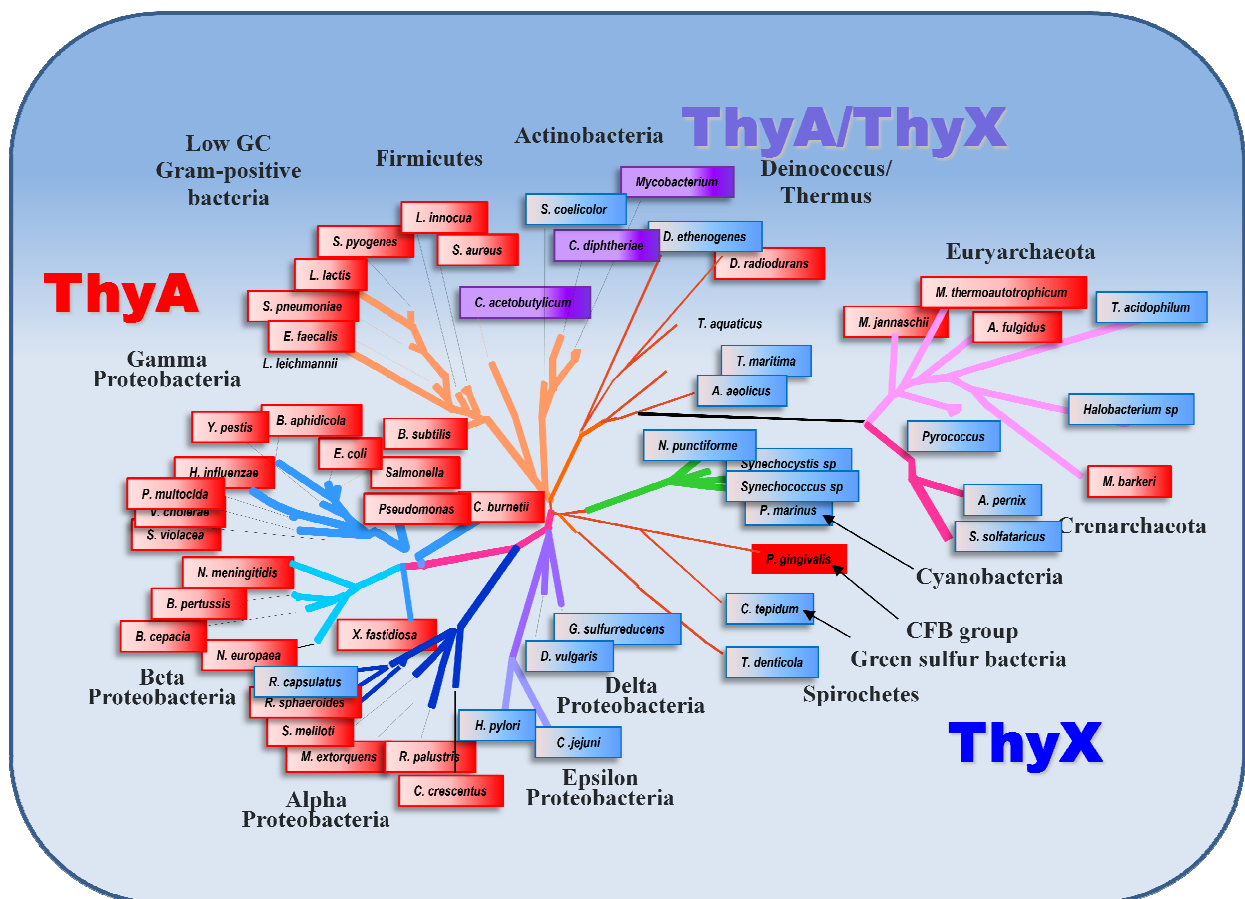


Figure 14 : Distributions phylogénétiques de *thyX* et *thyA* dans la plupart des groupes procaryotes [d'après (Leduc et coll., 2004b)].

**II. Article II : « Flavin Dependent Thymidylate Synthase X Limits Chromosomal DNA Replication »**

**Frédéric Escartin, Stéphane Skouloubris, Ursula Liebl, Hannu Myllykallio**

Article accepté le 6 mai 2008 dans  
*Proceedings of the National Academy of Sciences*

# Flavin-dependent thymidylate synthase X limits chromosomal DNA replication

Frédéric Escartin\*<sup>†‡</sup>, Stéphane Skouloubris<sup>‡</sup>, Ursula Liebl\*<sup>†</sup>, and Hannu Myllykallio\*<sup>§</sup>

\*Laboratoire d'Optique et Biosciences, Centre National de la Recherche Scientifique, École Polytechnique, F-91128 Palaiseau, France; <sup>†</sup>Institut National de la Santé et de la Recherche Médicale, Unité 696, F-91128 Palaiseau, France; and <sup>‡</sup>Université Paris-Sud 11, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 8621, Institut de Génétique et de Microbiologie, F-91405 Orsay, France

Edited by Robert M. Stroud, University of California, San Francisco, CA, and approved May 5, 2008 (received for review February 11, 2008)

We have investigated the hitherto unexplored possibility that differences in the catalytic efficiencies of thymidylate synthases ThyX and ThyA, enzymes that produce the essential DNA precursor dTMP, have influenced prokaryotic genome evolution. We demonstrate that DNA replication speed in bacteria and archaea that contain the low-activity ThyX enzyme is up to 10-fold decreased compared with species that contain the catalytically more efficient ThyA. Our statistical studies of >400 genomes indicated that ThyA proteins are preferred for the replication of large genomes, providing further evidence that the thymidylate metabolism is limiting expansion of prokaryotic genomes. Because both ThyX and ThyA participate in frequent reciprocal gene replacement events, our observations indicate that the bacterial metabolism continues to modulate the size and composition of prokaryotic genomes. We also propose that the increased kinetic efficiency of thymidylate synthesis has contributed to extending the prokaryotic evolutionary potential.

genome evolution | gene replacement | nucleotide metabolism | PBCV-1 | thymidylate synthesis

Contrarily to the deoxyribonucleotides dATP, dCTP, and dGTP, which can be produced directly by ribonucleotide reductase, *de novo* synthesis of dTTP requires the formation of thymidylate (dTMP). This essential DNA precursor is produced by the methylation of dUMP by thymidylate synthase. The mechanisms for *de novo* synthesis of DNA precursors were thought to be conserved in all free-living prokaryotes and eukaryotes. However, we identified a second family of thymidylate synthases [ThyX, also known as flavin-dependent thymidylate synthase (FDTS)] that is evolutionarily unrelated to the canonical thymidylate synthases ThyA. Approximately 30% of microbial species (based on completed genome sequences) depend on FDTS ThyX (EC 2.1.1.148) (1), whereas the canonical dTMP-forming enzyme ThyA (EC 2.1.1.45) is present in ≈70% of microorganisms. The key difference in ThyA and ThyX catalysis is related to the reductive mechanisms used for the reduction of the methylene group that serves as carbon source in the reaction. During ThyX catalysis, NAD(P)H is used as reductant, whereas ThyA uses methylenetetrahydrofolate (MTHF) not only as a one-carbon donor, but also as a source of reductive power (reviewed in ref. 2). Consequently, ThyX catalysis leads to the production of reduced tetrahydrofolate (THF), whereas ThyA produces oxidized dihydrofolate. These mechanistic differences of the two thymidylate synthases have clearly raised the possibility that the two enzymes might not be fully interchangeable *in vivo* (3).

Several arguments suggest that the evolutionary trajectories of ThyX and ThyA proteins differ. For instance, ThyA proteins form a highly conserved protein family, whereas ThyX sequences are much more divergent (1, 2). Notably, the high level of sequence conservation of ThyA proteins does not result from functional constraints because extensive mutagenesis studies performed on ThyA proteins identified only five critical residues for catalytic activity (4). Moreover, the sporadic and almost

**Table 1. Catalytic parameters of PBCV-1 ThyX and *Lactobacillus casei* ThyA**

Parameter*	PBCV-1 ThyX	<i>L. casei</i> ThyA
Reductant	NADPH/FAD	THF
$k_{cat}$ , $s^{-1}$	0.35 (10%)	3.6 (100%)
$K_m$ , mM		
dUMP	15	5
MTHF	20	10
NADPH	43	
$k_{cat}/K_m^{dUMP}$ , $s^{-1}\cdot\mu M^{-1}$	0.023 (3.2%)	0.72 (100%)
$k_{cat}/K_m^{MTHF}$ , $s^{-1}\cdot\mu M^{-1}$	0.0175 (4.8%)	0.36 (100%)

\*Data for ThyX and ThyA enzymes are extracted from Graziani *et al.* (8) and Liu and Santi (31), respectively.

mutually exclusive phylogenetic distribution patterns of *thyX* and *thyA* are indicative of frequent lateral gene transfer and/or nonorthologous gene displacement events (1, 5). Environmental factors might also influence thymidylate synthase utilization because our sequence-similarity searches indicated that *thyX* genes are over-represented in the genomes of (hyper)thermophilic, microaerophilic, and anaerobic microorganisms (6). However, *thyX* can also be found in a number of mesophilic and aerobic organisms suggesting that these “environmental” factors are not sufficient to explain the complex distribution patterns of the two thymidylate synthases.

Using experimental and statistical analyses, we have addressed why these two analogous enzyme families appeared, and why both thymidylate synthase families have been maintained in current-day organisms.

## Results and Discussion

Systematic activity measurements under aerobic and anaerobic conditions using both viral and bacterial ThyX proteins have suggested that flavin adenine dinucleotide (FAD) reduction by NAD(P)H or conformational changes associated with this reductive step limit the catalytic efficiency of ThyX catalysis *in vitro* compared with ThyA (7–11). This point is illustrated in Table 1 which shows that the catalytic efficiency of the viral ThyX protein is 3–5% of that which has been described for the *Lactobacillus casei* ThyA enzyme. This observation raised the unexplored possibility that intrinsic biochemical properties of ThyA and ThyX enzymes have influenced thymidylate synthase phylogeny and possibly the dynamics of prokaryotic genome

Author contributions: F.E., S.S., U.L., and H.M. designed research; F.E. performed research; F.E., U.L., and H.M. analyzed data; and F.E., S.S., U.L., and H.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>§</sup>To whom correspondence should be addressed. Email: hannu.myllykallio@igmors.u-psud.fr.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0801356105/DCSupplemental](http://www.pnas.org/cgi/content/full/0801356105/DCSupplemental).

© 2008 by The National Academy of Sciences of the USA







strikingly large, AT-rich genomes suggests that alternative strategies to enlarge prokaryotic genomes exist.

In conclusion, our combined experimental results and statistical analyses have indicated that metabolic enzymes can shape prokaryotic genomes. Because the work reported here indicates that ThyX enzymes limit DNA replication kinetics *in vivo*, our observations suggest that the increased catalytic efficiency of ThyA enzymes was necessary to increase the prokaryotic evolutionary potential.

## Materials and Methods

**Bacterial Strains and Growth Conditions.** Bacterial strains used in this study are listed in Table S2. *E. coli* cells were grown at 37°C in either thymidine-containing LB medium or thymidine-free M9 minimal medium supplemented with 0.1 mM CaCl<sub>2</sub>, 2 mM MgSO<sub>4</sub>, 0.2% casamino acids (Difco) or thymidine-free L+ medium (18). When necessary, thymidine (50 μg/ml) and isopropyl β-D-thiogalactopyranoside (1 mM) were added to the growth medium. *H. pylori* 26695 was grown on either blood agar base two (Oxoid) plates supplemented with 10% defibrinated horse blood (Oxoid) or in liquid culture in brain-heart infusion broth (Oxoid) supplemented with 10% FBS (Gibco BRL). An antibiotics–fungicide mix consisting of vancomycin (final 10 μg/ml), polymyxin B (2.5 units per liter), trimethoprim (5 μg/ml), and fungizone (2.5 μg/ml) was added. Plates were incubated at 37°C in a microaerobic atmosphere in jars by using CampyGen (Oxoid). Liquid cultures were shaken at 175 rpm.

**Molecular Biology Techniques.** The replacement of the complete *thyA* gene of *E. coli* strain MG1655 by the *Paramecium bursaria* *Chlorella* virus 1 (PBCV-1) *thyX* gene, encoding the most catalytically active ThyX protein isolated to date, was performed as follows. First, *thyX* and 500-bp DNA fragments flanking *thyA* were amplified by PCR (primers used are indicated in Table S1). The kanamycin cassette *aphA-3* was obtained by *Sma*I digestion of pUC18K2. The deletion construct (Fig. 1A) was assembled in two sequential PCRs. The obtained PCR product was introduced into the chromosome of *E. coli* by the use of the lambda-red recombination system (19, 20). This system uses the thermosensitive pKOBEG plasmid that carries the λ phage *red $\gamma$  $\beta$  $\alpha$*  operon under the control of a pBAD promoter. Transformants were selected with kanamycin in the presence of thymidine (replacement efficiency ≈ 7.10<sup>3</sup> CFU/μg of linear DNA). The *thyA*-deleted strain, FE013, was constructed as above without insertion of *thyX*. Both constructs were confirmed by PCR and DNA sequencing (data not shown).

**Fluorescence Microscopy.** *E. coli* strains MG1655 and FE010 were diluted in thymidine-deprived L+ medium to an OD<sub>600</sub> ≈ 0.2–0.3 and incubated at 37°C with agitation. One-milliliter culture aliquots were removed and stained as described (21). Briefly, cell membranes were stained with *N*-(3-triethylammonium-propyl)-4-(6-(4-(diethylamino)phenyl)hexatrienyl)pyridinium dibromide (FM 4-64) at 0.01 μg.μl<sup>-1</sup>, and the nucleoid was stained with 4,6-

diamidino-2-phenylindole dihydrochloride (DAPI) at 2 μg.ml<sup>-1</sup>. FM 4-64 stains lipid membranes and emits a red fluorescence (excitation/emission ≈ 515/640 nm); DAPI is a nucleoid stain with blue fluorescence (excitation/emission ≈ 350/470 nm). The stained cells were observed by using a Leica DM RXA microscope. Images were captured with a CCD camera 5 MHz Micromax 1300Y (Roper Instruments). The final reconstructed images were obtained by deconvoluting Z-series with Metamorph software (Universal Imaging).

**Runout Replication Assay.** *E. coli* strains were grown in LB medium for 12–14 h and diluted 10-fold by using M9 minimal medium. At the beginning of the exponential phase, rifampicin (50 μg/ml final; Sigma), uridine (50 μg/ml final; Sigma), and 3 μCi of [*methyl*-<sup>3</sup>H]-thymidine (84 Ci/mmol; Amersham) were added. Incorporation of radioactive thymidine was measured by scintillation counting after trichloroacetic acid precipitation of whole cells.

**Determination of Ori/Ter Ratio and Replication Period (C-Period).** For the experimental quantification of the relative ratios of *H. pylori* origin and terminus regions [previously identified by using bioinformatic analyses (22)], DNA was extracted from midexponential cultures (OD<sub>600</sub> ≈ 0.8) by using QIAamp DNA Mini kit (Qiagen). The primers used (oEF102, oEF103, oEF106, and oEF107) for the quantification reactions are indicated in Table S1. Quantitative PCR was performed in a MiniOpticon real-time PCR detection system (Bio-Rad) by using SYBR green PCR MasterMix Q-PCR (Bio-Rad). Observed cycle thresholds and PCR efficiencies allowed determination of a calibrator-normalized relative quantification of the replication origin over terminus for each sample by using Opticon Monitor Software version 3.1.32 (Bio-Rad). The replication time (C in the following equation) was calculated from the origin/terminus ratio by using the formula O/T = 2<sup>C/τ</sup>, where τ is the doubling time (23, 24).

**Statistical Analysis.** We analyzed several genomic characteristics of 425 complete prokaryotic genomes (Table S3) that were retrieved in March 2006 from the Integrated Microbial Genomes (25) and the Cluster of Orthologous Groups databases (26, 27). Automated analyses were completed and validated by manual investigation. In particular, the presence and/or absence of both thymidylate synthases was confirmed by using BLAST and PSI-BLAST tools (28, 29). The maximum growth rates of prokaryotes used in this study were taken from ref. 30. Statistical analyses were performed by using GraphPad Prism version 4.00 for Windows.

**ACKNOWLEDGMENTS.** We thank B. Dalmais and E. Bentchikou for technical advice, H. de Reuse and J. M. Ghigo (Pasteur Institute, Paris) for *H. pylori* strain 26695 and pKOBEG plasmid, respectively. F.E. was supported by the Ministère de la Recherche Française and the Fondation pour la Recherche Médicale. This work was supported by the Centre National de la Recherche Scientifique Program Microbiologie Fondamentale (U.L. and H.M.) and the Institut National de la Santé et de la Recherche Médicale AVENIR program and the Fondation Bettencourt Schuller (H.M.).

1. Mylykallio H, et al. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* 297:105–107.
2. Carreras CW, Santi DV (1995) The catalytic mechanism and structure of thymidylate synthase. *Annu Rev Biochem* 64:721–762.
3. Leduc D, et al. (2007) Flavin-dependent thymidylate synthase ThyX activity: Implications for the folate cycle in bacteria. *J Bacteriol* 189:8537–8545.
4. Finer-Moore JS, Santi DV, Stroud RM (2003) Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: Thymidylate synthase mutagenesis, function, and structure. *Biochemistry* 42:248–256.
5. Galperin MY, Koonin EV (2000) Who's your neighbor? New computational approaches for functional genomics. *Nat Biotechnol* 18:609–613.
6. Leduc D, et al. (2004) Two distinct pathways for thymidylate (dTMP) synthesis in (hyper)thermophilic Bacteria and Archaea. *Biochem Soc Trans* 32:231–235.
7. Leduc D, et al. (2004) Functional evidence for active site location of tetrameric thymidylate synthase X at the interphase of three monomers. *Proc Natl Acad Sci USA* 101:7252–7257.
8. Graziani S, et al. (2004) Functional analysis of FAD-dependent thymidylate synthase ThyX from *Paramecium bursaria* *Chlorella* virus-1. *J Biol Chem* 279:54340–54347.
9. Graziani S, et al. (2006) Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J Biol Chem* 281:24048–24057.
10. Zhong J, Skouloubris S, Dai Q, Mylykallio H, Barbour AG (2006) Function and evolution of plasmid-borne genes for pyrimidine biosynthesis in *Borrelia* spp. *J Bacteriol* 188:909–918.
11. Chernyshev A, Fleischmann T, Koehn EM, Lesley SA, Kohen A (2007) The relationships between oxidase and synthase activities of flavin dependent thymidylate synthase (FDTS). *Chem Commun* 27:2861–2863.
12. Helmsstetter, C (1996) Timing of synthetic activities in the cell cycle. *Escherichia coli* and *Salmonella: Cellular and Molecular Biology*, eds Neidhardt FC, et al. (Am Soc Microbiol, Washington, DC), pp 1627–1639.
13. Rocha EP, Danchin A (2002) Base composition bias might result from competition for metabolic resources. *Trends Genet* 18:291–294.
14. Bentley SD, Parkhill J (2004) Comparative genomic structure of prokaryotes. *Annu Rev Genet* 38:771–792.
15. Nakabachi A, et al. (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267.
16. Blanc G, et al. (2007) Reductive genome evolution from the mother of Rickettsia. *PLoS Genet* 3:e14.
17. Marais GA, Calteau A, Tenaillon O (2007) Mutation rate and genome reduction in endosymbiotic and free-living bacteria. *Genetica* 10.1007/s10709-007-9226-6.
18. Giladi M, Bitan-Banin G, Mevarech M, Ortenberg R (2002) Genetic evidence for a novel thymidylate synthase in the halophilic archaeon *Halobacterium salinarum* and in *Campylobacter jejuni*. *FEMS Microbiol Lett* 216:105–109.
19. Derbise A, Lesic B, Dacheux D, Ghigo JM, Carniel E (2003) A rapid and simple method for inactivating chromosomal genes in *Yersinia*. *FEMS Immunol Med Microbiol* 38:113–116.
20. Chaverche MK, Ghigo JM, d'Enfert C (2000) A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res* 28:E97.
21. Servant P, et al. (2007) The ClpPX protease is required for radioresistance and regulates cell division after gamma-irradiation in *Deinococcus radiodurans*. *Mol Microbiol* 66:1231–1239.
22. Hallin PF, Ussery DW (2004) CBS Genome Atlas Database: A dynamic storage for bioinformatic results and sequence data. *Bioinformatics* 20:3682–3686.
23. Riber L, Lobner-Olesen A (2005) Coordinated replication and sequestration of oriC and dnaA are required for maintaining controlled once-per-cell-cycle initiation in *Escherichia coli*. *J Bacteriol* 187:5605–5613.
24. Bremer H, Churchward G (1977) An examination of the Cooper-Helmstetter theory of DNA replication in bacteria and its underlying assumptions. *J Theor Biol* 69:645–654.



25. Markowitz VM, et al. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Res* 34:D344–D348.
26. Tatusov RL, et al. (2001) The COG database: New developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 29:22–28.
27. Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. *Science* 278:631–637.
28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410.
29. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
30. Couturier E, Rocha EP (2006) Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol Microbiol* 59:1506–1518.
31. Liu L, Santi DV (1993) Asparagine 229 in thymidylate synthase contributes to, but is not essential for, catalysis. *Proc Natl Acad Sci USA* 90:8604–8608.
32. Lambden PR, Pickett MA, Clarke IN (2006) The effect of penicillin on *Chlamydia trachomatis* DNA replication. *Microbiology* 152:2573–2578.
33. Burbage CD, Binder BJ (2007) Relationship between cell cycle and light-limited growth rate in oceanic *Prochlorococcus* (mit9312) and *Synechococcus* (wh8103) (cyanobacteria). *J Phycol* 43:266–274.
34. Myllykallio H, et al. (2000) Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* 288:2212–2215.
35. Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R (2004) Three replication origins in *Sulfolobus* species: Synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci USA* 101:7046–7051.
36. Sharpe ME, Hauser PM, Sharpe RG, Errington J (1998) *Bacillus subtilis* cell cycle as studied by fluorescence microscopy: Constancy of cell length at initiation of DNA replication and evidence for active nucleoid partitioning. *J Bacteriol* 180:547–555.
37. Altenbern RA (1973) Chromosomal mapping of *Brucella abortus*, strain 19. *Can J Microbiol* 19:109–112.
38. Higgins ML, Koch AL, Dicker DT, Daneo-Moore L (1986) Autoradiographic studies of chromosome replication during the cell cycle of *Streptococcus faecium*. *J Bacteriol* 168:541–547.
39. Meacock PA, Pritchard RH (1975) Relationship between chromosome replication and cell division in a thymineless mutant of *Escherichia coli* B/r. *J Bacteriol* 122:931–942.
40. Altenbern RA (1968) Chromosome mapping in *Staphylococcus aureus*. *J Bacteriol* 95:1642–1646.
41. Rasmussen T, Jensen RB, Skovgaard O (2007) The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. *EMBO J* 26:3124–3131.

1 **Supplemental table 1.** Bacterial strains used in this work.

Strain	Parental strain	Species	Relevant Genotype	Reference
MG1655	K-12	<i>E. coli</i>	wild type	(42)
FE010	MG1655	<i>E. coli</i>	$\Delta$ thyA:: PBCV1-thyX-aphA-3	this work
FE013	MG1655	<i>E. coli</i>	$\Delta$ thyA::aphA-3	this work
FE023	FE013	<i>E. coli</i>	$\Delta$ thyA::aphA-3 thyX in trans	this work and (10)
26695		<i>H. pylori</i>	wild-type	(43)

2  
3 **Supplemental table 2.** Oligonucleotides used in this work.

Name	Sequence
oEF011	GGTGATGATTATCTTCGCCCGCCGTACTAAACG
oEF012	CCAGTCATGCTAGCCATGGTTCCTCAGGAAACGTGTTG
oEF013	ATGGCTAGCATGACTGGTTCGCAAAGCTC
oEF015	GGATGAATTGTTTTAGTTACGAAACATCCTGCCAGAGCC
oEF016	GCGAAGGTGTCAGGTCGCTCATTTCTGACTTC
oEF017	GTTAGTCACCCGGTACTTAGACCCAATCAAGTGCTTTTGC
oEF031	GCCTGATTGGCGTTATCGTG
oEF032	GCGCAGGCCAAAGAAAGCAA
oEF048	TTGTTTTTAAGCTTTTAGTTGCCACGACTTCCTGG
oEF065	CCATGCATTCTTCCAGTTCT
oEF066	TCGAAACGGTAGTCGAAGAT
oEF102	TTGAGCCGGCATAGAATAGC
oEF103	GGGGTAAAAAGCTCTCCACTTC
oEF106	GCCAGAGCGCTTTAAACAAC
oEF107	TGTAGCGCGTTTGCTCTATG

5

6

Domain	Name	doubling time	Genes	AT Perc	Bases	GC Count	AT count	ThyA	ThyX
Archaea	Aeropyrum pernix K1		2751	43,69%	1669695	940227	729468	0	1
Archaea	Archaeoglobus fulgidus DSM 4304		2456	51,42%	2178400	1058303	1120097	1	0
Archaea	Cenarchaeum symbiosum		2059	42,53%	2023713	1162895	860818	0	1
Archaea	Ferroplasma acidarmanus fer1		1742	62,99%	1971391	728816	1242575	0	1
Archaea	Haloarcula marismortui ATCC 43049		4300	38,88%	4274642	2612610	1662032	0	1
Archaea	Halobacterium salinarum NRC-1		2657	34,07%	2571010	1695157	875853	0	1
Archaea	Methanocaldococcus jannaschii DSM 2661		1813	68,71%	1739927	544485	1195442	1	0
Archaea	Methanococcoides burtonii DSM6242		2551	59,21%	2597419	1059593	1537826	1	0
Archaea	Methanococcus maripaludis S2		1770	66,90%	1661137	549842	1111295	1	0
Archaea	Methanoculleus marisnigri		2466	37,93%	2460155	1526918	933237	1	0
Archaea	Methanopyrus kandleri AV19		1734	38,84%	1694969	1036629	658340	1	0
Archaea	Methanosarcina acetivorans C2A		4609	57,32%	5751492	2454788	3296704	1	0
Archaea	Methanosarcina barkeri Fusaro		3830	60,77%	4873766	1912156	2961610	1	0
Archaea	Methanosarcina mazei Go1		3438	58,52%	4096345	1699365	2396980	1	0
Archaea	Methanospirillum hungatei JF-1		3369	54,83%	3535022	1596563	1938459	1	0
Archaea	Methanothermobacter thermautotrophicus Delta H		1915	50,46%	1751377	867701	883676	1	0
Archaea	Nanoarchaeum equitans Kin4-M		579	68,44%	490885	154921	335964	0	1
Archaea	Natronomonas pharaonis DSM 2160		2892	36,88%	2749696	1735599	1014097	1	0
Archaea	Picrophilus torridus DSM 9790		1628	64,03%	1545895	556065	989830	0	1
Archaea	Pyrobaculum aerophilum IM2		2655	48,64%	2222430	1141496	1080934	0	1
Archaea	Pyrococcus abyssi GE5		1881	55,29%	1765118	789253	975865	0	1
Archaea	Pyrococcus furiosus DSM 3638		2115	59,23%	1908256	777994	1130262	0	1
Archaea	Pyrococcus horikoshii OT3		2112	58,12%	1738505	728068	1010437	0	1
Archaea	Sulfolobus acidocaldarius DSM 639		2322	63,29%	2225959	817096	1408863	0	1
Archaea	Sulfolobus solfataricus P2		3050	64,21%	2992245	1070840	1921405	0	1
Archaea	Sulfolobus tokodaii 7		2878	67,21%	2694756	883522	1811234	0	1
Archaea	Thermococcus kodakaraensis KOD1		2358	48,00%	2088737	1086049	1002688	0	1
Archaea	Thermofilum pendens		1932	42,36%	1811389	1044172	767217	1	0
Archaea	Thermoplasma acidophilum DSM 1728		1549	54,01%	1564906	719769	845137	0	1
Archaea	Thermoplasma volcanium GSS1		1575	60,08%	1584804	632630	952174	0	1
Bacteria	Acinetobacter sp. ADP1	0,5	3422	59,57%	3598621	1454907	2143714	1	0
Bacteria	Actinobacillus pleuropneumoniae 4074		2193	58,62%	2292348	948502	1343846	1	0
Bacteria	Actinobacillus succinogenes 130Z		1938	55,11%	2046146	918573	1127573	1	0
Bacteria	Agrobacterium tumefaciens C58 Cereon	3	5366	40,96%	5673465	3349785	2323680	1	0
Bacteria	Agrobacterium tumefaciens C58 Dupont	3	5467	40,96%	5674064	3350042	2324022	1	0
Bacteria	Alkaliimnicola ehrlichei MLHE-1		2923	32,46%	3272789	2210429	1062360	1	0
Bacteria	Alkaliphilus metalliredigenes QYMF		4233	63,36%	4410303	1616081	2794222	0	1
Bacteria	Anabaena sp. PCC 7120		6214	58,73%	7211789	2976185	4235604	0	1
Bacteria	Anabaena variabilis ATCC29413		5805	58,59%	7105752	2942474	4163278	0	1
Bacteria	Anaeromyxobacter dehalogenans 2CP-C		4416	25,09%	5013479	3755360	1258119	0	1
Bacteria	Anaplasma marginale St. Maries		1005	50,24%	1197687	596001	601686	0	1
Bacteria	Aquifex aeolicus VF5		1603	56,70%	1590791	688838	901953	0	1
Bacteria	Arthrobacter sp. FB24		4591	34,60%	5011599	3277672	1733927	1	0
Bacteria	Azoarcus sp. EbN1	4,3	4668	35,30%	4727255	3058421	1668834	1	0
Bacteria	Azotobacter vinelandii AvOP		5058	34,31%	5352434	3515639	1836795	1	0
Bacteria	Bacillus anthracis A2012		5390	64,78%	5093554	1779131	3314423	1	0
Bacteria	Bacillus anthracis Ames		5439	64,62%	5227293	1849251	3378042	1	1
Bacteria	Bacillus anthracis Ames 0581		5745	64,76%	5503926	1939746	3564180	1	1
Bacteria	Bacillus anthracis Sterne		5415	64,62%	5228663	1849872	3378791	1	1
Bacteria	Bacillus cereus ATCC 10987		6008	64,51%	5432652	1928302	3504350	1	0
Bacteria	Bacillus cereus ATCC 14579	0,3	5402	64,71%	5427083	1915128	3511955	1	0
Bacteria	Bacillus cereus E33L		5269	64,65%	5300915	1873974	3426941	1	1
Bacteria	Bacillus cereus G9241		6316	64,75%	5934942	2091877	3843065	1	0
Bacteria	Bacillus cereus NVH391-98		4033	64,18%	3915845	1402645	2513200	1	0
Bacteria	Bacillus clausii KSM-K16		4205	55,25%	4303871	1926125	2377746	1	0
Bacteria	Bacillus halodurans C-125	0,6	4172	56,31%	4202352	1835819	2366533	1	0
Bacteria	Bacillus licheniformis Goettingen	0,58	4289	53,81%	4222645	1950640	2272005	1	0
Bacteria	Bacillus licheniformis Novozymes		4245	53,80%	4222334	1950541	2271793	1	0
Bacteria	Bacillus subtilis 168	0,43	4225	56,48%	4214630	1834115	2380515	1	0
Bacteria	Bacillus thuringiensis konkukian 97-27		5264	64,59%	5237682	1854419	3383263	1	0
Bacteria	Bacteroides fragilis NCTC9343	0,63	4400	56,89%	5241700	2259916	2981784	1	0
Bacteria	Bacteroides fragilis YCH46		4717	56,80%	5310990	2294576	3016414	1	0
Bacteria	Bacteroides thetaiotaomicron VPI-5482	1,47	4902	57,14%	6293399	2697369	3596030	1	0
Bacteria	Bartonella henselae Houston-1	3	1664	61,77%	1931047	738246	1192801	1	0
Bacteria	Bartonella quintana Toulouse	3	1357	61,20%	1581384	613527	967857	1	0
Bacteria	Bdellovibrio bacteriovorus HD100	1,4	3625	49,35%	3782950	1915894	1867056	1	0
Bacteria	Bifidobacterium longum DJO10A		2020	39,92%	2375286	1403005	972281	1	0
Bacteria	Bifidobacterium longum NCC2705		1802	39,87%	2260272	1359041	901231	1	0
Bacteria	Blochmannia floridanus	36	635	72,62%	705557	193204	512353	0	0
Bacteria	Blochmannia pennsylvanicus BPEN		692	70,44%	791654	234050	557604	0	0
Bacteria	Bordetella bronchiseptica RB50		5070	31,92%	5339179	3634754	1704425	1	0
Bacteria	Bordetella parapertussis 12822		4465	31,90%	4773551	3250786	1522765	1	0
Bacteria	Bordetella pertussis Tohama I	6	3866	32,28%	4086189	2767191	1318998	1	0
Bacteria	Borrelia burgdorferi B31	4	1677	71,83%	1519856	428130	1091726	0	1
Bacteria	Borrelia garinii Pbi		972	71,87%	986914	277589	709325	0	0
Bacteria	Bradyrhizobium japonicum USDA 110	20	8371	35,94%	9105828	5833137	3272691	1	0
Bacteria	Bradyrhizobium sp. BTAi1		7518	35,17%	8422430	5460210	2962220	1	0
Bacteria	Brevibacterium linens BL2		4038	37,21%	4510745	2747796	1762949	0	1
Bacteria	Brucella abortus 2308	2	3414	42,78%	3278307	1875908	1402399	1	0
Bacteria	Brucella abortus 9-941	2	3149	42,78%	3286445	1880482	1405963	1	0
Bacteria	Brucella melitensis 16M	2	3264	42,78%	3294931	1885435	1409496	1	0
Bacteria	Brucella suis 1330	2	3337	42,75%	3315175	1897981	1417194	1	0
Bacteria	Buchnera aphidicola APS		609	73,64%	655725	172881	482844	1	0
Bacteria	Buchnera aphidicola Bp		542	74,66%	618379	156707	461672	1	0
Bacteria	Buchnera aphidicola Sg	36	590	74,65%	654174	165834	488340	1	0
Bacteria	Burkholderia ambifaria AMMD		6617	33,20%	7503613	5012715	2490898	1	0
Bacteria	Burkholderia cenocepacia AU 1054		6555	33,04%	7249477	4854030	2395447	1	0
Bacteria	Burkholderia cenocepacia H12424		7146	33,62%	8139086	5402682	2736404	1	0

Bacteria	Burkholderia mallei 10229		5201	31,83%	6065244	4134472	1930772	1	0
Bacteria	Burkholderia mallei 10399		4947	31,54%	5933917	4062138	1871779	1	0
Bacteria	Burkholderia mallei ATCC 23344	0,4	4831	31,51%	5835527	3996491	1839036	1	0
Bacteria	Burkholderia mallei FMH		2687	31,77%	2273401	1551210	722191	1	0
Bacteria	Burkholderia mallei GB8 horse 4		4884	31,53%	5804858	3974520	1830338	1	0
Bacteria	Burkholderia mallei JHU		2608	31,81%	2190539	1493749	696790	1	0
Bacteria	Burkholderia mallei NCTC 10247		5146	31,54%	5976679	4091621	1885058	1	0
Bacteria	Burkholderia mallei SAVP1		4438	31,63%	5283468	3612235	1671233	1	0
Bacteria	Burkholderia pseudomallei 1655		5465	32,00%	7029452	4780314	2249138	1	0
Bacteria	Burkholderia pseudomallei 1710a		5565	32,03%	7319487	4975220	2344267	1	0
Bacteria	Burkholderia pseudomallei 1710b		5553	32,02%	7390459	5023952	2366507	1	0
Bacteria	Burkholderia pseudomallei 668		5495	31,69%	7071938	4830559	2241379	1	0
Bacteria	Burkholderia pseudomallei K96243	0,7	5928	31,94%	7247547	4932588	2314959	1	0
Bacteria	Burkholderia pseudomallei Pasteur		5653	31,98%	7346573	4997318	2349255	1	0
Bacteria	Burkholderia pseudomallei S13		5771	31,90%	7388320	5031135	2357185	1	0
Bacteria	Burkholderia sp. 383		7824	33,73%	8676277	5749981	2926296	1	0
Bacteria	Burkholderia vietnamiensis G4		7992	34,25%	8410976	5529991	2880985	1	0
Bacteria	Burkholderia xenovorans LB400		9012	37,36%	9703676	6078164	3625512	1	0
Bacteria	Caldicellulosiruptor saccharolyticus DSM 8903		2647	64,79%	2788317	981765	1806552	0	1
Bacteria	Campylobacter coli RM2228		2019	68,88%	1860666	578951	1281715	0	1
Bacteria	Campylobacter jejuni NCTC 11168	1,5	1707	69,45%	1641481	501449	1140032	0	1
Bacteria	Campylobacter jejuni RM1221		1891	69,69%	1777831	538792	1239039	0	1
Bacteria	Campylobacter lari RM2100		1651	70,44%	1562926	461952	1100974	0	1
Bacteria	Campylobacter upsaliensis RM3195		1984	65,73%	1773834	607817	1166017	0	1
Bacteria	Carboxydotherrnus hydrogenoformans Z-2901		2682	57,95%	2401520	1009832	1391688	0	1
Bacteria	Caulobacter crescentus CB15	1,5	3794	32,79%	4016947	2699950	1316997	1	0
Bacteria	Chlamydia muridarum Nigg		954	59,70%	1080451	435471	644980	0	1
Bacteria	Chlamydia trachomatis D/UW-3/CX	24	937	58,69%	1042519	430636	611883	0	1
Bacteria	Chlamydia abortus S26/3	24	1002	60,13%	1144377	456222	688155	0	1
Bacteria	Chlamydia caviae GPIC	24	1047	60,81%	1181356	462922	718434	0	1
Bacteria	Chlamydia pneumoniae AR39	24	1151	59,43%	1229853	498986	730867	0	1
Bacteria	Chlamydia pneumoniae CWL029		1118	59,42%	1230230	499199	731031	0	1
Bacteria	Chlamydia pneumoniae J138		1134	59,42%	1226565	497758	728807	0	1
Bacteria	Chlamydia pneumoniae TW-183		1154	59,42%	1225935	497438	728497	0	1
Bacteria	Chlorobium limicola DSMZ 245		2483	48,67%	2761915	1417363	1344552	0	1
Bacteria	Chlorobium phaeobacteroides BS1		3845	54,47%	4444192	2023430	2420762	0	1
Bacteria	Chlorobium phaeobacteroides DSMZ 266		2840	51,66%	3114286	1505082	1609204	0	1
Bacteria	Chlorobium tepidum TLS	2	2308	43,47%	2154946	1218184	936762	0	1
Bacteria	Chlorobium vibrioforme f. thiosulfatophilum DSMZ 265		1798	47,01%	1980186	1049116	931070	0	1
Bacteria	Chlorochromatium aggregatum		2081	55,72%	2572079	1138870	1433209	0	1
Bacteria	Chloroflexus aurantiacus J-10-fl		3994	43,29%	5193782	2945582	2248200	0	1
Bacteria	Chromobacterium violaceum ATCC 12472	0,8	4530	35,17%	4751080	3080303	1670777	1	0
Bacteria	Chromohalobacter salexigens DSM3043		3422	36,06%	3671024	2347366	1323658	1	0
Bacteria	Clostridium acetobutylicum ATCC 824	0,58	3955	69,07%	4132880	1278098	2854782	1	1
Bacteria	Clostridium beijerinckii NCIMB 8052		5210	70,18%	5952522	1775063	4177459	1	0
Bacteria	Clostridium perfringens 13	0,2	2849	71,49%	3085740	879807	2205933	0	1
Bacteria	Clostridium tetani E88		2504	71,36%	2873333	822936	2050397	0	1
Bacteria	Clostridium thermocellum ATCC 27405		3235	60,80%	3894953	1526454	2368499	1	0
Bacteria	Colwellia psychrerythraea 34H		5114	61,99%	5373180	2042100	3331080	1	0
Bacteria	Corynebacterium diphtheriae NCTC 13129		2389	46,52%	2488635	1330885	1157750	1	1
Bacteria	Corynebacterium efficiens YS-314		3069	37,03%	3219505	2027445	1192060	1	1
Bacteria	Corynebacterium glutamicum ATCC 13032 Kalinowski	1,2	3136	46,16%	3282708	1767468	1515240	1	1
Bacteria	Corynebacterium glutamicum ATCC 13032 Nakagawa		3177	46,19%	3309401	1780797	1528604	1	1
Bacteria	Corynebacterium jeikeium K411		2229	38,65%	2476822	1519564	957258	1	1
Bacteria	Coxiella burnetii RSA 493	19	2091	57,40%	2032668	865891	1166777	1	0
Bacteria	Croceibacter atlanticus HTCC2559		2756	66,08%	2954711	1002094	1952617	1	0
Bacteria	Crocospaera watsonii WH8501		6011	62,88%	6285399	2332562	3952837	1	0
Bacteria	Cupriavidus metallidurans CH34		6377	36,46%	6887566	4376367	2511199	1	0
Bacteria	Cupriavidus necator (Ralstonia eutropha) JMP134		6613	35,57%	7255290	4674830	2580460	1	0
Bacteria	Cytophaga hutchinsonii ATCC 33406		3640	61,15%	4433218	1722148	2711070	1	0
Bacteria	Dechloromonas aromatica RCB		4280	40,75%	4501104	2666779	1834325	1	0
Bacteria	Dehalococcoides ethenogenes 195	19	1629	51,15%	1469720	717962	751758	1	1
Bacteria	Dehalococcoides sp. CBDB1		1508	52,97%	1395502	662647	739255	1	1
Bacteria	Deinococcus geothermalis DSM11300		3058	33,58%	3164085	2101652	1062433	0	1
Bacteria	Deinococcus radiodurans R1		3159	33,39%	3284156	2187673	1096483	1	0
Bacteria	Desulfotobacterium hafniense DCB-2		4463	48,96%	6083768	3104777	2978991	0	1
Bacteria	Desulfotalea psychrophila L.Sv54	0,37	3343	53,37%	3659634	1706546	1953088	0	1
Bacteria	Desulfovibrio desulfuricans G20		3862	42,16%	3730232	2157422	1572810	0	1
Bacteria	Desulfovibrio vulgaris Hildenborough	14	3615	36,72%	3773159	2387614	1385545	0	1
Bacteria	Desulfuromonas acetoxidans DSM 684		3065	48,12%	3528973	1830939	1698034	0	1
Bacteria	Ehrlichia canis Jake		983	71,04%	1315030	380803	934227	0	1
Bacteria	Ehrlichia chaffeensis sapulpa		834	70,02%	1005812	301539	704273	0	1
Bacteria	Ehrlichia ruminantium Gardel		1019	72,49%	1499920	412667	1087253	0	1
Bacteria	Ehrlichia ruminantium Welgevonden ARC-OVI		982	72,52%	1516355	416676	1099679	0	1
Bacteria	Ehrlichia ruminantium Welgevonden CIRAD		1015	72,52%	1512977	415749	1097228	0	1
Bacteria	Enterococcus faecalis V583	0,5	3344	62,62%	3359974	1255997	2103977	1	0
Bacteria	Enterococcus faecium DO		2794	62,09%	2848380	1079676	1768704	1	0
Bacteria	Erwinia carotovora subsp. atroseptica SCRI1043	0,2	4590	49,03%	5064019	2580995	2483024	1	0
Bacteria	Erythrobacter litoralis HTCC2594		3088	36,93%	3052801	1925389	1127412	1	0
Bacteria	Escherichia coli K12	0,35	4468	49,27%	4738834	2404240	2334594	1	0
Bacteria	Escherichia coli O157:H7 EDL933		5571	49,62%	5620522	2829558	2790964	1	0
Bacteria	Escherichia coli O6 CFT073		5489	49,52%	5231428	2640581	2590847	1	0
Bacteria	Escherichia coli Sakai O157:H7		5571	49,51%	5591171	2822890	2768281	1	0
Bacteria	Exiguobacterium sp. 255-15		3073	52,21%	2894116	1383089	1511027	1	0
Bacteria	Francisella tularensis SCHU S4		1863	67,74%	1892819	610610	1282209	1	0
Bacteria	Frankia sp. Ccl3		4611	29,93%	5381405	3770668	1610737	0	1
Bacteria	Frankia sp. EAN1pec		8039	29,06%	9081415	6441809	2639606	0	1
Bacteria	Fusobacterium nucleatum nucleatum ATCC 25586	0,72	2129	72,85%	2174500	590411	1584089	0	1
Bacteria	Fusobacterium nucleatum subsp. vincentii ATCC 49256		2327	72,66%	2118259	578995	1539264	1	0
Bacteria	Geobacillus kaustophilus HTA426		3654	48,02%	3592666	1867483	1725183	0	1
Bacteria	Geobacter metallireducens GS-15		3631	40,51%	4011182	2386067	1625115	0	1
Bacteria	Geobacter sulfurreducens PCA	6	3502	39,06%	3814139	2324233	1489906	0	1
Bacteria	Gloeobacter violaceus PCC 7421		4478	38,00%	4659019	2888492	1770527	0	1
Bacteria	Gluconobacter oxydans 621H	1	2732	39,18%	2922384	1777303	1145081	0	1
Bacteria	Haemophilus ducreyi 35000HP	1,8	1784	61,78%	1698955	649349	1049606	1	0

Bacteria	Haemophilus influenzae 86-028NP	0,5	1869	61,85%	1913428	730017	1183411	1	0
Bacteria	Haemophilus influenzae R2846		1759	62,03%	1824242	692696	1131546	1	0
Bacteria	Haemophilus influenzae R2866		1892	61,92%	1933340	736202	1197138	1	0
Bacteria	Haemophilus influenzae Rd KW20		1785	61,85%	1830138	698171	1131967	1	0
Bacteria	Haemophilus somnus 129PT		1845	62,80%	2008359	747055	1261304	1	0
Bacteria	Haemophilus somnus 2336		2131	62,59%	2247676	840662	1407014	1	0
Bacteria	Helicobacter hepaticus ATCC 51449	4,2	1939	64,07%	1799146	646456	1152690	0	1
Bacteria	Helicobacter pylori 26695	2,44	1609	61,13%	1667867	648358	1019509	0	1
Bacteria	Helicobacter pylori J99		1531	60,81%	1643831	644196	999635	0	1
Bacteria	Idiomarina loihiensis L2TR		2697	52,96%	2839318	1335622	1503696	1	0
Bacteria	Jannaschia sp.CCS1		4337	37,76%	4399650	2738389	1661261	0	1
Bacteria	Kineococcus radiotolerans SRS30216		4618	25,79%	4893957	3631695	1262262	1	0
Bacteria	Lactobacillus acidophilus NCFM	1,8	1944	65,29%	1993564	692059	1301505	1	0
Bacteria	Lactobacillus brevis ATCC367		1600	53,58%	1880818	809760	1071058	1	0
Bacteria	Lactobacillus casei ATCC334		2529	53,31%	2760660	1135750	1624910	0	0
Bacteria	Lactobacillus delbrueckii bulgaricus ATCCBAA-365		1622	49,98%	1629380	744783	884597	1	0
Bacteria	Lactobacillus gasseri ATCC3323		1757	64,96%	1950210	642151	1308059	1	0
Bacteria	Lactobacillus johnsonii NCC 533	0,9	1915	65,39%	1992676	686990	1302986	1	0
Bacteria	Lactobacillus plantarum WCFS1	1,6	3137	55,53%	3308274	1471079	1837195	1	0
Bacteria	Lactobacillus sakei sakei 23K		1968	58,74%	1884661	777566	1107095	1	0
Bacteria	Lactococcus lactis cremoris SK11		2690	64,36%	2613164	807395	1805769	1	0
Bacteria	Lactococcus lactis lactis IL1403	0,7	2346	64,67%	2365589	835748	1529841	1	0
Bacteria	Legionella pneumophila Lens		2999	61,58%	3345687	1285245	2060442	1	0
Bacteria	Legionella pneumophila Paris		3134	61,63%	3503610	1344403	2159207	1	0
Bacteria	Legionella pneumophila Philadelphia 1	3,3	2997	61,73%	3397754	1300304	2097450	1	0
Bacteria	Leifsonia xyli xyli CTCB07	5	2078	32,32%	2584158	1748856	835302	1	0
Bacteria	Leptospira interrogans serovar Copenhageni Fiocruz	9	3700	64,96%	4627366	1621630	3005736	0	1
Bacteria	Leptospira interrogans serovar Lai 56601	9	4766	64,98%	4691184	1642656	3048528	0	1
Bacteria	Leuconostoc mesenteroides ATCC 8293		1921	62,47%	1976579	738378	1238201	1	0
Bacteria	Listeria innocua Clip11262	0,6	3145	62,61%	3093113	1156458	1936655	1	0
Bacteria	Listeria monocytogenes 4b F2365		2906	61,96%	2905310	1105222	1800088	1	0
Bacteria	Listeria monocytogenes EGD-e	1	2940	62,02%	2944528	1118362	1826166	1	0
Bacteria	Listeria monocytogenes str. 1/2a F6854		3052	62,15%	2950285	1116567	1833718	1	0
Bacteria	Listeria monocytogenes str. 4b H7858		3191	62,05%	2972254	1127902	1844352	1	0
Bacteria	Magnetococcus sp. MC-1		3770	45,80%	4628740	2508716	2120024	0	1
Bacteria	Magnetospirillum magnetotacticum MS-1		4357	36,04%	4503280	2879463	1623817	1	0
Bacteria	Mannheimia succiniciproducens MBEL55E		2463	57,46%	2314078	984362	1329716	1	0
Bacteria	Marinobacter aquaeolei VT8		4210	43,06%	4647952	2646714	2001238	1	0
Bacteria	Mesoplasma florum L1		716	72,98%	793224	214333	578891	1	0
Bacteria	Mesorhizobium loti MAFF303099		7341	37,49%	7596297	4748293	2848004	1	0
Bacteria	Mesorhizobium sp. BNC1		4537	39,02%	4977960	3035181	1942779	1	0
Bacteria	Methylobacillus flagellatus KT		2553	44,22%	2851330	1590394	1260936	1	0
Bacteria	Methylococcus capsulatus Bath		3011	36,42%	3304554	2101185	1203369	1	0
Bacteria	Moorella thermoacetica ATCC 39073		2577	44,21%	2628784	1466639	1162145	0	1
Bacteria	Mycobacterium avium paratuberculosis k10	10	4399	30,70%	4829781	3346971	1482810	1	1
Bacteria	Mycobacterium bovis AF2122/97		4002	34,37%	4345492	2852076	1493416	1	1
Bacteria	Mycobacterium leprae TN	240	2768	42,20%	3268203	1888915	1379288	1	1
Bacteria	Mycobacterium tuberculosis CDC1551	24	4238	34,39%	4403837	2889246	1514591	1	1
Bacteria	Mycobacterium tuberculosis H37Rv		4047	34,39%	4411532	2894588	1516944	1	1
Bacteria	Mycoplasma gallisepticum R	1	765	68,55%	996422	313375	683047	1	0
Bacteria	Mycoplasma genitalium G-37	12	519	68,31%	580074	183836	396238	1	0
Bacteria	Mycoplasma hyopneumoniae 232		724	71,44%	892758	254982	637776	0	0
Bacteria	Mycoplasma hyopneumoniae 7448		726	71,51%	920079	262120	657959	0	0
Bacteria	Mycoplasma hyopneumoniae J		728	71,48%	897405	255960	641445	0	0
Bacteria	Mycoplasma mobile 163K		666	75,05%	777079	193891	583188	1	0
Bacteria	Mycoplasma mycoides SC PG1		1052	76,03%	1211703	290393	921310	0	0
Bacteria	Mycoplasma penetrans HF-2		1072	74,28%	1358633	349407	1009226	1	0
Bacteria	Mycoplasma pneumoniae M129	6	728	59,99%	816394	326623	489771	1	0
Bacteria	Mycoplasma pulmonis UAB CTIP	2	814	73,36%	963879	256744	707135	1	0
Bacteria	Mycoplasma synoviae 53		747	71,50%	799476	227812	571664	1	0
Bacteria	Neisseria gonorrhoeae FA 1090	1,1	2069	47,31%	2153922	1134894	1019028	1	0
Bacteria	Neisseria meningitidis MC58	1	2097	48,47%	2272351	1170894	1101457	1	0
Bacteria	Neisseria meningitidis Z2491		2191	48,19%	2184406	1131722	1052684	1	0
Bacteria	Nitrobacter hamburgensis X14		4728	38,39%	5010322	3086778	1923544	1	0
Bacteria	Nitrobacter winogradskyi Nb-255		3195	37,95%	3402093	2110965	1291128	1	0
Bacteria	Nitrosococcus oceani ATCC 19707		3183	49,66%	3522111	1773098	1749013	1	0
Bacteria	Nitrosomonas europaea ATCC 19718	18,5	2619	49,28%	2812094	1426243	1385851	1	0
Bacteria	Nitrosomonas eutropha C71		2588	51,53%	2711982	1314561	1397421	1	0
Bacteria	Nitrosospora multiformis ATCC25196		2873	46,12%	3234309	1742553	1491756	1	0
Bacteria	Nocardia farcinica IFM 10152	3	5999	29,31%	6292344	4448204	1844140	1	0
Bacteria	Nocardioides sp. JS614		4954	28,86%	5394058	3837048	1557010	1	0
Bacteria	Nostoc punctiforme PCC73102		7465	58,65%	9059191	3746385	5312806	0	1
Bacteria	Novosphingobium aromaticivorans DSM 12444		3823	34,89%	4226597	2751686	1474911	1	0
Bacteria	Oceanicaulis alexandrii HTCC2633		3077	36,41%	3168201	2014697	1153504	1	0
Bacteria	Oceanobacillus iheyensis HTE831		3588	64,32%	3630528	1295420	2335108	1	0
Bacteria	Oenococcus oeni PSU-1		1834	62,21%	1782786	668573	1114213	1	0
Bacteria	Onion yellows phytoplasma OY-M		792	72,26%	860631	238710	621921	1	0
Bacteria	Parachlamydia sp. UWE25	48	2075	65,28%	2414465	838236	1576229	0	1
Bacteria	Paracoccus denitrificans PD1222		5157	33,20%	5175736	3457359	1718377	0	1
Bacteria	Pasteurella multocida subsp. multocida Pm70	1	2090	59,60%	2257487	912135	1345352	1	0
Bacteria	Pediococcus pentosaceus ATCC25745		1722	62,81%	1814631	670195	1144436	1	0
Bacteria	Pelagibacter ubique HTCC1062		1406	70,32%	1308759	388480	920279	0	1
Bacteria	Pelobacter carbinolicus DSM 2380		3205	44,89%	3662252	2018354	1643898	0	1
Bacteria	Pelobacter propionicus DSM 2379		3896	41,94%	4466736	2573686	1893050	0	1
Bacteria	Pelodictyon luteolum DSMZ 273(T)		2187	42,67%	2364842	1355765	1009077	0	1
Bacteria	Pelodictyon phaeooclathratiforme BU-1		2814	51,92%	3000217	1442307	1557910	0	1
Bacteria	Photobacterium profundum SS9	2,5	5697	58,27%	6403280	2672333	3730947	1	0
Bacteria	Photorhabdus luminescens laumondii TTO1	0,5	5012	57,17%	5688987	2436329	3252658	1	0
Bacteria	Polaromonas naphthalenivorans CJ2		5022	38,30%	5340188	3295063	2045125	1	0
Bacteria	Polaromonas sp. JS666		5709	38,03%	5912295	3663567	2248728	1	0
Bacteria	Porphyromonas gingivalis W83	2,7	1974	51,71%	2343476	1131616	2118660	1	0
Bacteria	Prochlorococcus marinus MED4		1756	69,20%	1657990	510647	1147343	0	1
Bacteria	Prochlorococcus marinus MIT 9312		1811	68,79%	1709204	533451	1175753	0	1
Bacteria	Prochlorococcus marinus MIT 9313		2324	49,26%	2410873	1223271	1187602	0	1
Bacteria	Prochlorococcus marinus sp. NATL2A		1936	64,88%	1842899	647282	1195617	0	1



Bacteria	<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> CCMP1375	17	1926	63,56%	1751080	638131	1112949	0	1
Bacteria	<i>Prochlorococcus</i> sp. CC9605 (oligotrophic)		2747	40,78%	2510659	1486753	1023906	0	1
Bacteria	<i>Prochlorococcus</i> sp. CC9902 (coastal)		2372	45,84%	2234828	1210483	1024345	0	1
Bacteria	<i>Propionibacterium acnes</i> KPA171202	5,1	2351	39,99%	2560265	1536434	1023831	1	0
Bacteria	<i>Prosthecochloris aestuarii</i> SK413		2359	49,98%	2563197	1282162	1281035	0	1
Bacteria	<i>Pseudoalteromonas atlantica</i> T6c		4297	55,40%	5094958	2272596	2822362	1	0
Bacteria	<i>Pseudoalteromonas haloplanktis</i> TAC 125		3621	59,91%	3850272	1543638	2306634	1	0
Bacteria	<i>Pseudomonas aeruginosa</i> PAO1	0,5	5642	33,44%	6264403	4169319	2095084	1	0
Bacteria	<i>Pseudomonas aeruginosa</i> UCBPP-PA14		5754	33,68%	6524142	4326566	2197576	1	0
Bacteria	<i>Pseudomonas fluorescens</i> Pf-5		6295	36,70%	7074893	4478756	2596137	1	0
Bacteria	<i>Pseudomonas fluorescens</i> PfO-1		5837	39,48%	6438405	3896570	2541835	1	0
Bacteria	<i>Pseudomonas putida</i> F1	1,1	5341	38,11%	5925059	3666802	2258257	1	0
Bacteria	<i>Pseudomonas putida</i> KT2440		5445	38,48%	6181863	3803335	2378528	1	0
Bacteria	<i>Pseudomonas syringae</i> B728a		5217	40,77%	6093698	3609468	2484230	1	0
Bacteria	<i>Pseudomonas syringae</i> phaseolicola 1448A		5312	42,09%	6112448	3539920	2572528	1	0
Bacteria	<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000	1,47	5688	41,66%	6538260	3814368	2723892	1	0
Bacteria	<i>Psychrobacter arcticum</i> 273-4		2208	57,20%	2650701	1134506	1516195	1	0
Bacteria	<i>Psychrobacter cryohalolentis</i> K5		2538	57,75%	3081078	1301763	1779315	1	0
Bacteria	<i>Ralstonia solanacearum</i> GM1000	4	5189	33,02%	5810922	3891890	1919032	1	0
Bacteria	<i>Rhodobacter sphaeroides</i> 2.4.1		4367	31,21%	4603060	3166329	1436731	1	0
Bacteria	<i>Rhodobacter sphaeroides</i> ATCC 17025		4296	31,76%	4479360	3056644	1422716	1	0
Bacteria	<i>Rhodobacter sphaeroides</i> ATCC 17029		4019	30,99%	4461982	3079434	1382548	1	0
Bacteria	<i>Rhodoferax ferrireducens</i> DSM 15236		4427	40,36%	4926722	2938433	1988289	1	0
Bacteria	<i>Rhodopirellula baltica</i> SH 1		7403	44,60%	7145576	3958884	3186692	1	0
Bacteria	<i>Rhodopseudomonas palustris</i> BisA53	9	4969	35,56%	5502424	3545828	1956596	1	0
Bacteria	<i>Rhodopseudomonas palustris</i> BisB18		5007	35,05%	5504960	3575687	1929273	1	0
Bacteria	<i>Rhodopseudomonas palustris</i> BisB5		4443	35,19%	4890427	3169726	1720701	1	0
Bacteria	<i>Rhodopseudomonas palustris</i> CGA009		4897	34,97%	5467640	3555665	1911975	1	0
Bacteria	<i>Rhodopseudomonas palustris</i> HaA2		4736	33,96%	5331481	3520829	1810652	1	0
Bacteria	<i>Rhodospirillum rubrum</i> ATCC 11170		3917	34,62%	4406557	2880951	1525606	0	1
Bacteria	<i>Rickettsia akari</i> Hartford		1253	67,67%	1231060	398060	833000	0	1
Bacteria	<i>Rickettsia conorii</i> Malish 7	4,1	1410	67,56%	1268755	411569	857186	0	1
Bacteria	<i>Rickettsia felis</i> URRWXC2		1548	67,49%	1587240	516066	1071174	0	1
Bacteria	<i>Rickettsia prowazekii</i> Madrid E	10	870	71,00%	1111523	322345	789178	0	1
Bacteria	<i>Rickettsia rickettsii</i>		1346	67,53%	1257710	408346	849364	0	1
Bacteria	<i>Rickettsia sibirica</i> 246		1270	67,53%	1250021	405874	844147	0	1
Bacteria	<i>Rickettsia typhi</i> Wilmington	10	874	71,08%	1111496	321438	790058	0	1
Bacteria	<i>Roseovarius nubinihibens</i> ISM		3597	36,09%	3668667	2344580	1324087	0	1
Bacteria	<i>Rubrivivax gelatinosus</i> PM1		4548	31,22%	4643669	3194002	1449667	1	0
Bacteria	<i>Rubrobacter xylanophilus</i> DSM 9941		3039	30,04%	3299423	2307838	991585	0	1
Bacteria	<i>Saccharophagus degradans</i> 2-40		4064	54,17%	5057531	2317668	2739863	1	0
Bacteria	<i>Salmonella enterica</i> Paratyphi-A SARB42		4197	47,84%	4585229	2391855	2193374	1	0
Bacteria	<i>Salmonella enterica</i> Typhi CT18		5082	48,13%	5133713	2662865	2470848	1	0
Bacteria	<i>Salmonella enterica</i> Typhi Ty2		4423	47,95%	4791961	2494292	2297669	1	0
Bacteria	<i>Salmonella enterica</i> choleraesuis SC-B67		4773	47,87%	4944000	2577430	2366570	1	0
Bacteria	<i>Salmonella typhimurium</i> LT2	0,4	4692	47,76%	4951371	2586532	2364839	1	0
Bacteria	<i>Shewanella amazonensis</i> SB2B		3702	46,38%	4264533	2286549	1977984	1	0
Bacteria	<i>Shewanella baltica</i> OS155		4458	53,75%	5084318	2351473	2732845	1	0
Bacteria	<i>Shewanella denitrificans</i> OS217		3840	54,87%	4491474	2026975	2464499	1	0
Bacteria	<i>Shewanella frigidimarina</i> NCMB400		4124	58,43%	4782427	1988048	2794379	1	0
Bacteria	<i>Shewanella oneidensis</i> MR-1	0,66	4908	54,11%	5131416	2354767	2776649	1	0
Bacteria	<i>Shewanella putrefaciens</i> CN-32		4095	55,56%	4577697	2034233	2543464	1	0
Bacteria	<i>Shewanella</i> sp. ANA-3		4353	52,04%	5100729	2446144	2654585	1	0
Bacteria	<i>Shewanella</i> sp. MR-4		3928	52,07%	4579182	2194734	2384448	1	0
Bacteria	<i>Shewanella</i> sp. MR-7		3920	52,08%	4546355	2178696	2367659	1	0
Bacteria	<i>Shewanella</i> sp. PV-4		3811	46,24%	4474526	2405423	2069103	1	0
Bacteria	<i>Shewanella</i> sp. W3-18-1		4241	55,35%	4754010	2122432	2631578	1	0
Bacteria	<i>Shigella boydii</i> sv4 Sb227		4403	48,89%	4646520	2374795	2271725	1	0
Bacteria	<i>Shigella dysenteriae</i> sv1 Sd197		4607	49,01%	4551958	2321087	2230871	1	0
Bacteria	<i>Shigella flexneri</i> 2a 2457T		4195	49,09%	4599354	2341336	2258018	1	0
Bacteria	<i>Shigella flexneri</i> 2a 301		4822	49,35%	4828821	2446008	2382813	1	0
Bacteria	<i>Shigella sonnei</i> Ss046		4581	49,23%	5039661	2558405	2481256	1	0
Bacteria	<i>Silicibacter pomeroyi</i> DSS-3		4314	35,93%	4601053	2947878	1653175	0	1
Bacteria	<i>Silicibacter</i> sp. TM1040		3932	40,13%	4198271	2512579	1685692	1	0
Bacteria	<i>Sinorhizobium meliloti</i> 1021	1,5	6270	37,83%	6691694	4160497	2531197	1	0
Bacteria	<i>Solibacter usitatus</i> Ellin6076		8157	38,10%	9949872	6158997	3790875	0	1
Bacteria	<i>Sphingopyxis alaskensis</i> RB2256		3247	34,50%	3343420	2189865	1153555	1	0
Bacteria	<i>Staphylococcus aureus</i> COL		2690	67,19%	2813862	923317	1890545	1	0
Bacteria	<i>Staphylococcus aureus</i> MRSA252		2820	67,19%	2902619	952267	1950352	1	0
Bacteria	<i>Staphylococcus aureus</i> MSSA476		2722	67,18%	2820454	925603	1894851	1	0
Bacteria	<i>Staphylococcus aureus</i> MW2		2738	67,20%	2841116	931807	1909309	1	0
Bacteria	<i>Staphylococcus aureus</i> Mu50		2809	67,16%	2903636	953661	1949975	1	0
Bacteria	<i>Staphylococcus aureus</i> N315	0,4	2702	67,19%	2839469	931559	1907910	1	0
Bacteria	<i>Staphylococcus epidermidis</i> ATCC 12228	0,8	2561	67,95%	2564615	821855	1742760	1	0
Bacteria	<i>Staphylococcus epidermidis</i> RP62A		2606	67,85%	2643840	849969	1793871	1	0
Bacteria	<i>Staphylococcus haemolyticus</i> JCSC1435		2813	67,21%	2685015	880486	1804529	1	0
Bacteria	<i>Staphylococcus saprophyticus</i> saprophyticus ATCC 15305		2595	66,81%	2577899	855500	1722399	1	0
Bacteria	<i>Streptococcus agalactiae</i> 2603V/R		2225	64,35%	2160267	770062	1390205	1	0
Bacteria	<i>Streptococcus agalactiae</i> la A909		2097	64,38%	2127839	757988	1369851	1	0
Bacteria	<i>Streptococcus agalactiae</i> NEM316	1,8	2235	64,37%	2211485	787895	1423590	1	0
Bacteria	<i>Streptococcus mutans</i> UA159	0,75	2040	63,17%	2030921	747887	1283034	1	0
Bacteria	<i>Streptococcus pneumoniae</i> R6		2113	60,28%	2038615	809656	1228959	1	0
Bacteria	<i>Streptococcus pneumoniae</i> TIGR4	0,5	2164	60,30%	2160837	857790	1303047	1	0
Bacteria	<i>Streptococcus pyogenes</i> M1 GAS	0,4	1774	61,49%	1852441	713439	1139002	1	0
Bacteria	<i>Streptococcus pyogenes</i> M1 MGAS5005		2017	61,47%	1838554	708475	1130079	1	0
Bacteria	<i>Streptococcus pyogenes</i> M28 MGAS6180		2044	61,65%	1897573	727780	1169793	1	0
Bacteria	<i>Streptococcus pyogenes</i> MGAS10394		1973	61,31%	1899877	734970	1164907	1	0
Bacteria	<i>Streptococcus pyogenes</i> MGAS315		1946	61,41%	1900521	733462	1167059	1	0
Bacteria	<i>Streptococcus pyogenes</i> MGAS8232		1930	61,45%	1895017	730465	1164552	1	0
Bacteria	<i>Streptococcus pyogenes</i> SSI-1		1930	61,45%	1894275	730248	1164027	1	0
Bacteria	<i>Streptococcus suis</i> 89/1591		1938	58,97%	1978218	811614	1166604	1	0
Bacteria	<i>Streptococcus thermophilus</i> CNRZ 1066		2000	60,92%	1796226	701991	1094235	1	0
Bacteria	<i>Streptococcus thermophilus</i> LMD-9		1919	60,97%	1842121	677066	1165055	1	0
Bacteria	<i>Streptococcus thermophilus</i> LMG 18311	0,43	1974	60,91%	1796846	702417	1094429	1	0
Bacteria	<i>Streptomyces avermitilis</i> MA-4680		7761	29,30%	9119895	6447889	2672006	1	1

Bacteria	<i>Streptomyces coelicolor</i> A3(2)	2,2	8299	28,00%	9054847	6519342	2535505	0	1
Bacteria	<i>Sulfitobacter</i> sp. EE-36		3529	39,69%	3547243	2139377	1407866	0	1
Bacteria	<i>Sulfitobacter</i> sp. NAS-14.1		4014	39,95%	4002069	2403046	1599023	0	1
Bacteria	<i>Symbiobacterium thermophilum</i> IAM 14863	4,2	3454	31,33%	3566135	2448808	1117327	1	0
Bacteria	<i>Synechococcus elongatus</i> BP-1		2520	46,08%	2593857	1398554	1195303	0	1
Bacteria	<i>Synechococcus</i> sp. PCC 6301		2576	44,52%	2696255	1495999	1200256	0	1
Bacteria	<i>Synechococcus</i> sp. PCC 7942 (elongatus)		2663	44,53%	2695903	1495418	1200485	0	1
Bacteria	<i>Synechococcus</i> sp. WH8102	6	2576	40,59%	2434428	1446357	988071	0	1
Bacteria	<i>Synechocystis</i> sp. PCC 6803		3614	52,63%	3947019	1869553	2077466	0	1
Bacteria	<i>Syntrophobacter fumaroxidans</i> MPOB		4084	40,02%	4848841	2908179	1940662	0	1
Bacteria	<i>Syntrophomonas wolfei</i> Goettingen		2524	55,16%	2845772	1275930	1569842	0	1
Bacteria	<i>Syntrophus aciditrophicus</i>		3219	48,54%	3179300	1636069	1543231	0	1
Bacteria	<i>Thermoanaerobacter ethanolicus</i> 39E		2311	65,47%	2282740	788337	1494403	0	1
Bacteria	<i>Thermoanaerobacter tengcongensis</i> MB4	1,1	2652	62,43%	2689445	1010449	1678996	0	1
Bacteria	<i>Thermobifida fusca</i> YX		3182	32,50%	3642249	2458622	1183627	0	1
Bacteria	<i>Thermotoga maritima</i> MSB8	1,2	1895	53,75%	1860725	860543	1000182	0	1
Bacteria	<i>Thermus thermophilus</i> HB27	2,5	2263	30,59%	2127482	1476627	650855	0	1
Bacteria	<i>Thermus thermophilus</i> HB8		2042	30,48%	1859064	1292334	566730	0	1
Bacteria	<i>Thiobacillus denitrificans</i> sp. ATCC 25259		2876	33,93%	2909809	1922430	987379	1	0
Bacteria	<i>Thiomicrospira crunigena</i> XCL-2		2254	56,87%	2427734	1047081	1380653	1	0
Bacteria	<i>Thiomicrospira denitrificans</i> ATCC 33889		2160	65,54%	2201561	758701	1442860	0	1
Bacteria	<i>Treponema denticola</i> ATCC 35405	5	2817	62,13%	2843201	1076827	1766374	0	1
Bacteria	<i>Treponema pallidum</i> Nichols	33	1082	47,23%	1138011	600568	537443	0	1
Bacteria	<i>Trichodesmium erythraeum</i> IMS101		4801	64,72%	8191365	2886300	5305065	0	1
Bacteria	<i>Tropheryma whippelii</i> TW08/27		838	53,69%	925938	428839	497099	0	1
Bacteria	<i>Tropheryma whippelii</i> Twist		861	53,67%	927303	429638	497665	0	1
Bacteria	<i>Ureaplasma parvum</i> serovar 3 ATCC 700970	0,9	647	74,50%	751719	191686	560033	0	0
Bacteria	<i>Vibrio cholerae</i> O1 biovar eltor N16961	0,2	3951	52,51%	4033464	1915379	2118085	1	0
Bacteria	<i>Vibrio fischeri</i> ES114	0,3	3903	61,65%	4238201	1625497	2612704	1	0
Bacteria	<i>Vibrio parahaemolyticus</i> RIMD 2210633	0,2	4992	54,63%	5165770	2343877	2821893	1	0
Bacteria	<i>Vibrio vulnificus</i> CMCP6	0,16	4676	53,31%	5126798	2393740	2733058	1	0
Bacteria	<i>Vibrio vulnificus</i> YJ016	0,16	5168	53,32%	5260086	2455465	2804621	1	0
Bacteria	<i>Wigglesworthia glossinidia</i> endosymbiont of <i>Glossina</i>		651	77,52%	697724	156841	540883	1	0
Bacteria	<i>Wolbachia</i> endosymbiont of <i>Drosophila ananassae</i>		1833	64,29%	1440750	514492	926258	0	1
Bacteria	<i>Wolbachia</i> endosymbiont of <i>Drosophila simulans</i>		782	64,60%	1063100	376358	686742	0	1
Bacteria	<i>Wolbachia</i> endosymbiont strain TRS of <i>Brugia malayi</i>	14	842	65,82%	1080084	369163	710921	0	1
Bacteria	<i>Wolbachia pipientis</i> wMel		1232	64,77%	1267782	446701	821081	0	1
Bacteria	<i>Wolinella succinogenes</i> DSM 1740	1	2090	51,54%	2110355	1022636	1087719	0	1
Bacteria	<i>Xanthomonas axonopodis</i> citri 306	7	4487	35,29%	5274174	3413008	1861166	1	0
Bacteria	<i>Xanthomonas campestris</i> campestris 8004	3	4389	35,04%	5148708	3344448	1804260	1	0
Bacteria	<i>Xanthomonas campestris</i> campestris ATCC 33913		4241	34,93%	5076188	3302992	1773196	1	0
Bacteria	<i>Xanthomonas campestris</i> vesicatoria 85-10		4786	35,44%	5420152	3499053	1921099	1	0
Bacteria	<i>Xanthomonas oryzae</i> KACC10331	2	4697	36,31%	4941439	3147365	1794074	1	0
Bacteria	<i>Xylella fastidiosa</i> 9a5c	96	2888	47,38%	2731750	1437372	1294378	1	0
Bacteria	<i>Xylella fastidiosa</i> Ann-1		2411	47,97%	2622328	1364351	1257977	1	0
Bacteria	<i>Xylella fastidiosa</i> Dixon		4757	48,06%	5118933	2658905	2460028	1	0
Bacteria	<i>Xylella fastidiosa</i> Temecula1		2091	48,22%	2521148	1305386	1215762	1	0
Bacteria	<i>Yersinia pestis</i> CO92	1,25	4307	52,36%	4829855	2301052	2528803	1	0
Bacteria	<i>Yersinia pestis</i> KIM		4187	52,36%	4600755	2191853	2408902	1	0
Bacteria	<i>Yersinia pestis</i> biovar Medievalis 91001		4237	52,33%	4803217	2289686	2513531	1	0
Bacteria	<i>Yersinia pseudotuberculosis</i> IP 32953	0,5	4226	52,45%	4840899	2301869	2539030	1	0
Bacteria	<i>Zymomonas mobilis</i> ZM4		2086	53,67%	2056416	952703	1103713	1	0

### III. Discussion

#### A. La thymidylate synthase *ThyX* limite la vitesse de réplication d'*Escherichia coli*, un organisme utilisant naturellement *ThyA*.

Au cours de ce travail, j'ai pu montrer que le remplacement non-orthologue de *ThyA* par *ThyX* est contre-sélectionné dans un organisme tel qu'*E. coli*. La souche FE010 ( $\Delta thyA::thyX$ ) cultivé en absence de thymidine présente des phénotypes caractéristiques d'une limitation en dTTP : taux de croissance réduit, perte de viabilité, filamentation des cellules, augmentation du taux d'apparition de mutations spontanées. Une hypothèse pour expliquer ce résultat aurait pu être que la protéine *ThyX* n'était pas correctement exprimée dans cette souche. Afin de tester cette idée, j'ai complété la souche FE013 ( $\Delta thyA$ ) avec le vecteur d'expression inductible pQE80-L (Qiagen) portant le gène *thyX* de *Borrelia hermsii* sous le contrôle du promoteur T5. J'ai pu déterminer qu'à partir d'une concentration de 67  $\mu\text{M}$  d'IPTG dans le milieu de culture, on n'observe plus de phénomène de filamentation (Figure 15).

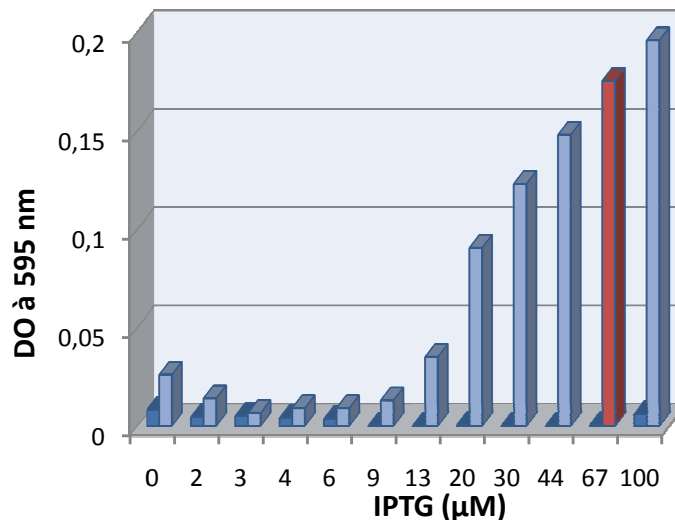


Figure 15 : Effet de l'induction de *thyX* sur la croissance de la souche d'*E. coli* FE013 ( $\Delta thyA$ ). La souche témoin FE013/pQE80-L (bleu foncé) et la souche FE023 ( $\Delta thyA$ , *thyX*<sup>+</sup> *trans*) (bleu clair) ont été cultivées en milieu minimum M9 sans thymidine à des concentrations croissances d'inducteur (IPTG). La croissance a été quantifiée par la mesure de la DO à 595nm après 16h de culture. La barre rouge correspond à la quantité d'IPTG nécessaire pour obtenir une morphologie sauvage.

Il est intéressant de noter qu'à cette concentration, l'analyse par bio-analyseur (Experion, Bio-Rad) des protéines contenues dans un extrait brut de cette souche a montré que la protéine ThyX représente près d'un tiers du protéome total. De plus, les expériences de mesure de la synthèse d'ADN par incorporation de précurseur marqué (article II), ont permis de montrer que cette souche présentait toujours un défaut de réplication. Ceci indique que le phénomène observé ici est indépendant du niveau d'expression de la protéine. L'ensemble de ses résultats montre que la faible efficacité catalytique de ThyX ne permet pas un approvisionnement en dTMP suffisamment rapide pour permettre la correcte réplication chez *E. coli*.

***B. Les organismes procaryotes utilisant ThyX répliquent lentement leur ADN et ont une croissance lente.***

Nous avons vu dans l'introduction que les bactéries ont généralement une seule origine de réplication par chromosome et seulement deux fourches actives par cycle de réplication. Ainsi toute accumulation d'ADN conduit à un allongement de la durée de réplication et peut par conséquent ralentir le taux de multiplication rendant l'organisme moins compétitif. Les bactéries sont donc sous une pression de sélection constante afin de limiter l'accumulation d'ADN non-codant. Il paraît alors que la correcte réplication du chromosome est l'une des étapes majeures pour qu'un organisme procaryote se développe, persiste et reste compétitif. En d'autres termes, la géométrie du chromosome limite la vitesse de réplication qui va elle-même limiter l'expansion des génomes procaryotes. Dans l'article II, nous avons établi que l'efficacité catalytique de l'étape de synthèse du dTMP, précurseur essentiel de l'ADN, limite également la vitesse de réplication chez les bactéries. Cette idée n'est pas complètement nouvelle puisque, comme nous l'avons vu, l'inhibition de cette étape est une stratégie de lutte contre la prolifération de bactéries pathogènes ou de cellules cancéreuses depuis longtemps utilisée [introduction p 26]. Dans ce cas-là, on crée une limitation artificielle de la réplication de l'ADN qui conduit à la mort cellulaire. Ce qui est nouveau dans notre analyse, c'est que nous avons montré que cette limitation a lieu naturellement chez les procaryotes et qu'elle est dépendante du type de thymidylate synthase utilisée. La mesure *in vivo* chez un

organisme utilisant naturellement ThyX, *H. pylori*, ainsi que les données issues de la littérature ont permis de mettre en évidence que la vitesse de réplication chez les procaryotes n'est pas uniforme. Les organismes utilisant la thymidylate synthase la moins active – ThyX – montrent une vitesse de réplication plus lente, un taux de croissance plus réduit que ceux qui utilisent le couple ThyA/FolA. L'ensemble de ces résultats indiquent que l'approvisionnement en l'une des quatre matières premières va limiter l'étape de fabrication de la molécule d'ADN.

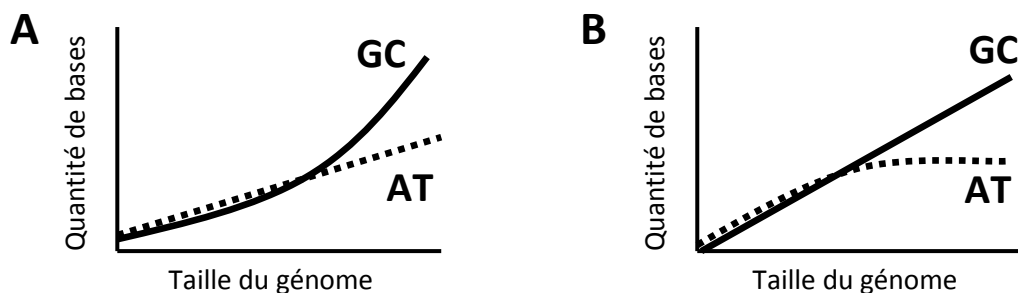
### ***C. La thymidylate synthase ThyA est nécessaire à la réplication de grands génomes***

La vitesse de réplication ( $V$ ) est le quotient de la taille du génome, ou quantité d'ADN à répliquer ( $Q$ ), sur le temps de réplication ( $TR$ ). Cette relation peut également s'écrire ainsi :  $TR = Q/V$ . Or nous savons que chez un organisme procaryote, l'augmentation du temps de réplication ( $TR$ ) va limiter sa capacité à croître et le rendre moins compétitif. En d'autres termes, pour être compétitif, un organisme doit limiter le rapport  $Q/V$ . Dans le cas d'un organisme possédant *thyX* nous avons vu que la valeur  $V$  est basse, par conséquent pour  $Q$  constant,  $TR$  va être élevé. Comme nous l'avons vu, ceci représente un coût pour la cellule. Une possibilité pour contourner ce problème serait de réduire également la valeur de  $Q$ . On peut alors prédire que les organismes utilisant la thymidylate synthase ThyX devraient globalement avoir un génome petit. Afin de tester cette idée, j'ai examiné les données génomiques de plus de 400 organismes procaryotes. Cette analyse m'a permis de révéler que le génome des organismes utilisant la thymidylate synthase ThyX était majoritairement restreint à des petites tailles, contrairement à celui des organismes ThyA. Les données analysées sont donc compatibles avec le modèle proposé. On peut également supposer que la limitation de la réplication chez les organismes utilisant ThyX, restreint leur capacité à acquérir de l'ADN et ainsi limite la taille de leur génome. Ainsi un organisme possédant un large génome devra minimiser la durée de la période de réplication en augmentant la vitesse de réplication. Les résultats

présentés dans l'article II suggèrent que cette augmentation nécessite un mécanisme efficace de synthèse du dTMP catalysé par le couple ThyA/FolA.

#### ***D. La cinétique de synthèse du thymidylate limite l'expansion des génomes procaryotes***

Lors de ce travail, j'ai également réexaminé la relation positive décrite entre la taille du génome et le GC% (Figure 7, p 42). La lecture classique de cette courbe est qu'à mesure que les génomes s'élargissent, ils deviennent relativement plus riches en bases G et C qu'en bases A et T. Cet enrichissement peut s'expliquer de deux façons : soit l'augmentation de AT est linéaire et la quantité de GC augmente plus rapidement (Figure 16 cas A) ; soit à l'inverse l'augmentation de GC est linéaire et la quantité de AT tend à se stabiliser (Figure 16 cas B). J'ai mis en évidence que c'est cette dernière solution qui explique la tendance générale observée.



**Figure 16 : Deux scénarii possibles expliquant la corrélation positive entre la taille du génome et la quantité de bases.**

Ainsi la quantité de bases AT dans les génomes apparaît être un facteur limitant à partir d'une certaine taille de génome (> 4Mb). Considérant le rôle majeur de l'ATP dans la cellule, il paraît peu probable que la disponibilité en adénine soit en cause, il est donc plus vraisemblable que ce soit la disponibilité en thymidine qui limite l'ajout de bases AT. Je propose que cette limitation en thymidine résulte d'une relative inefficacité des thymidylate synthases à alimenter la machinerie de réplication en thymidine assez rapidement. Logiquement on observe alors une augmentation du GC% à mesure que les génomes grossissent. Le GC% est lui-même limité à environ 70% probablement car le dépassement de cette valeur induirait un biais trop important pour certains acides aminés

(appauvrissement en phénylalanine, tyrosine, isoleucine, asparagine et lysine), et augmenterait l'énergie nécessaire pour dérouler la molécule d'ADN. Il apparaît ainsi que l'efficacité de synthèse du thymidylate et la limitation du GC% représentent des forces antagonistes. Je propose que l'apparente limite de taille des génomes procaryotes à environ 10 Mpb soit, au moins en partie, sous l'influence de ces deux forces.

## Chapitre III. Etude de deux étapes clés du métabolisme des pyrimidines chez *H. pylori*

Les travaux qui vont être présentés dans ce chapitre sont le fruit du travail que j'ai effectué chez la bactérie *H. pylori*. On peut séparer ce travail en deux parties. En premier lieu, je vais présenter les travaux que j'ai effectués au début de ma thèse : il s'agissait alors de mieux comprendre quels étaient les processus cellulaires dans lesquels la protéine ThyX était impliquée. La deuxième partie de ce chapitre est consacrée à l'étude des premières étapes de synthèse des pyrimidines et plus particulièrement la synthèse de l'uridine monophosphate (UMP).

### I. La synthèse du dTMP

#### A. Introduction

Historiquement, la première thymidylate synthase de la famille X dont l'activité de synthèse de dTMP a été prouvée *in vitro* fut celle de la bactérie *H. pylori* (Mylykallio *et coll.*, 2002). L'analyse du génome de la bactérie, outre la présence du gène *hp1533* codant ThyX, nous indiquait l'absence de gène homologue à *thyA*, *folA* et *tdk* codant respectivement la thymidylate synthase ThyA, la dihydrofolate réductase (DHFR) et la thymidine kinase (TK). Ceci indiquait que la synthèse *de novo* du dTMP devait être sous la seule dépendance de ThyX, et que sa formation à partir de thymidine exogène devait être impossible. Des données biochimiques antérieures avaient montré que *H. pylori* n'était pas capable d'incorporer de thymidine exogène (Menz *et coll.*, 1994). Considérant ces informations, ThyX apparaît comme un « talon d'Achille » dans la synthèse des pyrimidines chez cet organisme. Compte tenu de l'absence de ThyX chez l'Homme et les différences mécanistiques et structurales des deux familles de thymidylate synthases, ThyX représente une cible de choix pour inhiber la prolifération de *H. pylori* (voir introduction p 27). Afin de prouver expérimentalement les prédictions ci-dessus, j'ai entrepris de construire une souche de *H. pylori* dans laquelle l'expression de ThyX serait conditionnée à la présence d'une molécule inductrice.



Comme nous l'avons vu dans l'introduction, rien n'était connu sur la régulation de l'expression de ThyX. On pouvait s'interroger en particulier sur la possibilité que ThyX puisse « communiquer » avec des protéines impliquées dans le métabolisme des nucléotides, ou dans un autre processus cellulaire comme la réplication de l'ADN par exemple. Une façon d'aborder cette question est de rechercher les partenaires d'interactions de la protéine. A l'époque, les seules données disponibles concernant d'éventuels partenaires de ThyX étaient issues de la carte d'interaction protéine-protéine de *H. pylori* réalisée par double-hybride (Rain *et coll.*, 2001). La nature des protéines identifiées comme potentiellement interagissant avec ThyX (GroEL : Chaperonne et protéine de choc thermique ; PheT : phénylalananyl-tRNA synthétase sous-unité bêta ; HemG : S-adenosylméthionine-dépendante méthyltransférase) étant peu informative, j'ai alors entrepris des expériences de recherche de complexes protéiques *in vivo* et d'interaction protéine-protéine *in vitro*.

Les travaux que je vais vous présenter n'auraient pas été possibles sans l'assistance scientifique et technique de Hilde de Reuse (Institut Pasteur) qui m'a fourni les différentes souches de *H. pylori* et outils génétiques (alors non publiés) utilisés dans cette étude. Elle m'a aimablement accueilli dans son laboratoire où j'ai effectué les expériences de « TAP-tag » grâce à l'expertise de Kerstin Stingl.

## **B. Matériels et Méthodes**

### **a. Souches bactériennes et conditions de croissances**

*Helicobacter pylori* 26695 et N6 ont été cultivées à la fois sur milieu solide « blood agar base2 » supplémenté avec du sang de cheval 10% (Oxoid) ; en milieux liquides « Brain Heart Infusion » supplémenté avec du Sérum de Veau Fœtal 10% (GibcoBRL) ou « Brucella » supplémenté de  $\beta$ -cyclodextrine 10% (Sigma). Un mélange antibiotiques-fongicide composé de vancomycine ( $10 \mu\text{g.ml}^{-1}$ ), polymyxine ( $2,5 \text{ IU.l}^{-1}$ ), triméthoprime ( $5 \mu\text{g.ml}^{-1}$ ) et fungizone ( $2,5 \mu\text{g.ml}^{-1}$ ) a été utilisé. Les boîtes ont été incubées à  $37^\circ\text{C}$  en condition microaérobie dans des jarres CampyGen (Oxoid). Les

cultures liquides ont été agitées à 175 tours par minutes. Lorsque nécessaire l'isopropyl-beta-D-thiogalactopyranoside (IPTG) (Sigma) a été utilisée à la concentration de 1 mM.

Les souches d'*Escherichia coli* ont été cultivées à 37°C dans du bouillon Luria Bertani (LB) ou dans du milieu minimum M9. Les antibiotiques et suppléments ont été ajoutés, lorsque nécessaires, aux concentrations suivantes : ampicilline 100 µg.ml<sup>-1</sup> (Sigma), IPTG 1 mM (Sigma), uridine 40 µg.ml<sup>-1</sup> (Fluka), uracile 40 µg.ml<sup>-1</sup> (Fluka), tétracycline 5 µg.ml<sup>-1</sup> (Sigma), thiamine 5 µg.ml<sup>-1</sup> (Sigma), biotine 1 µg.ml<sup>-1</sup> (Sigma), 5-fluorouracile 20 µg.ml<sup>-1</sup> (Sigma). La souche *E. coli* HB101 a été utilisée comme bactérie hôte pour la construction des plasmides.

#### b. Techniques moléculaires

Les réactions de PCR ont été réalisées en utilisant la Phusion High Fidelity DNA Polymerase (Finnzymes) sur l'ADN génomique de *H. pylori* 26695 et N6. Les séquences génomiques ont été extraites du navigateur de génomes GenoList (<http://genolist.pasteur.fr/>) PyloriGene. La cassette de délétion pour le gène *hp1533* et les plasmides utilisés ont été construits comme décrit dans Résultats et Discussion. Les amorces, souches et plasmides sont présentés dans les tableaux ci-dessous.

Tableau 6 : Listes des souches et plasmides utilisées dans ce travail.

Nom	Origine	Plasmide	Informations ou génotypes pertinents	Référence
<b>Plasmides</b>				
pILL851	pILL570		cassette <i>tap-aphA-3</i> cloné entre les sites ClaI et PstI	Stingl, K.
pHM0702	pILL851		<i>hp1533-tap-aphA-3</i>	ce travail
pILL2150		pHeL2	Expression inductible à l'IPTG de ThyX	(Boneca <i>et coll.</i> , 2008)
pILL2157		pHeL2	Expression inductible à l'IPTG de ThyX	(Boneca <i>et coll.</i> , 2008)
pHM0700		pILL2150	Expression inductible à l'IPTG de ThyX	ce travail
pHM0701		pILL2157	Expression inductible à l'IPTG de ThyX	ce travail
<b>Souches bactériennes</b>				
<i>H. pylori</i>				
26695			sauvage	(Tomb <i>et coll.</i> , 1997)
N6			sauvage	souche de laboratoire
FE008		26695	<i>hp1533-tap-aphA-3</i>	ce travail
<i>E. coli</i>				
HB101			sauvage	souche de laboratoire

Tableau 7 : Listes des amorces utilisées dans ce travail.

gene ciblé	Nom	Restriction	Séquence
<i>hp1533</i> (clonage)	oEF001 (f)	NdeI	agtcaacatatggaagtgattgtaagca
	oEF002 (r)	BamHI	attgtggatccctaatagcttcaacaatctt
<i>hp1533</i> (délétion)	oEF003 (f)		gaaagtagagcctaattgcatgaaaaaat
	oEF004 (r)		gttagtcacccgggtactatttgactcctatatcgctt
	oEF007 (f)		tacctggagggaataatgatttaaaaaataccg
	oEF008 (r)		gtcaaggacaaccgctat
<i>hp1533</i> (tap-tag)	oEF009 (f)		ctctttccatggatccatgcttcaacaatcttcaacaataac
	oEF010 (r)	PstI	aactgcaggataaagaatgatccacagggtgggg
	oEF026 (f)		tagtacctggagggaataatgacaataaccgcttaatagctcattc
	oEF027 (r)	Clal	ccatcgatcaatcaataataaaagagcttaattttaac

### c. « Tandem Affinity Purification » de protéines interagissant avec ThyX

Un culot bactérien correspondant à 2 litres d'une culture de FE008 (16 h) a été suspendu dans 20 ml d'un tampon aqueux composé de glycérol 8% ; NaCl 0,1 M ; Tris-HCl 0,1 M (pH7,4) ; additionné d'un cocktail d'inhibiteurs de protéases (Roche). Les cellules ont été lysées à haute pression à l'aide d'une presse de French. Les débris cellulaires ont été éliminés par centrifugation (22000 g à 4°C). Le surnageant a été ensuite incubé 2 h à 4°C sous lente agitation avec 0,2 ml de Sépharose-IgG (GE Healthcare) en présence de 0,1% NP-40. Après lavage avec 10 ml de tampon de fixation (Tris-HCl 50 mM pH 7,4 ; 0,1 M NaCl ; 0,1% NP-40 ; 0,5 mM DTT), les billes de Sépharose ont été incubées 2 h à 16°C en présence de protéase TEV (100 U dans 250 µl). Après élution, l'extrait protéique est incubé 1 h à 4°C en présence de billes de calmoduline (Stratagene) et de 2 mM de CaCl<sub>2</sub>. Après lavage avec le tampon de fixation additionné de 2 mM de CaCl<sub>2</sub>, les protéines ont été éluées avec 400 µl de tampon d'élution (Tris-Hcl 20 mM pH 8; NaCl 50 mM ; EGTA 5 mM).

## C. Résultats et Discussion

### a. Tentative d'inactivation conditionnelle du gène thyX

La stratégie que nous avons retenue pour prouver l'essentialité du ThyX chez *H. pylori* peut se décomposer en trois étapes. Dans un premier temps, on doit apporter à la bactérie un plasmide répliquatif portant le gène *thyX* sous le contrôle d'un promoteur inductible. Dans un second temps, en

présence d'inducteur, on inactive la copie chromosomique par remplacement allélique à l'aide d'un fragment linéaire d'ADN portant une cassette de résistance à la kanamycine flanquée des régions amont et aval du gène ciblé. Enfin, la souche ainsi obtenue sera cultivée en absence d'inducteur : la perte de viabilité indiquant l'essentialité du gène.

La première étape a été d'amplifier par PCR le gène *hp1533* codant la protéine ThyX à partir de l'ADN chromosomique de la souche séquencée 26695. Le fragment obtenu a été cloné parallèlement dans deux vecteurs navettes *E. coli-H. pylori*, pILL2150 et pILL2157 (Boneca *et coll.*, 2008), tous deux dérivés du plasmide pHeL2 (Heuermann et Haas, 1998). Les deux plasmides expriment constitutivement le répresseur LacI<sup>q</sup> sous le contrôle du promoteur du gène *amiE*. L'expression du gène d'intérêt est contrôlée par le promoteur *pTac* dans le plasmide pILL2150, et par une version modifiée du promoteur du gène *ureI* contenant deux sites de fixation pour LacI<sup>q</sup> entre le site de fixation du ribosome et le codon de démarrage ATG. Ainsi le plasmide pILL2150, finement régulé par l'IPTG, permet une faible expression bien contrôlée, alors que pILL2157 permet une forte expression. Les plasmides pHM0700 et pHM0701, dérivés de pILL2150 et pILL2157 respectivement, ont été introduits dans la souche de *H. pylori* N6 par transformation naturelle, donnant les souches FE014 et FE009 respectivement. Le choix de cette souche, plutôt qu'une des deux souches séquencées 26695 et J99, se justifie par la capacité qu'à cette souche à correctement maintenir les plasmides réplcatifs contrairement à ces dernières (communication personnelle de Hilde de Reuse).

Parallèlement, j'ai entrepris la synthèse d'un fragment linéaire d'ADN correspondant à une cassette de résistance à la kanamycine flanquée des régions amont et aval du gène *thyX* à partir de l'ADN chromosomique de la souche N6 par la technique de PCR, dite « 3-step PCR » (Derbise *et coll.*, 2003) [pour plus d'informations voir le site du groupe de JM Ghigo à l'Institut Pasteur : <http://www.pasteur.fr/recherche/unites/Ggb/>]. J'ai dessiné les amorces oligonucléotidiques à partir de la séquence de la souche 26695. J'ai pu amplifier la région amont du gène mais je n'ai obtenu aucune amplification pour la région aval. En comparant, les régions bornant le gène *thyX* dans les

deux souches séquencées, je me suis aperçu que la région aval du gène était variable. J'ai alors cartographié cette région par PCR en utilisant une amorce directe sortante dans la région 3' de *thyX* et une amorce inverse dans la première ORF commune rencontrée dans les deux souches séquencées. Après séquençage, j'ai ainsi pu obtenir la séquence de la région aval de *thyX* dans la souche N6 sur environ 400 pb. Les étapes de construction du fragment de délétion sont schématisées sur la Figure 17.

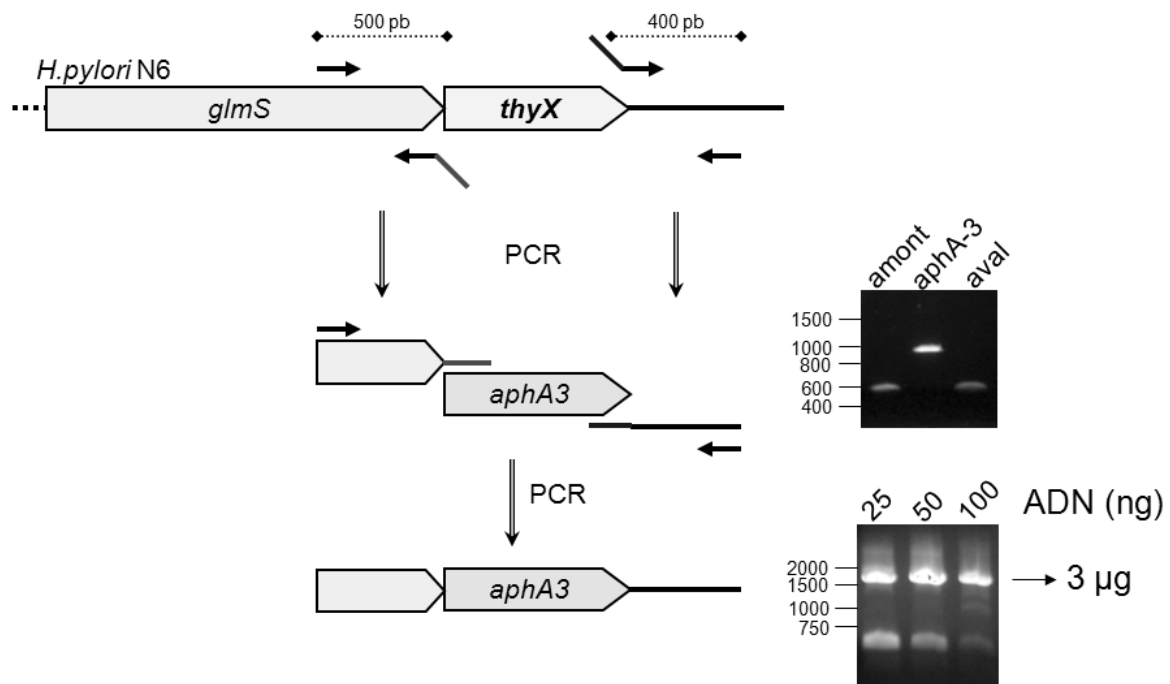


Figure 17 : Différentes étapes de construction du fragment de délétion du gène *thyX*.

Le fragment ainsi obtenu après deux réactions successives de PCR a été introduit par transformation naturelle dans les souches FE014 et FE009 cultivées en milieu riche additionné d'IPTG. Après 5 jours, une vingtaine de colonies par boîte étaient visibles en présence de kanamycine. J'ai testé ensuite leur croissance en absence d'IPTG afin de tester l'essentialité du gène *thyX*. De façon surprenante toutes les souches testées ont été capables de pousser en présence de kanamycine et absence d'IPTG que ce soit sur milieu solide [Blood Agar base 2, Sérum de Veau Foetal (SVF) 10%] ou en culture liquide [Brain Heart Infusion (BHI), SVF 10% et Brucella,  $\beta$ -cyclodextrine]. J'ai alors voulu vérifier si la délétion était correcte. Les résultats obtenus par PCR montraient une

situation aberrante. J'ai soupçonné la présence d'une duplication du gène *thyX* ailleurs sur le chromosome. J'ai alors réalisé l'inactivation du gène directement dans la souche parentale N6 en suivant la même procédure. Dans ce cas, j'ai également obtenu des clones résistants à la kanamycine. J'ai extrait l'ADN chromosomique de plusieurs clones et j'ai utilisé un couple d'oligonucléotides afin d'amplifier un fragment interne du gène *thyX*. Cette dernière expérience m'a permis de constater que malgré l'insertion de la cassette de sélection, une copie du gène était toujours présente ailleurs sur le chromosome.

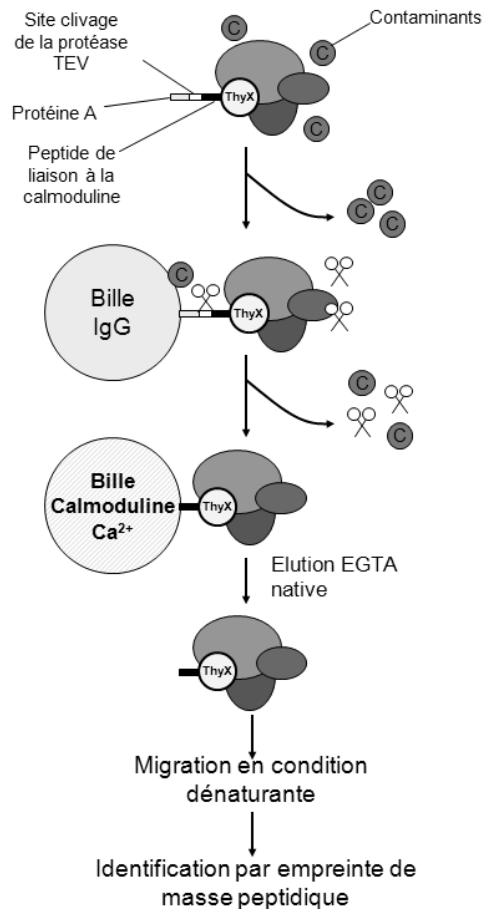
Plusieurs hypothèses peuvent expliquer ce résultat. Le génome de la souche N6 pourrait posséder originellement deux gènes codant ThyX. Cette situation paraît très fortement improbable considérant que sur 148 génomes séquencés portant *thyX*, aucun d'entre eux n'a plus d'une copie. Une autre hypothèse pourrait être que c'est la localisation de *thyX* sur le chromosome qui gêne la correcte délétion. En effet si on place les gènes codant ThyX dans les souches 26695 et J99, on se rend compte qu'ils sont respectivement à 56 et 81 kb de l'origine de réplication (Hallin et Ussery, 2004). Il est envisageable que des contraintes s'exercent sur cette région du chromosome induisant des réarrangements chromosomiques. Dans tous les cas, il n'a pas été possible de créer un mutant conditionnel du gène *thyX* chez *H. pylori*. Malgré tout, les résultats obtenus chez *R. capsulatus* (article I) et l'absence d'incorporation de thymidine exogène suggèrent fortement que *thyX* est un gène essentiel chez *H. pylori*.

#### b. Recherche de protéines interagissant avec la thymidylate synthase X.

##### 1. « Tandem Affinity Purification »

La deuxième partie de ce projet était de rechercher des protéines pouvant interagir avec ThyX. La première stratégie retenue fut une méthode de purification de complexes protéiques dite de "purification en tandem" (Tandem Affinity Purification) ou TAP-Tag développée par Rigault et collaborateurs et adaptée par Kerstin Stingl à *H. pylori* (Puig *et coll.*, 2001; Rigault *et coll.*, 1999). Cette méthode permet de purifier un complexe en conditions natives en maintenant le niveau d'expression

physiologique de la protéine. Cette méthode repose sur une double purification par affinité. La ci-dessous présente le principe général de la méthode "TAP-Tag".



**Figure 18 :** Représentation schématique des différentes étapes de la purification de complexe protéique par la technique du TAP-tag.

Le marqueur est constitué de trois éléments fusionnés permettant la double purification. Il comporte un peptide de liaison à la calmoduline lié à la partie C-terminale de la protéine ; un site de clivage aux protéases TEV (Tobacco Etch Virus) et enfin la protéine A. La méthodologie consiste à fusionner l'étiquette à la région 3' du gène directement sur le chromosome (Figure 18). L'expression de la protéine recombinante est donc conservée à son taux naturel. Il est alors possible de purifier la protéine et ses partenaires à partir d'extraits cellulaires par purification d'affinité sur des billes d'IgG qui vont fixer la protéine A. Après des étapes de lavages, le complexe peut être libéré en ajoutant la protéase TEV qui coupera spécifiquement au niveau du site de clivage de l'étiquette. Le complexe élué est purifié une seconde fois par passage sur une colonne de billes de calmoduline par

reconnaissance de la partie peptidique liant la calmoduline (CBP). Le complexe natif est finalement purifié par élution en conditions non dénaturantes.

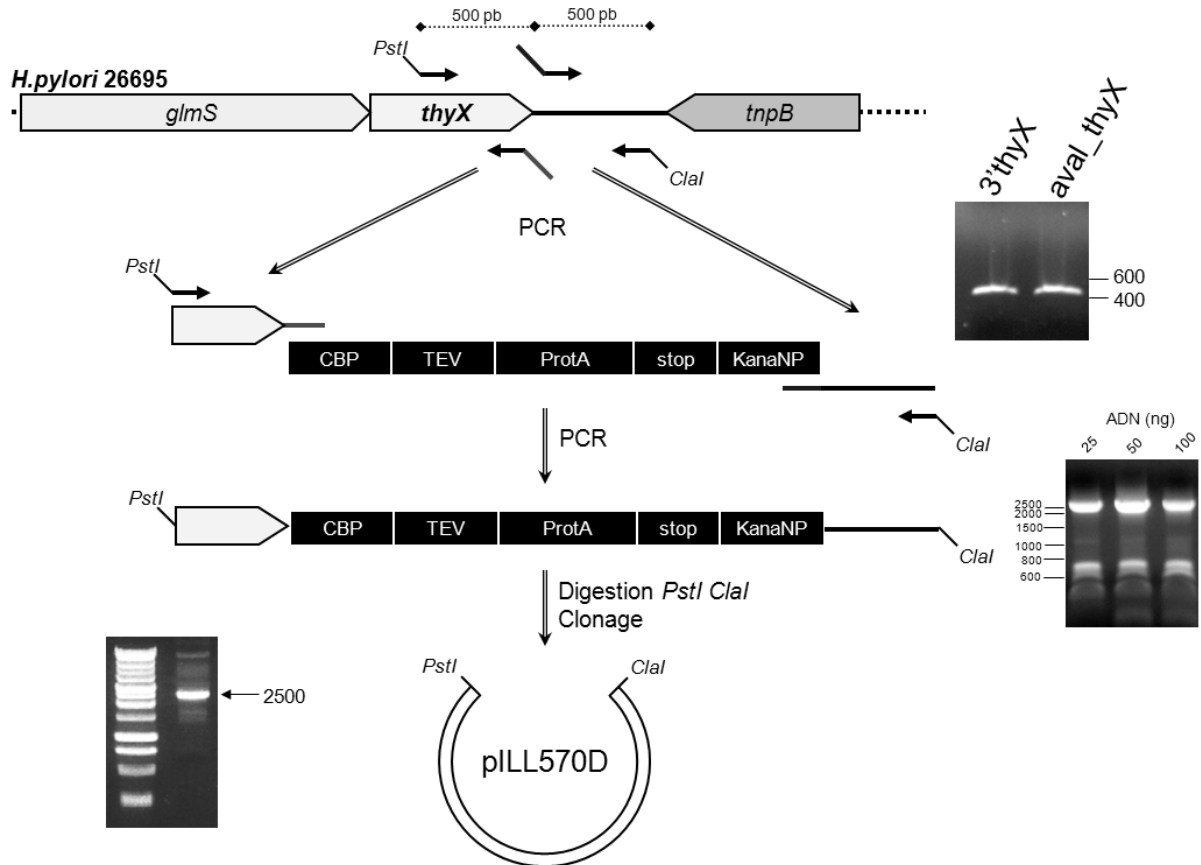
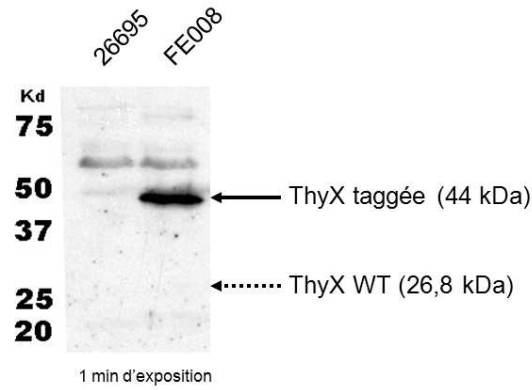


Figure 19 : Différentes étapes de construction du plasmide portant la fusion du gène *thyX* et de l'étiquette TAP. CBP : Calmoduline Binding Protein ; TEV : site de coupure de la protéase TEV ; ProtA : protéine A ; KanaNP : cassette *aphA-3* non polaire.

En utilisant cette approche, j'ai pu construire la souche FE008 portant l'insertion de l'étiquette TAP en 3' du gène *thyX* (détails expérimentaux dans Matériel et Méthodes).



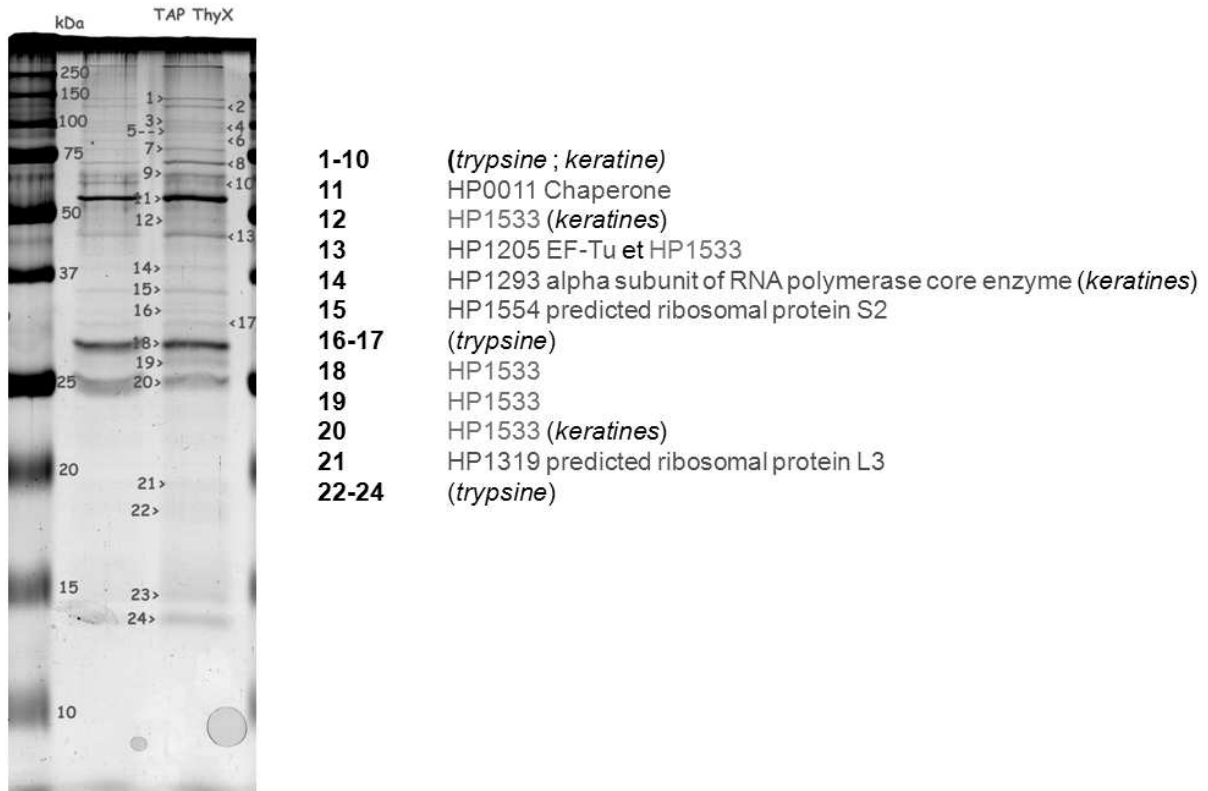


Culture 24h DO<sub>600</sub> 1.3  
 Anticorps I<sup>R</sup> : anti-proteinA SIGMA (1/5000)  
 Anticorps II<sup>R</sup> : anti-rabbit peroxydase (1/20000)

**Figure 20 : Immuno-blot sur des extraits cellulaires des souches 26695 et FE008 à l'aide d'un anticorps ciblant la protéine A et par conséquent l'étiquette TAP.**

Afin de vérifier si l'étiquette n'était pas dégradée dans la cellule, j'ai réalisé un immuno-blot sur des extraits bruts des souches 26695 (sauvage) et FE008 en utilisant un anticorps primaire dirigé contre la protéine A. La Figure 20 montre la présence d'une protéine uniquement présente dans la souche FE008. La taille observée à environ 40 kDa correspond bien à la taille prédite pour la protéine ThyX fusionné à l'étiquette TAP (44 kDa).

Les détails de la démarche expérimentale lors de la purification sont présentés dans Matériels et Méthodes. L'éluât final obtenu à partir d'une culture de deux litres de *H. pylori* FE008 a été séparé sur gel SDS-PAGE afin de séparer ThyX de ses éventuels partenaires (Figure 21). La nature des différentes bandes apparaissant sur le gel a été analysée par la technique d'empreinte peptidique en spectrométrie de masse MALDI-TOF.



**Figure 21 :** Migration sur gel SDS-PAGE de l'éluat protéique issu de la purification TAP. Les différentes bandes ont été identifiées par empreinte peptidique en spectrométrie de masse MALDI-TOF.

L'analyse des résultats a mise en évidence la présence de ThyX (HP1533) dans 5 bandes sur les 24 analysées. Les seules protéines identifiées par la technique sont des protéines classiquement décrites comme des contaminants (Séraphin *et coll.*, 2001). Dans les conditions testées, il n'a pas été possible de détecter par cette approche d'éventuels partenaires de ThyX.

## 2. « Pull-down »

Le problème lié à un résultat négatif est qu'il impossible de conclure. J'ai décidé d'utiliser une approche complémentaire au « TAP » mais qui à la particularité d'être *in vitro* : le « pull-down ». Cette technique utilise une protéine « appât » (ici ThyX) couplée à un support (gel) afin de capturer une protéine « proie » contenue dans un lysat cellulaire. Dans notre cas, nous avons utilisé cette approche pour identifier de nouvelles interactions entre une protéine connue et une protéine cible inconnue. Afin d'utiliser ThyX comme appât, j'ai préalablement biotinyllé la protéine purifiée afin de la fixer par interaction forte à une résine de streptavidine. Le principe de l'expérience est semblable à

celui du TAP. Par cette approche, je n'ai pas pu identifier sur gel SDS-PAGE de protéine spécifiquement liée à ThyX.

Aucune protéine interagissant avec ThyX n'a pu être identifiée en utilisant deux approches différentes. L'interprétation d'un tel résultat doit toujours être faite avec prudence. Nous pouvons expliquer ce résultat de deux façons. 1) La protéine ThyX n'aurait pas de partenaire protéique dans la cellule. Nous savons que la protéine ThyX - seule - est suffisante à la biosynthèse de dTMP et que son activité ne requiert pas la présence d'enzyme DHFR. Métaboliquement parlant, cette enzyme est donc autonome, il est donc envisageable qu'elle ne soit pas impliquée dans un complexe protéique. 2) Il existerait des protéines partenaires de ThyX mais elles n'auraient pas été détectées car leurs interactions avec ThyX seraient faibles et/ou transitoires.

## II. La synthèse de l'UMP

### A. Introduction

Les travaux que je vais présenter maintenant ne concernent pas directement la thymidylate synthase ThyX, mais portent sur une étape précoce du métabolisme des pyrimidines : la synthèse de l'uridine monophosphate (UMP). Ce métabolite a rôle capital car il est le précurseur de tous les ribo- et désoxyribonucléotides pyrimidiques. C'est la lecture de l'article de Mendz et collaborateurs concernant l'étude de la synthèse des nucléotides pyrimidiques chez *H. pylori* qui m'a amené à m'intéresser à la synthèse de l'UMP (Menz *et coll.*, 1994). Comme nous l'avons vu dans l'introduction (voir p 35), les expériences d'incorporation de précurseurs radio-marqués ont montré que la bactérie était incapable d'utiliser la thymidine exogène, mais pouvait pas en revanche incorporer l'uracile. L'absence d'incorporation de la thymidine s'expliquera aisément à la publication du génome de la bactérie (Tomb *et coll.*, 1997) par l'absence de gène *tdk* codant la thymidine kinase. L'analyse de la séquence génomique révéla également, de façon plus surprenante, l'absence des gènes codant les protéines impliquées dans la récupération de l'uracile, et en particulier le gène *upp* codant l'uracile phosphorybosyltransférase (UPRTase). Pourtant, les auteurs ont mesuré une faible

activité UPRTase à partir de lysats bactériens, cette activité correspond à la conversion de l'uracile et du 5-phosphoribosyl- $\alpha$ -1-pyrophosphate (PRPP) en UMP et pyrophosphate. Ces résultats semblaient indiquer qu'une voie peu efficace de récupération non identifiée de l'uracile existait chez *H. pylori*. J'ai décidé de tenter d'identifier cette voie par des approches génétiques et biochimiques.

## **B. Matériels et Méthodes**

### a. Souches bactériennes et conditions de croissance

*Helicobacter pylori* 26695 a été cultivée à la fois sur milieu solide « blood agar base2 » supplémenté avec du sang de cheval 10% (Oxoid) ; et en milieu liquide « Brain Heart Infusion » supplémenté avec du Sérum de Veau Fœtal 10% (GibcoBRL). Un mélange antibiotiques-fongicide composé de vancomycine ( $10 \mu\text{g.ml}^{-1}$ ), polymyxine ( $2,5 \text{ IU.l}^{-1}$ ), triméthoprime ( $5 \mu\text{g.ml}^{-1}$ ) et fungizone ( $2,5 \mu\text{g.ml}^{-1}$ ) a été utilisé. Les boîtes ont été incubées à  $37^{\circ}\text{C}$  en condition microaérobie dans des jarres CampyGen (Oxoid). Les cultures liquides ont été agitées à 175 tours par minutes. Le 5-fluorouracile a été utilisés aux concentrations suivantes : 20, 200, 500 et  $1000 \mu\text{g.ml}^{-1}$ .

Les souches d'*Escherichia coli* ont été cultivées à  $37^{\circ}\text{C}$  dans du bouillon Luria Bertani (LB) ou en dans du milieu minimum M9. Les antibiotiques et suppléments ont été ajoutés, lorsque nécessaire, aux concentrations suivantes : ampicilline  $100 \mu\text{g.ml}^{-1}$  (Sigma), isopropyl-beta-D-thiogalactopyranoside (IPTG) 1 mM (Sigma), uridine  $40 \mu\text{g.ml}^{-1}$  (Fluka), uracile  $40 \mu\text{g.ml}^{-1}$  (Fluka), tétracycline  $5 \mu\text{g/ml}^{-1}$  (Sigma), thiamine  $5 \mu\text{g/ml}^{-1}$  (Sigma), biotine  $1 \mu\text{g/ml}^{-1}$  (Sigma), 5-fluorouracile  $20 \mu\text{g/ml}^{-1}$  (Sigma). La souche *E. coli* HB101 a été utilisée comme bactérie hôte pour la construction des plasmides. La souche S06710 (F<sup>-</sup>, *upp*, *udp*, *thi*, *bio*, *argA*, *lysA*, *car-96::Tn10*) (Lundegaard et Jensen, 1999), utilisée comme souche parentale dans les expériences de complémentation. Elle est déficiente dans la première étape de la voie *de novo* de synthèse de l'UMP, et nécessite de l'uridine, mais pas de l'uracile, comme source de pyrimidine à cause des mutations *upp* et *udp*.

## b. Techniques moléculaires

Les gènes codant l'OPRTase de *H. pylori*, *hp1257* ; la XGPRTase, *hp0735* ; l'OPRTase d'*E. coli*, *pyrE* et l'UPRTase d'*E. coli*, *upp*, ont été clonés dans le vecteur d'expression pQE80-L (Invitrogen). Brièvement, les phases ouvertes de lecture ont été amplifiées par PCR en utilisant la Phusion High Fidelity DNA Polymerase (Finnzymes) sur l'ADN génomique de *H. pylori* 26695 et *E. coli* MG1655. Les fragments obtenus ont été digérés par BamHI (SphI pour *pyrE*) et HindIII (Fermentas), puis ligaturés dans le vecteur pQE80-L digéré par les mêmes enzymes. Le produit de ligature a été transformé par électroporation dans la souche HB101 d'*E. coli*. Les bactéries transformées ont été sélectionnées sur boîtes LB ampicilline. L'insertion correcte a été vérifiée par PCR et séquençage. Les souches et plasmides sont présentés dans le Tableau 8.

Tableau 8 : Liste des plasmides et souches bactériennes utilisées.

Nom	Origine	Plasmide	Informations ou génotypes pertinents	Référence
<b>Plasmides</b>				
pQE80L			Vecteur d'expression inductible à l'IPTG	Invitrogen
pHM0703	pQE80L		Expression inductible à l'IPTG de N-term 6His-Hp-OPRTase	ce travail
pHM0704	pQE80L		Expression inductible à l'IPTG de N-term 6His-Hp-XGPRTase	ce travail
pHM0705	pQE80L		Expression inductible à l'IPTG de N-term 6His-Ec-OPRTase	ce travail
pHM0707	pQE80L		Expression inductible à l'IPTG de N-term 6His-Ec-UPRTase	ce travail
<b>Souches bactériennes</b>				
<i>H. pylori</i>				
26695			sauvage	(Tomb <i>et coll.</i> , 1997)
FE029	26695		HP1175:: <i>aphA-3</i>	ce travail
<i>E. coli</i>				
HB101			sauvage	souche de laboratoire
SØ6710			<i>upp</i> <sup>-</sup>	(Lundegaard et Jensen, 1999)
FE017	HB101	pHM0703	Hp- <i>pyrE</i>	ce travail
FE020	SØ6710	pHM0703	<i>upp</i> <sup>-</sup> , Hp- <i>pyrE</i>	ce travail
FE021	SØ6710	pHM0704	<i>upp</i> <sup>-</sup> , Hp- <i>gpt</i>	ce travail
FE026	SØ6710	pQE80L	<i>upp</i> <sup>-</sup>	ce travail
FE027	SØ6710	pHM0705	<i>upp</i> <sup>-</sup> , Ec- <i>pyrE</i>	ce travail
GC002	SØ6710	pHM0707	Ec- <i>upp</i>	ce travail

Les cassettes de délétion pour les gènes *hp0005* et *hp1175* ont été construites comme présenté ci-après. Deux fragments d'environ 500 pb correspondant aux régions flanquantes du gène d'intérêt ont été amplifiés par PCR (la séquence des amorces est présentée dans le Tableau 9). La cassette de résistance à la kanamycine, *aphA-3* (Skouloubris *et coll.*, 1998), a été obtenue par digestion *Sma*I du plasmide pUC18K2 (Menard *et coll.*, 1993). Le fragment de délétion a été assemblé par deux réactions de PCR séquentielles. Le produit de PCR a été ensuite introduit dans la bactérie par transformation naturelle. Après transformation et sélection pour la résistance à la kanamycine, l'échange allélique correct a été vérifié par PCR à partir de colonies individuelles de *H. pylori*.

Tableau 9 : Liste des amorces utilisées. 1 f: forward ; r: reverse ; 2 site de restriction en minuscules

gene ciblé	Nom	Restriction	Séquence
<b>Hp-pyrE</b>	oEF047 (f)	BamHI	AGGTCgga <b>tc</b> ATGGATATTAAGGCATGTTATC
	oEF048 (r)	HindIII	TTGTTTTTaa <b>gctt</b> TTAGTTGCCACGACTTCCTGG
<b>Hp-gpt</b>	oEF051 (f)	BamHI	TAAGGgga <b>tc</b> ATGCATTATTCTTATGAAACCTTTT
	oEF052 (r)	HindIII	CTTTAACTaa <b>gctt</b> TAAATGGCTTTTTAAGTTTTTCAAATC
<b>Ec-pyrE</b>	oEF053 (f)	SphI	AGGAGCgca <b>tc</b> ATGAAACCATATCAGCGCC
	oEF054 (r)	HindIII	CGGCGAaa <b>gctt</b> TAAACGCCAAACTCTTCGCGAT
<b>Ec-upp</b>	oEF087 (f)	BamHI	GTATAATGGATCCATGAAGATCGTGGAAGTCAAACAC
	oEF088 (r)	HindIII	TATTTTAAAGCTTTTATTTTCGTACCAAAGATTTTGTC
<b>hp1175</b>	oEF089 (f)		CTAGAATATCGAAGGCTTTATAGCGTGTATC
	oEF090 (r)		GTTAGTCACCCGGGTACGTTAGATAACTCCTTACTATTTG
	oEF091 (f)		TACCTGGAGGGAATAATGGCTCTTTTAAAGGGCTTTG
	oEF093 (r)		GCAAGGTTACAAAGGGATTACCAGCGGTTTGC
	oEF094 (f)		CACGAATTATTATTCTTTTCTAGAC
	oEF095 (r)		AGTTGTGCCAAGCGAATTGAAC
	oEF096 (f)		GTGGTTTCATCATGGATTAATGAACGCTC
<b>hp0005</b>	oEF057 (r)		GTTAGTCACCCGGGTACGAAAGCCCTTAAAGTGTA
	oEF058 (f)		TACCTGGAGGGAATAATGATGCGCGTGTTAGAAACGATT
	oEF059 (r)		GATTAAAAGCTGTTGGGCGTCCTTTTTTGCC
	oEF060 (f)		TGCGGGTGGTTTTTAACTC
	oEF061 (r)		CAAATTATCGTCATCGCGCA
	oEF062 (f)		AGTTGTGCCAAGCGAATTGAAC

c. Séquences et recherche de motifs

La recherche de protéines PRTase putatives dans *H. pylori* a été effectuée à l'aide de la base de données InerPro (Mulder *et coll.*, 2007). Les séquences génomiques ont été extraites du navigateur de génomes GenoList (<http://genolist.pasteur.fr/>) : Colibri (*E. coli*) et PyloriGene (*H. pylori*).

#### d. Purification de Hp-OPRTase et mesure de l'activité enzymatique

L'OPRTase de *H. pylori* a été exprimée dans la souche HB101 d'*E. coli* à 37°C dans 750 ml de milieu LB contenant 100 µg.ml<sup>-1</sup> d'ampicilline. L'expression de la protéine a été induite par l'addition de 1 mM d'IPTG en phase exponentielle précoce (DO<sub>600</sub>=0,5) pendant 3 heures. Les protéines étiquetées 6-Histidine ont été purifiées par chromatographie par gravité sur une résine d'agarose Ni-NTA (Qiagen). Les échantillons protéiques obtenus ont été stockés à 4°C dans du tampon Tris-HCL 50 mM (pH 7,5) et se sont révélés stables pendant plusieurs semaines. Après purification les protéines ont été séparées sur gel SDS-PAGE 12% et visualisées par un marquage au bleu de Coomassie brillant. Pour les expériences de Western-blot, les protéines sur gel ont été transférées sur une membrane de nitrocellulose (Hybond PVDF Amersham) et traitées avec un anticorps mono-clonal anti-histag (Invitrogen).

Le mélange réactionnel pour la mesure des activités OPRTase et UPRTase est une solution aqueuse (25 µl) composée de 50 mM Tris-HCl (pH 7,5) ; 5 mM MgCl<sub>2</sub> ; 1 mM PRPP ; 12,5 à 500 µM orotate et/ou 100 à 5000 µM uracile. La réaction a été déclenchée par l'addition de l'enzyme Hp-OPRTase et incubée 5 minutes à 37°C. La réaction a été efficacement arrêtée en incubant la mélange 5 min à 95°C. L'échantillon a été récupéré par centrifugation et les composés (orotate, OMP, uracile) ont été séparés par chromatographie haute pression en phase liquide (HPLC system Beckman Gold) en utilisant une chromatographie en phase inverse (Pursuit® C18). L'élution isocratique a été réalisée à un débit de 1 ml.min<sup>-1</sup> en utilisant un tampon aqueux 10 mM KHPO<sub>4</sub>-H<sub>3</sub>PO<sub>4</sub> (pH 4) contenant 10 mM de KCl. L'analyse des données a été réalisée à l'aide du logiciel Laura Light 3 de LabLogic. Les paramètres cinétiques ont été déterminés par régressions non-linéaires à l'aide du logiciel GraphPad Prism version 4.00 pour Windows (GraphPad Software, San Diego California USA).

### C. Résultats

a. *Helicobacter pylori* est sensible à de fortes concentrations de 5-fluorouracile

*H. pylori* est décrit comme capable d'incorporer faiblement l'uracile exogène [ $0.52 \pm 0.07$  pmol.h<sup>-1</sup> pour 10<sup>6</sup> cellules] et posséder une activité UPRTase mesurable [ $0.01 \pm 0.005$  nmol.min<sup>-1</sup>.mg<sup>-1</sup> de protéine] (Menz *et coll.*, 1994). Ces expériences ont été réalisées avec la souche NCTC 11639 de *H. pylori* dont le génome n'est pas encore séquencé. On ne peut pas exclure la possibilité que cette souche contient un gène *upp* codant pour l'UPRTase. En effet, ce gène est présent chez d'autres  $\epsilon$ -protéobactéries comme certaines sous-espèces de *Campylobacter* ou encore *Wolinella succinogenes*. Afin d'éliminer cette hypothèse, j'ai décidé de confirmer par une approche alternative l'observation de Mendz et collaborateurs en utilisant la souche complètement séquencée 26695. Afin de tester la capacité à importer et métaboliser l'uracile, j'ai fait croître *H. pylori* 26695 en milieu riche additionné de 5-fluorouracile (5FU).

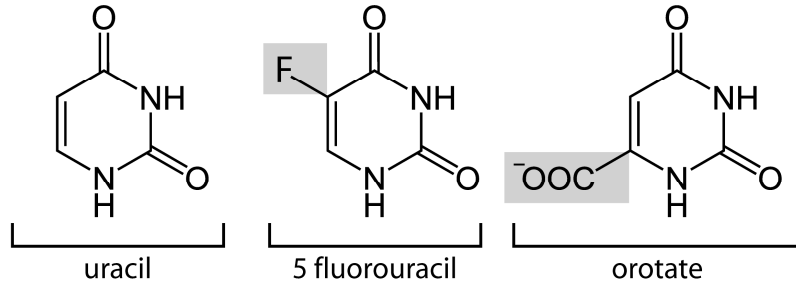


Figure 22 : Structure chimique de l'uracile, du 5-fluorouracile et de l'orotate.

Comme nous l'avons vu dans l'introduction, le 5FU est un analogue toxique de l'uracile qui possède un atome fluorine en position C-5 à la place de l'hydrogène Figure 22 : Structure chimique de l'uracile, du 5-fluorouracile et de l'orotate.. Après son entrée dans la cellule, le 5FU est converti trois métabolites actifs et toxiques : le fluoro-2'-désoxyuridine 5'-monophosphate (FdUMP) qui inhibe la réaction catalysée par la thymidylate synthase, le fluoro-2'-désoxyuridine 5'-triphosphate (FdUTP) et le 2'-uridine 5'-triphosphate (FUTP) qui seront mésincorporés dans l'ADN et l'ARN respectivement. Après l'entrée dans la cellule, la première étape est la conversion du 5FU en 5FUMP



par l'UPRTase ou par l'action successive de l'uridine phosphorylase (UDPase) et l'uridine kinase (UDKase). Il est attendu que *H. pylori* 26695, comme toutes les bactéries ne possédant pas *upp*, soit résistante au 5FU. Cette souche a été cultivée en milieu liquide (Brain Heart Infusion, 10% sérum de veau fœtal) en présence de 5FU, la croissance a été suivie par la mesure de l'absorbance à 600 nm.

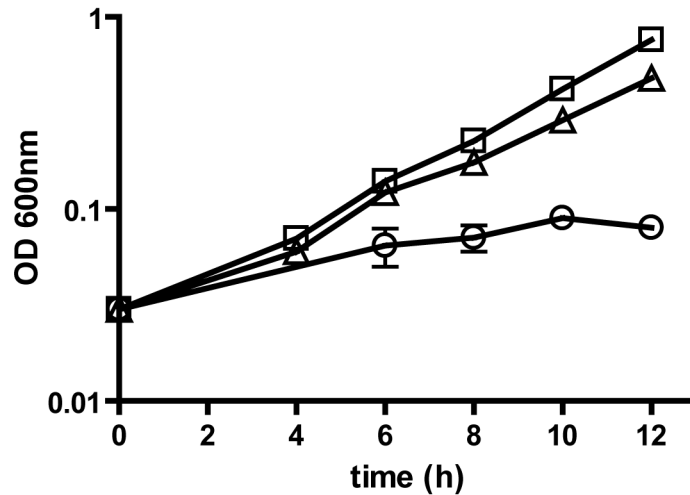


Figure 23 : La souche *H. pylori* 26695 a été cultivée en milieu BHI 10% SVF en absence (carrés) ou en présence de 5-fluorouracile à 500 (triangles) et 1000  $\mu\text{g}\cdot\text{ml}^{-1}$  (ronds).

J'ai pu ainsi observer que la croissance de *H. pylori* 26695 est inhibée par le 5FU. Cette inhibition est fonction de la concentration en 5FU, j'ai pu déterminer une Concentration Minimale d'Inhibition (CMI) de 1000  $\mu\text{g}\cdot\text{ml}^{-1}$  (Figure 23). Ce résultat indique clairement que *H. pylori* est capable de récupérer l'uracile exogène et de le convertir en une forme métaboliquement active, malgré l'absence de voie de récupération identifiée. Il est intéressant de noter que les concentrations nécessaires à l'inhibition de croissance sont élevées comparativement à celle inhibant la croissance d'*E. coli* (CMI 0,1  $\mu\text{g}\cdot\text{ml}^{-1}$ ). Cette dernière observation est en accord la faible activité préalablement mesurée, indiquant qu'au moins dans ces conditions, la récupération de l'uracile ne semble pas être une voie majeure pour produire de l'UMP.

b. Le mécanisme de transport de l'uracile doit être passif et non facilité

Mes résultats indiquent que l'uracile peut traverser les membranes de *H. pylori*. Chez *E. coli*, l'uracile pénètre dans la cellule par diffusion à travers la membrane cytoplasmique. Cette diffusion

est facilitée par l'action d'une protéine liée à la membrane : l'uracile perméase (UraA) codée par le gène *uraA* (Andersen *et coll.*, 1995). L'analyse du génome de *H. pylori* 26695 n'a pas révélée de gène codant une protéine homologue à UraA. Cependant, le gène *hp1175* code une protéine, appartenant à la famille des symporteurs Nucléobase/Cation, souvent annotée comme perméase à uracile/xanthine. Afin de tester si cette protéine HP1175 facilite l'entrée de l'uracile, j'ai construit une souche de *H. pylori* inactivée pour le gène *hp1175*.

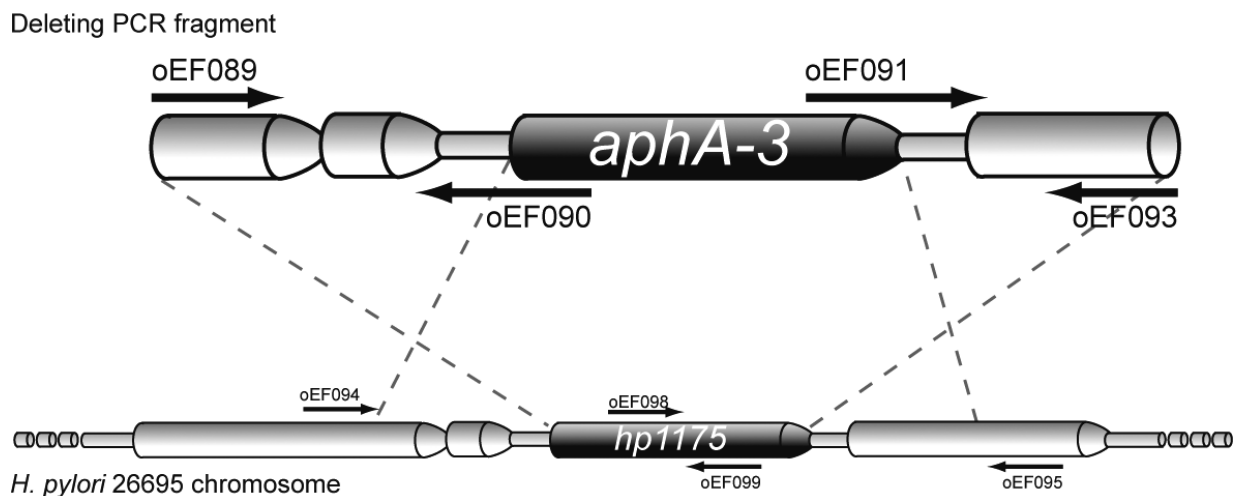


Figure 24 : Schéma de la délétion du gène *hp1175*.

La stratégie utilisée a été de construire, par PCR, un fragment d'ADN linéaire correspondant à une cassette non polaire de résistance à la kanamycine (*aphA-3*) flanquée par les régions amont et aval du gène *hp1175* (Figure 24). Après transformation naturelle, j'ai sélectionné les bactéries résistantes à la kanamycine et vérifié par PCR la correcte délétion. Si HP1175 est une perméase à uracile, la souche mutante (dénommée FE029) devrait être moins sensible au 5FU que la souche sauvage. Afin de tester cette hypothèse j'ai réalisé des cultures liquides de ces deux souches en présence de 5FU, mais je n'ai pas observé de différence significative dans la sensibilité au 5FU (donnée non montrée). Le fait que l'absence de *hp1175* n'affecte pas la sensibilité au 5FU suggère que HP1175 ne transporterait pas l'uracile mais un autre substrat, éventuellement la xanthine. En considérant ce résultat et l'absence d'autres transporteurs putatifs de nucléobases, je propose que l'uracile doive pénétrer la cellule par diffusion passive non-facilitée.

### c. Recherche bioinformatique de protéines contenant un motif « PRTase »

L'expérience de sensibilité au 5FU présentée ci-dessus indique que l'uracile est activement métabolisé après son entrée dans la cellule. Toutes les nucléobases, à l'exception de la thymine et de la cytosine, sont métabolisés par une phosphoribosyltransférase (PRTase) (Sinha et Smith, 2001). Les PRTases catalysent le transfert du ribose 5-phosphate provenant du PRPP vers l'atome N1 de différents substrats. L'analyse du génome a permis d'identifier trois membres de la famille PRTase chez *H. pylori* : deux sont impliquées dans le métabolisme des purines (adénine et xanthine-guanine PRTases) et une dans la voie *de novo* de synthèse des pyrimidines (orotate PRTase ou OPRTase). De façon intéressante j'ai pu noter que le substrat de l'OPRTase, l'orotate, ne diffère de l'uracile que par un groupement carboxyle (Figure 22). Ainsi l'OPRTase de *H. pylori* (Hp-OPRTase) codée par le gène *hp1257*, représentait le meilleur candidat pour expliquer l'activité UPRTase chez *H. pylori*.

### d. Hp-OPRTase compense l'absence d'UPRTase chez *E. coli*

Afin de tester si Hp-OPRTase est capable d'agir comme une UPRTase *in vivo*, j'ai réalisé un test génétique de complémentation hétérologue en utilisant une souche d'*E. coli* auxotrophe pour l'uridine : SØ6710 (F-, *upp*, *udp*, *thi*, *bio*, *argA*, *lysA*, *car-96::Tn10*) (Lundegaard et Jensen, 1999). Dans cette souche, la voie *de novo* de synthèse de l'UMP est bloquée à sa première étape et la récupération de l'uracile est impossible à cause des mutations dans les gènes codant l'UPRTase (*upp*) et l'UDPase (*udp*). En milieu minimum, cette souche ne peut être maintenue qu'en présence d'uridine. Le gène *hp1257* codant Hp-OPRTase a été amplifié par PCR puis cloné dans le vecteur pQE80-L d'expression inductible à l'IPTG (Qiagen). Le plasmide obtenu a été transformé dans la souche SØ6710, et la souche résultante a été nommée FE020. J'ai fait croître cette souche sur milieu minimum solide en utilisant l'uracile comme seule source de pyrimidines, en absence et présence d'inducteur.

Tableau 10 : Phénotypes de croissance des différentes souches d'*E. coli* utilisées en réponse à 20 µg.ml<sup>-1</sup> de 5FU.

Souche	Génotype pertinent	M9 uri	M9 ura	M9 ura +IPTG	LB 5FU	LB 5FU +IPTG
FE026	<i>upp<sup>-</sup></i>	+++	-	-	+++	+++
GC002	<i>upp<sup>-</sup> Ec-upp</i>	n.d.	n.d.	n.d.	++	-
FE020	<i>upp<sup>-</sup>, Hp-pyrE</i>	+++	-	++	+++	+/-
FE021	<i>upp<sup>-</sup>, Hp-gpt</i>	+++	-	-	n.d.	n.d.
FE027	<i>upp<sup>-</sup>, Ec-pyrE</i>	n.d.	n.d.	n.d.	+++	+/-

L'expression de Hp-OPRTase a permis la croissance de S06710 sur milieu minimum seulement en présence d'IPTG et d'uracile (Tableau 10), ce qui indique que Hp-OPRTase peut utiliser l'uracile comme substrat pour produire de l'UMP *in vivo*. Afin de tester si cette activité est spécifique à Hp-OPRTase et n'est pas commune aux autres PRTases, j'ai appliqué la même procédure pour exprimer la xanthine-guanine phosphoribosyltransférase (Hp-XGPTase) dans S06710. Cette enzyme a une spécificité de substrat pour les 6-oxopurines comme la guanine et la xanthine (Duckworth *et coll.*, 2006). Contrairement à FE020, la souche S06710 exprimant Hp-XGPRTase (FE021) n'est pas capable de croître en présence d'uracile et d'IPTG (Tableau 10). L'ensemble de ses résultats indiquent que Hp-OPRTase est capable de convertir l'uracile en une forme métaboliquement active *in vivo* chez *E. coli*.

e. L'expression de Hp-OPRTase induit une sensibilité au 5FU chez *E. coli*

Convaincu que Hp-OPRTase peut convertir l'uracile en UMP *in vivo*, j'ai recherché si cette activité était suffisante pour expliquer la sensibilité de *H. pylori* à de fortes doses de 5FU. Afin de tester cette idée, j'ai utilisé le même test de complémentation génétique pour estimer la sensibilité au 5FU dans différents contextes génétiques. En présence d'uridine, la croissance de la souche parentale S06710 ne devrait pas être inhibée par le 5FU, seule l'expression d'une enzyme possédant une activité UPRTase devrait conférer une sensibilité à la drogue. Afin de valider le test, nous avons construit une souche dérivée de S06710 exprimant l'UPRTase d'*E. coli* (Ec-UPRTase) qui sera nommée

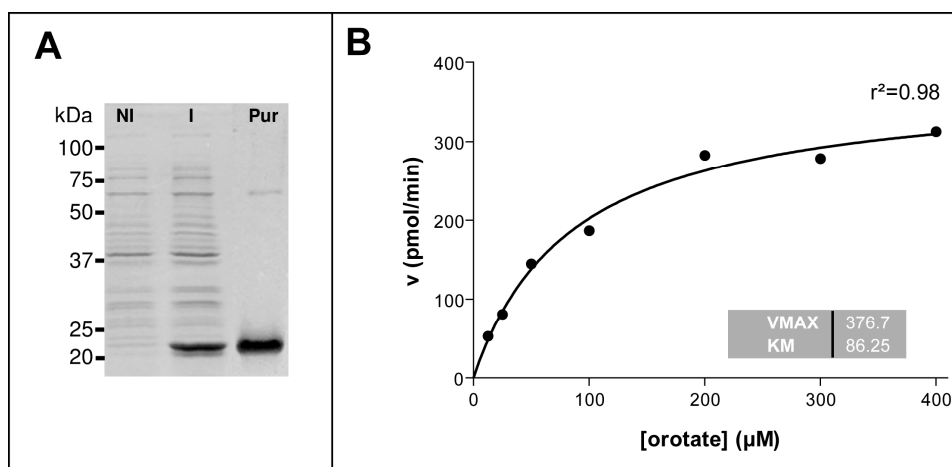
GC002. Comme prédit, la souche parentale montre un phénotype de résistance au 5FU, alors que la croissance de la souche GC002 est totalement inhibée en présence d'inducteur (Tableau 10). La faible inhibition de croissance de GC002 en absence d'IPTG résulte probablement de la fuite du promoteur du système d'expression. En présence d'IPTG, la souche exprimant Hp-OPRTase (FE020) montre une sensibilité à une faible concentration en 5FU ( $20 \mu\text{g}\cdot\text{ml}^{-1}$ ) (Tableau 10) en accord avec le résultat du test de complémentation. Considérant ce résultat, je propose que la sensibilité au 5FU observé chez *H. pylori* résulte d'une activité d'une activité secondaire de Hp-OPRTase.

- f. L'activité secondaire UPRTase de HP-OPRTase n'est pas suffisante pour permettre la croissance de *H. pylori* en absence de voie de novo

Une question se posait alors : la faible activité UPRTase de Hp-OPRTase identifiée *in vivo* est-elle suffisante pour permettre la croissance de *H. pylori* en utilisant que l'uracile comme seule source de pyrimidines, indépendamment de la voie *de novo* ? Pour répondre à cette question, j'ai utilisé une approche génétique pour bloquer la voie *de novo* en présence d'uracile en inactivant l'étape suivant celle catalysée par Hp-OPRTase : orotidine 5'-phosphate décarboxylase (ODCase) codée par *hp0005*. Ainsi, j'ai entrepris la construction d'un fragment d'ADN linéaire correspondant à la cassette kanamycine flanquée par les régions amont et aval de *hp0005*. Afin de favoriser l'inactivation, il était important d'être dans des conditions où j'étais sûr que l'uracile pouvait être métabolisé. En me basant sur les résultats d'inhibition de croissance par le 5FU, j'ai établi qu'une concentration de  $1 \text{ mg}\cdot\text{ml}^{-1}$  dans le milieu devrait être acceptable. Pour favoriser la présence d'uracile intracellulaire au moment de l'inactivation de *hp0005*, j'ai ajouté ce composé dans le milieu de culture dans toutes les étapes du processus de transformation. J'ai réalisé trois tentatives de transformation de la souche 26695, mais je n'ai pas pu obtenir de bactérie résistante à la kanamycine après 10 jours. Bien que l'absence de preuve ne soit pas preuve de l'absence, il est raisonnable de penser que *H. pylori* n'est pas capable de croître en utilisant l'uracile comme seule source de pyrimidines dans les conditions testées.

g. Test de la spécificité de substrat de Hp-OPRTase *in vitro*

Afin de mieux comprendre les résultats obtenus lors des études *in vivo*, j'ai tenté de mesurer l'activité OPRTase de Hp-OPRTase *in vitro*. La protéine Hp-OPRTase fusionnée à une étiquette oligohistidine en position N-terminal a été exprimée dans *E. coli*. La majorité de la protéine recombinante était présente dans les fractions solubles des lysats bactériens. Nous avons obtenu approximativement 25 mg de protéine purifiée par litre de culture. Hp-OPRTase purifiée a une masse moléculaire apparente de 23 kDa sur gel SDS-PAGE en accord avec la masse moléculaire prédite de 21,7 kDa (Figure 25A).



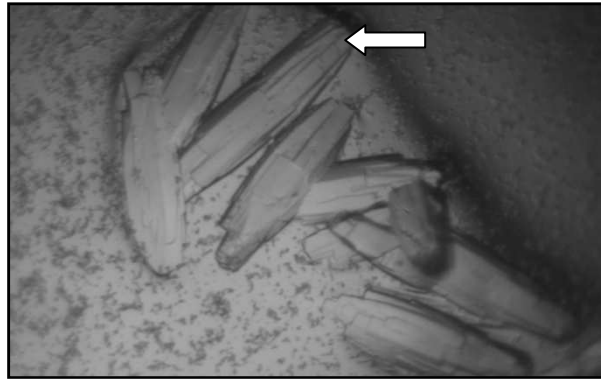
**Figure 25 :** A- Analyse SDS-PAGE de la protéine recombinante Hp-OPRTase. Les deux premières colonnes correspondent successivement à l'extrait brut avant induction (NI) et après induction (I). La fraction de protéine pure est visible dans la 3<sup>ème</sup> colonne. B- Cinétique de la catalyse de Hp-OPRTase pour son substrat naturel : l'orotate. L'activité a été mesurée par séparation de l'orotate et de l'OMP par HPLC (voir Matériels et Méthodes).

La stabilité de la protéine étiquetée a été vérifiée par Western-blot en utilisant des anticorps anti-histidine (donnée non montrée). Nous en premier lieu mesuré l'activité classique OPRTase de l'enzyme comme décrit dans la section matériels et méthodes. Le substrat orotate et le produit OMP ont été séparé par élution isocratique [10 mM  $\text{KHPO}_4\text{-H}_3\text{PO}_4$  buffer (pH 4), 10 mM KCl at 1 ml/min] en utilisant une chromatographie de phase inverse par chromatographie à haute-pression en phase liquide (HPLC). J'ai pu déterminer les constantes cinétiques sous des conditions où la concentration en PRPP est constante, et celle de l'orotate varie (Figure 25B). La valeur de constante d'affinité pour l'orotate ( $K_m^{\text{orotate}}$ ) de 86  $\mu\text{M}$  mesurée est similaire à celle mesurée pour l'OPRTase de *Thermus*

*thermophilus* [75  $\mu\text{M}$ ] (Bunnak *et coll.*, 1995), mais plus élevée que celle d'Ec-OPRTase et *Corynebacterium* OPRTase [30 et 33  $\mu\text{M}$  respectivement] (Shimosaka *et coll.*, 1985; Wang *et coll.*, 2007). J'ai pu déterminer également que l'efficacité catalytique correspondante  $K_{\text{cat}}/K_{\text{m}}$  était de 0,21  $\text{min}^{-1}\cdot\mu\text{M}^{-1}$ , indiquant que l'orotate est un substrat naturel de l'enzyme. Afin de tester si l'uracile pouvait être un substrat de l'enzyme, j'ai remplacé le substrat original par l'uracile dans la réaction. Dans les conditions testées, je n'ai pas détecté l'apparition d'un pic pouvant correspondre à l'UMP même en présence de 1 mM d'uracile (valeur correspondant à 10 fois le  $K_{\text{m}}^{\text{orotate}}$ ). De plus, je n'ai pas observé d'inhibition par l'uracile [jusqu'à 5 mM] de la réaction de formation d'OMP catalysée par Hp-OPRTase.

#### h. Cristallogénèse de Hp-OPRTase

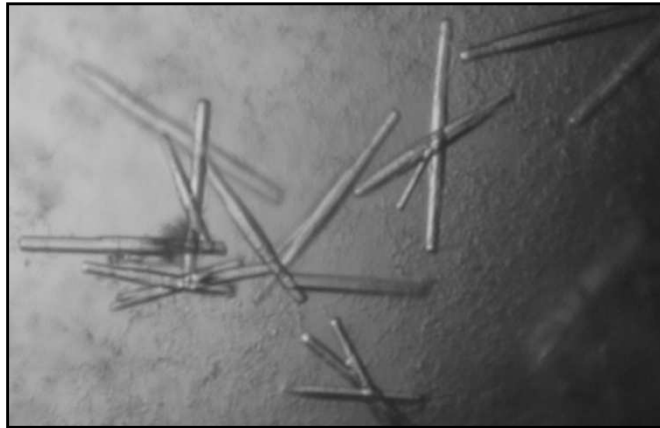
Il a été décrit deux familles d'OPRTases bactériennes : type « *E. coli* » et type « *Thermus* » ; les deux familles ne présentant que peu d'homologie à l'exception des domaines catalytiques (Bunnak *et coll.*, 1995). Il est intéressant de noter que plusieurs des bactéries qui possèdent une enzyme de la famille « *Thermus* », comme *H. pylori*, ne possèdent pas de gène *upp* codant l'UPRTase. Il est tentant d'imaginer que la capacité à métaboliser l'uracile soit commune aux membres de cette sous-famille. Il serait alors intéressant de visualiser si l'uracile peut se fixer dans la poche catalytique de l'enzyme. Bien que des structures tridimensionnelles de protéines de la famille « *E. coli* » ont déjà été résolues (Henriksen *et coll.*, 1996; Scapin *et coll.*, 1994), aucune structure de la deuxième famille n'est disponible. Des travaux préliminaires ont suggéré que Hp-OPRTase pouvait cristalliser dans des conditions définies (19% Polyéthylène glycol MME 5000, 0.1M sulfate d'ammonium) (Kim *et coll.*, 2003), mais cinq ans plus tard aucune structure résolue n'a été publiée. Nous avons établi une collaboration avec le cristallographe Laurent Terradot et son équipe (ESRF Grenoble) afin d'obtenir une structure tridimensionnelle de la protéine. Nous leur avons fourni la protéine Hp-OPRTase purifiée à homogénéité à l'aide d'une étiquette 6-Histidine en position N-terminale. Ils ont essayé de reproduire les cristaux décrits dans cette étude et y sont parvenus à partir de protéine concentrée à environ 21 mg/ml (Figure 26).



**Figure 26 : Cristaux de Hp-OPRTase obtenus dans du PEG 5kMME, 17% Ammonium Sulfate 0.1M (20 degrés/ 3 jours). La flèche indique l'endroit sur le cristal où on s'aperçoit de la multiplicité cristalline. Multiplicité confirmée lors de la diffraction**

Lorsque ces cristaux ont été testés sur la ligne de lumière du synchrotron (ESRF), leur diffraction s'est avérée insuffisante. En effet les cristaux obtenus sont en fait un empilement de cristaux (flèche Figure 26). Malgré plusieurs essais d'amélioration par des techniques de « seeding », ils ne sont pas parvenus à obtenir de meilleurs cristaux. Afin de rechercher de manière exhaustive de nouvelles conditions de cristallisation, ils ont procédé à un criblage à haut débit en utilisant la plateforme de l'EMBL de Grenoble testant trois concentrations différentes de protéines (méthode de diffusion de vapeur par gouttes assises). Ainsi ils ont réalisé 576 expériences de cristallisation indépendantes pour chaque concentration en utilisant la méthode de la diffusion de vapeur. Ils ont pu obtenir des conditions nouvelles de cristallisations, toujours en présence de polyéthylène glycol (de masse moléculaire différentes), mais à plusieurs pH différents. A partir de ces conditions obtenues au robot, ils ont essayé d'optimiser les plus prometteuses en variant quelques paramètres (pH, sels, concentration en PEG) et en faisant des gouttes plus grandes (2 à 3 $\mu$ l). Malheureusement aucun des cristaux obtenus n'a pu conduire à une bonne qualité de diffraction (plus de 30 cristaux testés sur la ligne de lumière) (Figure 27).





**Figure 27 : Cristaux de Hp-OPRTase obtenus dans du PEG 3350, 20%, 0.1M Hepes pH 7.5, 0.2M MgCl<sub>2</sub> reproduit à la main pour optimisation.**

Il se peut que l'étiquette 6-Histidine ayant servie à la purification de la protéine soit responsable de la mauvaise diffraction des cristaux. Afin d'éliminer ce problème, le gène *hp1257* codant Hp-OPRTase a été cloné dans un vecteur (pET-TOPO151D) exprimant la protéine en fusion avec un étiquette histidine clivable a la protéase TEV (Tobacco Etch Virus). Les premiers résultats utilisant la protéine dépourvue de l'étiquette sont encourageants et seront approfondis dans les mois qui viennent.

#### ***D. Discussion***

Je trouve toujours intéressant de découvrir l'existence de données contradictoires dans la littérature. Dans le cas présent, nous avons d'un côté des données expérimentales apparemment rigoureuses et fiables qui indiquaient la présence d'une voie de récupération de l'uracile chez *H. pylori* (Mendz *et coll.*, 1994), et de l'autre un génome séquencé et annoté qui prédisait son absence (Tomb *et coll.*, 1997). Les travaux présentés dans cette étude ont permis de confirmer la première observation, c'est-à-dire que *H. pylori* était capable de métaboliser l'uracile exogène. J'ai pu ainsi montrer que *H. pylori* est sensible à de fortes doses d'un analogue de l'uracile, le 5-fluorouracile (5FU). Ce résultat démontre que le 5FU est capable d'entrer dans la cellule et qu'il est ensuite métabolisé en 5FUMP.

Chez *E. coli*, l'entrée de l'uracile dans la cellule se fait de manière passive, mais se trouve facilitée par l'uracile perméase codée par le gène *uraA* (Andersen *et coll.*, 1995). L'analyse du génome d'*H. pylori* n'a pas permis de mettre en évidence de gène homologue à *uraA*. J'ai cependant testé l'hypothèse selon laquelle que la protéine codée par le gène *hp1175*, souvent annotée comme putative perméase à nucléobase, pouvait avoir une activité uracile perméase. L'inactivation du gène *hp1175*, n'a cependant pas diminué la sensibilité de *H. pylori* pour le 5FU, indiquant que cette protéine ne transporterait vraisemblablement pas l'uracile. Bien qu'il soit possible d'imaginer que la présence d'un autre transporteur de l'uracile masque le rôle de HP1175, il semble plus probable que ce composé pénètre la cellule par voie passive non facilitée.

L'uracile, une fois pénétré la cellule, peut être converti en UMP soit via un intermédiaire uridine par l'action concertée de l'uridine phosphorylase et de l'uridine kinase, soit directement par l'uracile phosphoribosyltransférase (UPRTase). Les travaux antérieurs nous donnaient deux informations : l'incorporation de l'uridine est quasi-inexistante, et l'activité enzymatique UPRTase est mesurable (Mendz *et coll.*, 1994). Ceci m'a laissé penser qu'il ne fallait pas chercher une nouvelle famille d'UPRTase, mais plutôt une enzyme catalysant le transfert d'un résidu ribosyl-phosphate d'une molécule de PRPP sur un substrat proche de l'uracile. La seule protéine répondant à ces critères fut l'orotate phosphoribosyltransférase (OPRTase) qui utilise l'orotate, ou 6-carboxyuracile, comme substrat. J'ai utilisé une souche d'*E. coli* ne possédant pas de voie de récupération de l'uracile fonctionnelle et inactivée dans la voie *de novo* de synthèse de l'UMP. J'ai pu montrer que l'expression de Hp-OPRTase restaure l'auxotrophie pour l'uracile de cette souche et lui confère une sensibilité au 5FU. Ces résultats indiquent que Hp-OPRTase a une activité UPRTase *in vivo*, et qu'elle est très probablement responsable de la sensibilité de *H. pylori* à des fortes concentrations d'uracile. Les études *in vitro* ont montré que Hp-OPRTase est capable de convertir l'orotate en OMP. Malheureusement, aucune activité UPRTase n'a pu être mesurée dans les conditions testées. Il est à noter que des travaux réalisés *in vitro* ont pu montrer que l'OPRTase de *Salmonella typhimurium* était capable d'utiliser l'uracile comme substrat avec une très faible efficacité ( $k_{cat}/K_m = 0,6 \text{ M}^{-1} \cdot \text{s}^{-1}$

pour l'uracile contre  $9,8 \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$  pour l'orotate) (Bhatia et Grubmeyer, 1993). Je pense alors qu'il serait intéressant de poursuivre la recherche des conditions nécessaires à la mesure d'une UPRTase de Hp-OPRTase *in vitro*, en faisant varier notamment le pH, la nature et la concentration en cation divalent. J'envisage également de tester si cette activité ne pourrait pas être activée allostériquement par le GTP, ou un autre nucléotide triphosphate, comme c'est le cas pour l'UPRTase d'*E. coli* (Jensen et Mygind, 1996) et de *Sulfolobus solfataricus* (Jensen *et coll.*, 2005).

Chalker et collaborateurs, en utilisant la technique de mutagénèse par remplacement allélique, avaient conclu que Hp-OPRTase est essentielle car aucun mutant de délétion de *hp1257* n'avait pu être obtenu, même en présence d'uracile (Chalker *et coll.*, 2001). Considérant les résultats présentés ci-dessus, cette observation peut être comprise si Hp-OPRTase est l'enzyme qui catalyse la conversion de l'uracile en UMP. Burns et collaborateurs, en utilisant une approche similaire, n'ont pas pu obtenir de mutant de *pyrB* qui code l'aspartate carbamoyltransférase (ACTase) qui catalyse la deuxième étape de la synthèse *de novo* même en présence d'uracile ou d'orotate (Burns *et coll.*, 2000). Finalement, il a été montré que des composés dérivés du pyrazole ciblant spécifiquement la dihydroorotate déshydrogénase (codée par *pyrD*) inhibent sélectivement la croissance de *H. pylori* (Copeland *et coll.*, 2000). Dans ce dernier cas, l'effet de l'uracile n'a pas été testé. En prenant en compte mes résultats et ceux présentés dans les travaux antérieurs, je propose que la voie de récupération de l'uracile soit présente chez *H. pylori* mais faiblement active, au moins en condition de laboratoire. Il n'est pas impossible que ces voies puissent être régulées dans des conditions physiologiques particulières. On ne peut pas exclure que la surexpression du gène *hp1257* codant Hp-OPRTase puisse permettre l'utilisation de l'uracile dans la croissance de *H. pylori*. Afin de tester cette hypothèse, le gène *hp1257* pourrait être cloné dans un vecteur d'expression dans le but de surexprimer Hp-OPRTase dans *H. pylori*. Il serait alors théoriquement possible d'inactiver la voie *de novo* de formation de l'UMP en présence d'uracile exogène.

La capacité à métaboliser l'uracile *in vivo* n'est peut-être pas spécifique de l'OPRTase de *H. pylori*. Il est intéressant de noter qu'une souche dérivée de SØ6710 surexprimant l'OPRTase d'*E. coli* (FE027) est également sensible au 5FU en présence d'IPTG (Tableau 10 : Phénotypes de croissance des différentes souches d'*E. coli* utilisées en réponse à 20  $\mu\text{g}\cdot\text{ml}^{-1}$  de 5FU.). Chez les mammifères, qui ne possèdent pas de gène codant l'UPRTase, la sensibilité au 5FU est décrite comme étant sous la dépendance de l'activité de l'OPRTase (Peters *et coll.*, 1986; Reyes, 1969). A ma connaissance, ce travail indique pour la première fois que la faible activité UPRTase des enzymes OPRTases semble être une caractéristique commune aux protéines eucaryotes et procaryotes. De plus des travaux préliminaires menés par Damien Leduc ont pu montrer que chez l'archée *Pyrococcus horikoshii*, en absence de gène homologue à *upp* codant l'UPRTase, la croissance est inhibée par 150  $\mu\text{g}\cdot\text{ml}^{-1}$  de 5FU (données non publiées). L'ensemble de ses résultats implique que toute cellule bactérienne, archée ou eucaryote ayant une OPRTase active serait virtuellement sensible à de hautes doses de 5FU.



## **DISCUSSION GENERALE ET PERSPECTIVES**

Régulièrement en Science, les projets initialement planifiés aboutissent à des résultats difficilement interprétables, et ce sont les projets secondaires qui ouvrent de nouvelles perspectives. C'est un peu ce qu'il s'est passé durant cette thèse. Le projet concernant l'étude de ThyX *in vivo* chez *H. pylori* était très excitant car il promettait l'identification de nouvelles protéines impliquées directement ou indirectement dans la synthèse de l'ADN. Malheureusement, l'obtention de résultats peu concluants a freiné notre ardeur. A l'inverse les autres projets, originellement secondaires, se sont révélés être plus féconds. Le travail que j'ai trouvé le plus passionnant fut certainement celui présenté dans l'article II.

## Thymidylate synthases

### I. Distribution des deux thymidylate synthases dans le monde procaryote

Que savions-nous au début de ma thèse de la distribution phylogénétique des deux thymidylate synthases au sein du vivant ? La première observation, qui fut d'ailleurs à l'origine de la découverte de ThyX, était que la distribution des gènes *thyA* et *thyX* est exclusive (à de rares exceptions près) (Myllykallio *et coll.*, 2002). Une autre observation était que cette distribution ne dépendait pas des relations phylogéniques indiquant des transferts horizontaux de gènes fréquents, et des déplacements non-orthologues (Murzin, 2002). Afin d'identifier quelles pouvaient être les forces de sélection sous-jacentes, Damien Leduc avait tenté de mettre en parallèle la distribution des deux thymidylate synthases avec les conditions de vie des organismes (Leduc, 2004; Leduc *et coll.*, 2004b). Cette étude portant sur une centaine d'organismes dont le génome était entièrement séquencé, a mis en évidence quelques niches écologiques préférentiellement occupées par des organismes utilisant ThyX. Il notera par exemple que les trois quart des organismes hyperthermophiles (croissance à des températures supérieures à 80°C) possédaient le seul gène *thyX*. Une corrélation plus intéressante montrait que la quasi-totalité des organismes croissant à des faibles concentrations en oxygène (microaérophiles, par exemple *H. pylori*) possédaient également le gène *thyX*. De plus, il venait d'être montré que la complémentation de l'absence de *thyA* par *thyX* de

*Campyobacter jejuni* n'était possible qu'en conditions microaérophiles (Giladi *et coll.*, 2002). Considérant l'ensemble de ces observations, il était alors tentant de penser que les conditions environnementales pouvaient, au moins partiellement, expliquer la distribution du gène *thyX*. Cependant, comme le faisait remarquer l'auteur dans sa thèse, la limite de cette étude réside dans le nombre restreint d'échantillons utilisés pour chaque mode de vie (Leduc, 2004). On ne peut pas exclure que l'environnement joue un rôle, cependant cela ne constitue pas une explication globale de la distribution de ces deux familles d'enzymes. La question qui persistait à l'issue de cette analyse était : quelles sont les caractéristiques communes aux organismes utilisant l'une ou l'autre des thymidylate synthases ?

Le travail présenté dans l'article II a permis de montrer que les organismes utilisant ThyA et ceux utilisant ThyX diffèrent en trois points. En premier lieu, j'ai pu montrer que la vitesse de réplication de l'ADN était comparativement plus basse chez les organismes utilisant ThyX que chez ceux utilisant le couple ThyA/FolA. J'ai mis en évidence que la vitesse de réplication était fortement réduite si on simulait le remplacement non-orthologue de *thyA* par *thyX* dans *E. coli*. En mettant en parallèle ces résultats et les différences d'efficacité catalytiques des deux enzymes démontrées *in vitro*, nous avons proposé que la faible efficacité de ThyX limite la vitesse moyenne de réplication de l'ADN en ne permettant pas un approvisionnement assez rapide en dTTP. La deuxième observation que j'ai pu faire est que les organismes utilisant ThyX présentent généralement un taux de croissance relativement lent comparativement aux organismes ThyA. Enfin, j'ai montré que les organismes utilisant ThyX ont globalement un génome de taille réduite (médiane = 2,3 Mpb). Il est intéressant de noter que les organismes hyperthermophiles et microaérophiles ont pour caractéristique commune la petitesse de leur génome avec une taille médiane de 1,9 et 1,8 Mpb respectivement. Ainsi, l'observation faite plus tôt, concernant la surreprésentation de *thyX* dans ces deux groupes, apparaît prévisible.



A la lumière des données apportées par les travaux présentés dans l'article II, nous pouvons décrire les organismes utilisant ThyX comme des organismes à réplication et croissance lente possédant un petit génome. Mes travaux ont ainsi permis de montrer que la distribution du gène *thyX* codant la thymidylate synthase X n'est pas sporadique, mais peut être définie selon des critères physiologiques (vitesse de réplication, temps de croissance) et structuraux (taille du génome).

## II. Simplement des corrélations ?

Il est tentant d'imaginer un lien de cause à effet entre la vitesse de synthèse du dTMP, la vitesse de réplication, la vitesse de croissance et la quantité d'ADN présente dans la cellule. Un critique qui a évalué l'article II a fait remarquer que « corrélation n'est pas causalité ». Je suis parfaitement d'accord avec cette remarque, et je pense qu'il faut être prudent. Quels sont alors les éléments qui me laissent penser qu'il y a relation de causalité ? En premier lieu, j'ai établi expérimentalement que l'étape de synthèse du dTMP est un facteur limitant de la vitesse de réplication et du taux de croissance. Ce fait est d'ailleurs bien connu chez *E. coli* puisqu'il participe au phénomène de « thymineless death » (voir introduction p 29). La relation entre le taux de croissance et la vitesse de réplication est décrite depuis longtemps chez *E. coli* (Helmstetter, 1996) et se vérifie également chez *Bacillus subtilis* (communication personnelle B. Dalmais, INRA Jouy-en-Josas). Comme nous l'avons vu, le temps de réplication est fonction de la vitesse de réplication et de la taille du génome (voir discussion article II p 91). Cependant, il est généralement admis que le temps de réplication est peu sensible à l'augmentation d'ADN à répliquer, est-ce vrai ?

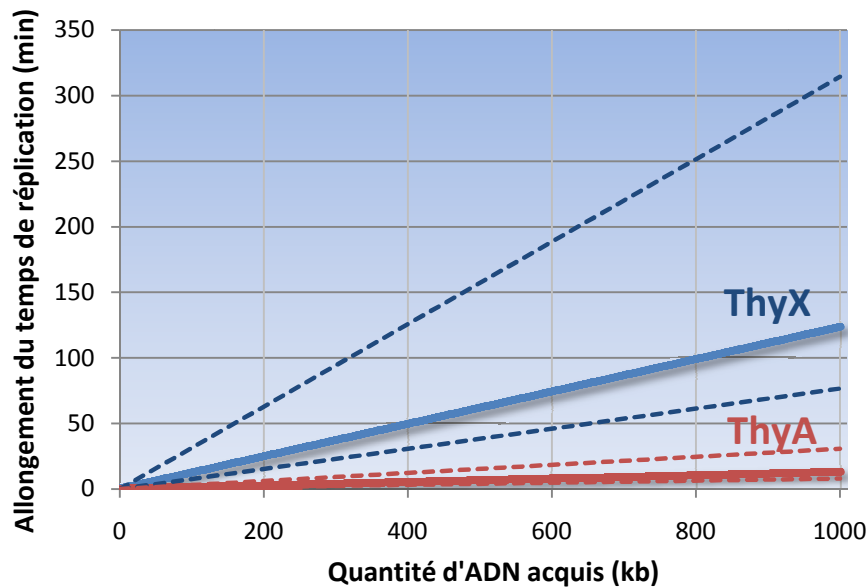


Figure 28 : Conséquence de l'expansion du génome sur le temps de réplication.

Ce graphique représente l'allongement théorique du temps de réplication en fonction de la quantité d'ADN nouvellement acquis. L'allongement du temps de réplication est calculé comme le quotient de la quantité d'ADN acquis par la vitesse de réplication ( $V$ ). Les valeurs de  $V$  utilisées pour les organismes ThyX (droite bleue) et ThyA (droite rouge) sont  $135 \pm 82$  et  $1301 \pm 762$  pb/sec respectivement [écarts-types en pointillés] (article II).

Si on considère les bactéries ayant une réplication rapide (i.e. *E. coli*), nous voyons qu'une augmentation de la taille du génome de 500 kb allongera le temps de réplication d'environ 6 minutes (Figure 28). Bien que faible, il est possible que cette variation ait tout de même un effet sur la physiologie générale de la bactérie. Pour preuve, il a été montré expérimentalement chez *E. coli* que l'augmentation de la quantité d'ADN de quelques kilobases porté par un plasmide conduisait à une diminution de la viabilité de la bactérie (Vellai *et coll.*, 1998). Il est cependant probable que la perte de fitness puisse être compensé par la cellule, en particulier si l'ADN nouvellement acquis apporte un avantage sélectif (résistance aux antibiotiques, nouvelles voies métaboliques, etc). En revanche, dans le cas des organismes qui utilisent ThyX, qui eux ont une réplication dix fois moins rapide (article II), la même augmentation va allonger le temps de réplication d'une heure (Figure 28). Il est raisonnable de penser qu'une augmentation aussi conséquente du temps de réplication aura des répercussions beaucoup plus drastiques sur le taux de croissance et la viabilité de la cellule.

Je pense que nous avons assez d'informations pour raisonnablement proposer que la faible activité de la thymidylate synthase X, en limitant la vitesse de réplication, freine l'expansion des

génomés procaryotes. Ainsi le couple ThyA/FolA, en offrant la possibilité d'augmenter la vitesse de réplication, permettrait d'augmenter le potentiel évolutif par l'acquisition de nouveaux gènes.

### III. Thymidylate synthase et eucaryotes ?

Nous avons vu dans l'introduction que les organismes eucaryotes possèdent généralement le couple ThyA/FolA pour produire le dTMP. Les deux seules exceptions connues à cette règle est la présence d'un gène homologue à *thyX* chez les amibes *Dictyostelium discoideum* (Dynes et Firtel, 1989) et *Physarum polycephalum*. De façon surprenante, la croissance *D. discoideum* est inhibée par l'antifolate méthotrexate (Deering et Michrina, 1982). Le fait que cette inhibition soit levée en présence de l'ajout de thymidine (conjointement à de l'adénine, guanine et sérine) semble indiquer que c'est l'action de la DHFR lors de la synthèse de dTMP qui est ciblée. Cet organisme possède une thymidine kinase et croit lentement (temps de doublement ~4h) en milieu riche ou sur un tapis bactérien (Fey *et coll.*, 2007). De plus, il a été montré chez *Physarum polycephalum* que la synthèse de dTMP durant les phases actives de croissance est majoritairement sous la dépendance de l'activité thymidine kinase. Il est alors envisageable que dans ces organismes la synthèse du dTMP ne soit que très partiellement dépendante de la thymidylate synthase ThyX.

Les résultats présentés dans l'article II indiquent que chez les organismes procaryotes la vitesse de synthèse du dTMP est un facteur limitant de la vitesse de réplication de l'ADN ce qui limite l'expansion des génomes. Si on applique aux eucaryotes le modèle proposé pour les procaryotes, on comprend aisément que la présence de *thyX* est contre-sélectionnée. En revanche, notre modèle indique également que l'efficacité du couple ThyA/FolA semble être un facteur limitant pour l'expansion des génomes procaryotes au-delà de 10 Mpb. Or les eucaryotes ont un génome beaucoup plus grand que celui des procaryotes (voir introduction p 43), comment expliquer alors que la réplication d'une si grande quantité d'ADN n'est pas limitée par la synthèse du dTMP ?

Une des différences communément admise entre les organismes de ces deux domaines est que la survie des eucaryotes n'est pas conditionnée à une croissance rapide, surtout chez les

pluricellulaires. Ainsi, un allongement du temps de réplication serait moins contre-sélectionné. Dans ces conditions, la « faible » efficacité de synthèse du dTMP pourrait être compensée par le ralentissement de la vitesse de réplication. Si les fourches de réplication chez les eucaryotes sont lentes [100 bp/sec], en revanche la vitesse de réplication ne l'est pas (Walther et Wold, 2001). En effet, la présence de plusieurs origines de réplication par chromosome, pouvant être répliquées simultanément, rend le temps de réplication peu dépendant de la quantité d'ADN à répliquer, donc peu dépendant de la taille du génome. La multiplication du nombre d'origines de réplication augmente virtuellement la vitesse de réplication, et permet aux eucaryotes d'avoir des temps de réplication comparable à certains procaryotes (Walther et Wold, 2001).

Comment alors les cellules eucaryotes ont pu dépasser le problème lié à la synthèse de dTMP par ThyA/FolA ? La solution la plus simple serait d'augmenter l'efficacité catalytique de ThyA/FolA, en stabilisant une forme active par des mutations. Chez certains protozoaires (Ivanetich et Santi, 1990) et chez les plantes dites supérieures (Lazar *et coll.*, 1993), on retrouve une fusion entre les gènes codant les protéines ThyA et FolA conduisant à un complexe enzymatique dans lequel le produit folique de la réaction catalysée par ThyA, le DHF, est directement pris en charge par la DHFR sans temps d'attente. Ce mécanisme appelé « substrate channeling » ne peut cependant pas fournir une explication générale compte tenu que ce type de fusion n'est pas commun à tous les eucaryotes, en particulier chez les mammifères. Une stratégie alternative et/ou complémentaire à l'augmentation de l'efficacité de la réaction de synthèse du thymidylate, est de réguler cette activité de façon transcriptionnelle ou traductionnelle. Nous avons vu dans l'introduction que chez les mammifères, et l'Homme en particulier, la transcription et la traduction de *thyA* sont augmentées au moment de l'entrée en phase S (voir introduction p 24). De plus des travaux réalisés sur des cellules ovariennes de hamster chinois (CHO) ont montré que l'activité dihydrofolate réductase (DHFR) augmente également durant la phase S (Mariani *et coll.*, 1981). Cette augmentation graduelle de l'activité DHFR suit une tendance parallèle à celle de la vitesse de réplication, suggérant que la vitesse de réplication chez les eucaryotes est également limitée par la disponibilité en nucléotides, en particulier en

thymine. Une étude réalisée sur des cellules humaines HeLa confirme cette proposition. Les auteurs ont montré que l'addition de désoxyribonucléotides induit une accélération de la réplication durant la phase S précoce (Malinsky *et coll.*, 2001). L'orchestration précise de la synthèse du dTMP et de la réplication suggère l'existence d'un système de contrôle. Vernis et collaborateurs ont mis en évidence que la cellule est capable de « sentir » l'absence de dTMP et qu'elle répond en bloquant, de façon réversible, son cycle en phase S (Vernis *et coll.*, 2003). Pour cela, les auteurs ont génétiquement modifié la levure *Saccharomyces cerevisiae*, afin de créer une souche auxotrophe pour la thymidine (*tdk<sup>+</sup>, ΔthyA*). Ils ont montré que la carence en dTTP dans la cellule induit la phosphorylation de la kinase Rad53, qui va activer le point de contrôle de la phase S (« S phase checkpoint »). L'ajout de thymidine permet aux cellules de redémarrer le cycle cellulaire rapidement.

Considérant les travaux ci-dessus, il semble que les eucaryotes peuvent répliquer de grandes quantités d'ADN à une vitesse élevée, grâce à l'existence de points de contrôle du cycle cellulaire et à une régulation fine des enzymes impliquées dans la synthèse des désoxyribonucléotides, en particulier la thymidylate synthase et la dihydrofolate réductase.

## **Le métabolisme des pyrimidines chez *H. pylori* : l'orotate phosphoribosyltransférase**

### **I. La thymidylate synthase ThyX : une protéine essentielle et solitaire ?**

Les résultats que j'ai pu obtenir concernant la thymidylate synthase ThyX chez *H. pylori* sont peu concluants. La première que l'on se posait était de savoir si ThyX était essentielle à la survie de la bactérie. Bien que je n'ai pas pu répondre directement à la question, plusieurs indices nous laissent penser que oui : 1) ThyX est essentielle chez *R. capsulatus* ; 2) la thymidine exogène n'est pas incorporé ; 3) Le premier inhibiteur spécifique de ThyX inhibe la croissance de *H. pylori*. Il est alors raisonnable de penser que l'activité de synthèse de dTMP est sous la seule dépendance de ThyX chez *H. pylori*.

La deuxième question était de savoir si ThyX avait des partenaires protéiques dans la cellule. Par deux approches complémentaires, je n'ai pas pu mettre en évidence de protéine interagissant spécifiquement avec ThyX. L'explication la plus simple est que ThyX est une protéine solitaire comme semble l'indiquer son rôle métabolique. Cependant on peut imaginer que dans certaines conditions ThyX interagit, peut-être de façon transitoire, avec d'autres protéines. Dans tous les cas, il est peu probable que cette interaction soit forte.

## II. Un double rôle pour les OPRTases bactériennes ?

L'orotate phosphoribosyltransférase (OPRTase) est une enzyme essentielle dans la voie *de novo* de formation des pyrimidines. Elle catalyse la réaction entre l'orotate et le 5-phosphoribosyl-1-pyrophosphate (PRPP) pour produire l'orotidine monophosphate (OMP). L'OMP est ensuite décarboxylé en uridine monophosphate (UMP) par l'OMP décarboxylase. Chez les procaryotes, ces deux étapes sont catalysées par deux enzymes distinctes, alors que chez les eucaryotes, elles sont l'œuvre d'un seul polypeptide portant les deux domaines, l'UMP synthase (Yablonski *et coll.*, 1996). Il est bien connu que chez ces derniers, l'UMP synthase est capable d'utiliser également l'uracile comme substrat pour former directement l'UMP. Ce phénomène est directement responsable de la sensibilité des cellules humaines au 5 fluorouracile (5FU) utilisé comme anticancéreux. J'ai montré, à travers les résultats présentés dans le chapitre III, que cette capacité à métaboliser l'uracile et le 5FU semblait être également partagée par les OPRTases bactériennes. Il semble cependant que cette activité UPRTase est accessoire et qu'elle ne représente vraisemblablement pas une réelle voie alternative à la voie *de novo*.

## III. Quelques perspectives...

### A. génétiques

Une façon de confirmer les résultats obtenus dans cette étude serait d'apporter la preuve directe que Hp-OPRTase est responsable de la sensibilité de *H. pylori* à de fortes concentrations de 5FU. Afin de le montrer, je propose de surexprimer la protéine dans la bactérie à l'aide d'un vecteur

d'expression inductible. Dans ces conditions, nous devrions observer une augmentation de la sensibilité au 5FU comme nous l'avons vu dans notre test chez *E. coli*. Il devrait être également possible dans ces conditions de bloquer la voie *de novo* en inactivant le gène codant l'OMP décarboxylase. J'ai amorcé la construction des outils nécessaires, mais des problèmes techniques de clonage ne m'ont pas permis de les mener à leur terme à ce jour.

### **B. biochimiques**

Il serait intéressant de rechercher s'il est possible d'augmenter l'affinité de Hp-OPRTase pour l'uracile. Pour cela je propose de mimer, de façon accélérée *in vitro*, le processus d'adaptation évolutionnaire en réalisant une évolution moléculaire dirigée par PCR de basse fidélité. Une banque de gènes *hp1257* aléatoirement mutés sera clonée dans un vecteur d'expression d'*E. coli*. La banque de plasmides sera alors transformée dans une souche dépourvue d'activité UPRTase comme la souche SØ6710 utilisée dans ce travail. Les transformants seront cultivés en milieu solide sur milieu minimum en présence d'uracile et d'une faible quantité d'IPTG. Cette quantité sera déterminée comme étant la quantité minimale d'inducteur nécessaire à la croissance normale d'une souche SØ6710 portant le gène codant l'UPRTase sur le plasmide. Seront alors sélectionnés les clones présentant une croissance accélérée. Le gène sera extrait du plasmide par digestion enzymatique et cloné dans le même vecteur vierge. Le plasmide sera transformé dans la souche SØ6710 afin de vérifier que le phénotype est bien dû à une ou plusieurs mutations dans la région codante. Les allèles obtenus seront séquencés afin de connaître la nature des mutations, puis transférés par échange allélique dans *H. pylori*. Nous pourrions alors tester la résistance au 5FU et la possibilité de croissance en utilisant l'uracile comme seule source de pyrimidines. L'activité des protéines correspondantes pourra également être testée *in vitro*.

## **REFERENCES BIBLIOGRAPHIQUES**



- Abby, S., and Daubin, V. (2007) Comparative genomics and the evolution of prokaryotes. *Trends Microbiol* **15**: 135-141.
- Agrawal, N., Lesley, S.A., Kuhn, P., and Kohen, A. (2004) Mechanistic studies of a flavin-dependent thymidylate synthase. *Biochemistry* **43**: 10295-10301.
- Ahmad, S.I., Kirk, S.H., and Eisenstark, A. (1998) Thymine metabolism and thymineless death in prokaryotes and eukaryotes. *Annu Rev Microbiol* **52**: 591-625.
- Albertson, N., Wenggren, I., and Sjoström, J.E. (1998) Growth and survival of *Helicobacter pylori* in defined medium and susceptibility to Brij 78. *J Clin Microbiol* **36**: 1232-1235.
- Allardet-Servent, A., Michaux-Charachon, S., Jumas-Bilak, E., Karayan, L., and Ramuz, M. (1993) Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J Bacteriol* **175**: 7869-7874.
- Andersen, P.S., Frees, D., Fast, R., and Mygind, B. (1995) Uracil uptake in *Escherichia coli* K-12: isolation of *uraA* mutants and cloning of the gene. *J Bacteriol* **177**: 2008-2013.
- Avery, O.T., MacLeod, C.M., and McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *The Journal of Experimental Medicine* **79**: 137-158.
- Batra, V., Kesavan, V., and Mishra, K.P. (2004) Modulation of enzymes involved in folate dependent one-carbon metabolism by gamma-radiation stress in mice. *J Radiat Res (Tokyo)* **45**: 527-533.
- Beadle, G.W., and Tatum, E.L. (1941) Genetic Control of Biochemical Reactions in *Neurospora*. *Proc Natl Acad Sci U S A* **27**: 499-506.
- Bentley, S.D., and Parkhill, J. (2004) Comparative genomic structure of prokaryotes. *Annu Rev Genet* **38**: 771-792.
- Bhatia, M.B., and Grubmeyer, C. (1993) The role of divalent magnesium in activating the reaction catalyzed by orotate phosphoribosyltransferase. *Arch Biochem Biophys* **303**: 321-325.
- Boneca, I.G., Ecobichon, C., Chaput, C., Mathieu, A., Guadagnini, S., Prevost, M.C., Colland, F., Labigne, A., and de Reuse, H. (2008) Development of inducible systems to engineer conditional mutants of essential genes of *Helicobacter pylori*. *Appl Environ Microbiol*.
- Brosius, J.L., and Colman, R.F. (2002) Three subunits contribute amino acids to the active site of tetrameric adenylosuccinate lyase: Lys268 and Glu275 are required. *Biochemistry* **41**: 2217-2226.
- Bunnak, J., Hamana, H., Ogino, Y., Saheki, T., Yamagishi, A., Oshima, T., Date, T., and Shinozawa, T. (1995) Orotate Phosphoribosyltransferase from *Thermus thermophilus*: Overexpression in *Escherichia coli*, Purification and Characterization. *Journal of Biochemistry* **118**: 1261.
- Burns, B.P., Hazell, S.L., Mendz, G.L., Kolesnikow, T., Tillett, D., and Neilan, B.A. (2000) The *Helicobacter pylori* *pyrB* gene encoding aspartate carbamoyltransferase is essential for bacterial survival. *Arch Biochem Biophys* **380**: 78-84.
- Carreras, C.W., and Santi, D.V. (1995) The catalytic mechanism and structure of thymidylate synthase. *Annu Rev Biochem* **64**: 721-762.
- Cavalier-Smith, T. (1978) Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. *J Cell Sci* **34**: 247-278.
- Chalker, A.F., Minehart, H.W., Hughes, N.J., Koretke, K.K., Lonetto, M.A., Brinkman, K.K., Warren, P.V., Lupas, A., Stanhope, M.J., Brown, J.R., and Hoffman, P.S. (2001) Systematic identification of selective essential genes in *Helicobacter pylori* by genome prioritization and allelic replacement mutagenesis. *J Bacteriol* **183**: 1259-1268.
- Chargaff, E. (1950) Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**: 201-240.
- Chase, M., and Hershey, A. (1952) Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage. *Journal of General Physiology* **36**: 39-56.
- Chernyshev, A., Fleischmann, T., Koehn, E.M., Lesley, S.A., and Kohen, A. (2007) The relationships between oxidase and synthase activities of flavin dependent thymidylate synthase (FDTS). *Chem Commun (Camb)*: 2861-2863.

- Cho, S., Lee, S., Lee, Y., and Rhie, H. (2006) Novel Thymidylate Synthase, ThyX, Endows an Intrinsic Resistance to Trimethoprim in *Corynebacterium glutamicum* ATCC 13032. In *International Meeting of the Microbiological Society of Korea* Korea.
- Chu, J., and Dolnick, B.J. (2002) Natural antisense (rTS $\alpha$ ) RNA induces site-specific cleavage of thymidylate synthase mRNA. *Biochim Biophys Acta* **1587**: 183-193.
- Cohen, S.S., and Barner, H.D. (1954) Studies On Unbalanced Growth In *Escherichia Coli*. *Proc Natl Acad Sci U S A* **40**: 885-893.
- Copeland, R.A., Marcinkeviciene, J., Haque, T.S., Kopcho, L.M., Jiang, W., Wang, K., Ecret, L.D., Sizemore, C., Amsler, K.A., Foster, L., Tadesse, S., Combs, A.P., Stern, A.M., Trainor, G.L., Slee, A., Rogers, M.J., and Hobbs, F. (2000) *Helicobacter pylori*-selective antibacterials based on inhibition of pyrimidine biosynthesis. *J Biol Chem* **275**: 33373-33378.
- Dahm, R. (2005) Friedrich Miescher and the discovery of DNA. *Developmental Biology* **278**: 274-288.
- Deering, R.A., and Michrina, C.A. (1982) Inhibitors of DNA precursor metabolism in *Dictyostelium discoideum*. *Antimicrob Agents Chemother* **21**: 764-769.
- Derbise, A., Lesic, B., Dacheux, D., Ghigo, J.M., and Carniel, E. (2003) A rapid and simple method for inactivating chromosomal genes in *Yersinia*. *FEMS Immunol Med Microbiol* **38**: 113-116.
- Duckworth, M., Menard, A., Megraud, F., and Mendz, G.L. (2006) Bioinformatic analysis of *Helicobacter pylori* XGPRTase: a potential therapeutic target. *Helicobacter* **11**: 287-295.
- Dynes, J.L., and Firtel, R.A. (1989) Molecular complementation of a genetic marker in *Dictyostelium* using a genomic DNA library. *Proc Natl Acad Sci U S A* **86**: 7966-7970.
- Eslick, G.D. (2000) *Helicobacter pylori* infection transmitted sexually via oral-genital contact: a hypothetical model. *Sex Transm Infect* **76**: 489-492.
- Faith, J.J., Driscoll, M.E., Fusaro, V.A., Cosgrove, E.J., Hayete, B., Juhn, F.S., Schneider, S.J., and Gardner, T.S. (2008) Many Microbe Microarrays Database: uniformly normalized Affymetrix compendia with structured experimental metadata. *Nucleic Acids Res* **36**: D866-870.
- Ferdows, M.S., and Barbour, A.G. (1989) Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. *Proc Natl Acad Sci U S A* **86**: 5969-5973.
- Fey, P., Kowal, A.S., Gaudet, P., Pilcher, K.E., and Chisholm, R.L. (2007) Protocols for growth and development of *Dictyostelium discoideum*. *Nat Protoc* **2**: 1307-1316.
- Finer-Moore, J.S., Santi, D.V., and Stroud, R.M. (2003) Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. *Biochemistry* **42**: 248-256.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., and et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496-512.
- Foerster, K.U., von Mering, C., Hooper, S.D., and Bork, P. (2005) Environments shape the nucleotide composition of genomes. *EMBO Rep* **6**: 1208-1213.
- Friedkin, M., and Roberts, D. (1956) Conversion of uracil deoxyriboside to thymidine of deoxyribonucleic acid. *J Biol Chem* **220**: 653-660.
- Gan, K., Sankaran, K., Williams, M.G., Aldea, M., Rudd, K.E., Kushner, S.R., and Wu, H.C. (1995) The *umpA* gene of *Escherichia coli* encodes phosphatidylglycerol:prolipoprotein diacylglycerol transferase (lgt) and regulates thymidylate synthase levels through translational coupling. *J Bacteriol* **177**: 1879-1882.
- Giladi, M., Bitan-Banin, G., Mevarech, M., and Ortenberg, R. (2002) Genetic evidence for a novel thymidylate synthase in the halophilic archaeon *Halobacterium salinarum* and in *Campylobacter jejuni*. *FEMS Microbiol Lett* **216**: 105-109.
- Giladi, M., Altman-Price, N., Levin, I., Levy, L., and Mevarech, M. (2003) FolM, a new chromosomally encoded dihydrofolate reductase in *Escherichia coli*. *J Bacteriol* **185**: 7015-7018.
- Glass, J.I., Lefkowitz, E.J., Glass, J.S., Heiner, C.R., Chen, E.Y., and Cassell, G.H. (2000) The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature* **407**: 757-762.

- Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S.G. (1996) Life with 6000 genes. *Science* **274**: 546, 563-547.
- Graziani, S., Xia, Y., Gurnon, J.R., Van Etten, J.L., Leduc, D., Skouloubris, S., Myllykallio, H., and Liebl, U. (2004) Functional analysis of FAD-dependent thymidylate synthase ThyX from *Paramecium bursaria* Chloroella virus-1. *J Biol Chem* **279**: 54340-54347.
- Graziani, S., Bernauer, J., Skouloubris, S., Graille, M., Zhou, C.Z., Marchand, C., Decottignies, P., van Tilbeurgh, H., Myllykallio, H., and Liebl, U. (2006) Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J Biol Chem* **281**: 24048-24057.
- Gregory, T.R. (2005) Synergy between sequence and size in large-scale genomics. *Nat Rev Genet* **6**: 699-708.
- Griffin, J., Roshick, C., Iliffe-Lee, E., and McClarty, G. (2005) Catalytic mechanism of *Chlamydia trachomatis* flavin-dependent thymidylate synthase. *J Biol Chem* **280**: 5456-5467.
- Griffith, F. (1928) The significance of pneumococcal types. *Journal of Hygiene* **27**: 113-159.
- Guarino, E., Salguero, I., Jimenez-Sanchez, A., and Guzman, E.C. (2007) Double-strand break generation under deoxyribonucleotide starvation in *Escherichia coli*. *J Bacteriol* **189**: 5782-5786.
- Hallin, P.F., and Ussery, D.W. (2004) CBS Genome Atlas Database: a dynamic storage for bioinformatic results and sequence data. *Bioinformatics* **20**: 3682-3686.
- Heidelberger, C., Chaudhuri, N.K., Danneberg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schnitzer, R.J., Plevin, E., and Scheiner, J. (1957) Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature* **179**: 663-666.
- Helmstetter, C. (1996) Timing of synthetic activities in the cell cycle. In *Escherichia coli and Salmonella typhimurium; Cellular and Molecular Biology*. al, N.F.e. (ed). Washington, DC: ASM, pp. 1627-1639.
- Henriksen, A., Aghajari, N., Jensen, K.F., and Gajhede, M. (1996) A flexible loop at the dimer interface is a part of the active site of the adjacent monomer of *Escherichia coli* orotate phosphoribosyltransferase. *Biochemistry* **35**: 3803-3809.
- Heuermann, D., and Haas, R. (1998) A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation. *Mol Gen Genet* **257**: 519-528.
- Hoagland, M.B., Stephenson, M.L., Scott, J.F., Hecht, L.I., and Zamecnik, P.C. (1958) A soluble ribonucleic acid intermediate in protein synthesis. *J Biol Chem* **231**: 241-257.
- Ivanetich, K.M., and Santi, D.V. (1990) Bifunctional thymidylate synthase-dihydrofolate reductase in protozoa. *Faseb J* **4**: 1591-1597.
- Jacob, F., and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-356.
- Jensen, K.F., and Mygind, B. (1996) Different oligomeric states are involved in the allosteric behavior of uracil phosphoribosyltransferase from *Escherichia coli*. *Eur J Biochem* **240**: 637-645.
- Jensen, K.F., Arent, S., Larsen, S., and Schack, L. (2005) Allosteric properties of the GTP activated and CTP inhibited uracil phosphoribosyltransferase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Febs J* **272**: 1440-1453.
- Ju, J., Pedersen-Lane, J., Maley, F., and Chu, E. (1999) Regulation of p53 expression by thymidylate synthase. *Proc Natl Acad Sci U S A* **96**: 3769-3774.
- Kanai, A., Sato, A., Imoto, J., and Tomita, M. (2006) Archaeal *Pyrococcus furiosus* thymidylate synthase 1 is an RNA-binding protein. *Biochem J* **393**: 373-379.
- Khorana, H.G. (1965) Polynucleotide synthesis and the genetic code. *Fed Proc* **24**: 1473-1487.
- Kieser, H.M., Kieser, T., and Hopwood, D.A. (1992) A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. *J Bacteriol* **174**: 5496-5507.

- Kim, M.K., Song, H.E., Kim, Y.S., Rho, S.H., Im, Y.J., Lee, J.H., Kang, G.B., and Eom, S.H. (2003) Crystallization and preliminary X-ray crystallographic analysis of orotate phosphoribosyltransferase from *Helicobacter pylori*. *Mol Cells* **15**: 361-363.
- Kompis, I.M., Islam, K., and Then, R.L. (2005) DNA and RNA synthesis: antifolates. *Chem Rev* **105**: 593-620.
- Kuhn, P., Lesley, S.A., Mathews, II, Canaves, J.M., Brinen, L.S., Dai, X., Deacon, A.M., Elsliger, M.A., Eshaghi, S., Floyd, R., Godzik, A., Grittini, C., Grzechnik, S.K., Guda, C., Hodgson, K.O., Jaroszewski, L., Karlak, C., Klock, H.E., Koesema, E., Kovarik, J.M., Kreuzsch, A.T., McMullan, D., McPhillips, T.M., Miller, M.A., Miller, M., Morse, A., Moy, K., Ouyang, J., Robb, A., Rodrigues, K., Selby, T.L., Spraggon, G., Stevens, R.C., Taylor, S.S., van den Bedem, H., Velasquez, J., Vincent, J., Wang, X., West, B., Wolf, G., Wooley, J., and Wilson, I.A. (2002) Crystal structure of thy1, a thymidylate synthase complementing protein from *Thermotoga maritima* at 2.25 Å resolution. *Proteins* **49**: 142-145.
- Lander, E.S., and Linton, L.M., and Birren, B., and Nusbaum, C., and Zody, M.C., and Baldwin, J., and Devon, K., and Dewar, K., and Doyle, M., and FitzHugh, W., and Funke, R., and Gage, D., and Harris, K., and Heaford, A., and Howland, J., and Kann, L., and Lehoczky, J., and LeVine, R., and McEwan, P., and McKernan, K., and Meldrim, J., and Mesirov, J.P., and Miranda, C., and Morris, W., and Naylor, J., and Raymond, C., and Rosetti, M., and Santos, R., and Sheridan, A., and Sougnez, C., and Stange-Thomann, N., and Stojanovic, N., and Subramanian, A., and Wyman, D., and Rogers, J., and Sulston, J., and Ainscough, R., and Beck, S., and Bentley, D., and Burton, J., and Clee, C., and Carter, N., and Coulson, A., and Deadman, R., and Deloukas, P., and Dunham, A., and Dunham, I., and Durbin, R., and French, L., and Grafham, D., and Gregory, S., and Hubbard, T., and Humphray, S., and Hunt, A., and Jones, M., and Lloyd, C., and McMurray, A., and Matthews, L., and Mercer, S., and Milne, S., and Mullikin, J.C., and Mungall, A., and Plumb, R., and Ross, M., and Shownkeen, R., and Sims, S., and Waterston, R.H., and Wilson, R.K., and Hillier, L.W., and McPherson, J.D., and Marra, M.A., and Mardis, E.R., and Fulton, L.A., and Chinwalla, A.T., and Pepin, K.H., and Gish, W.R., and Chissole, S.L., and Wendl, M.C., and Delehaunty, K.D., and Miner, T.L., and Delehaunty, A., and Kramer, J.B., and Cook, L.L., and Fulton, R.S., and Johnson, D.L., and Minx, P.J., and Clifton, S.W., and Hawkins, T., and Branscomb, E., and Predki, P., and Richardson, P., and Wenning, S., and Slezak, T., and Doggett, N., and Cheng, J.F., and Olsen, A., and Lucas, S., and Elkin, C., and Uberbacher, E., and Frazier, M., and Gibbs, R.A., and Muzny, D.M., and Scherer, S.E., and Bouck, J.B., and Sodergren, E.J., and Worley, K.C., and Rives, C.M., and Gorrell, J.H., and Metzker, M.L., and Naylor, S.L., and Kucherlapati, R.S., and Nelson, D.L., and Weinstock, G.M., and Sakaki, Y., and Fujiyama, A., and Hattori, M., and Yada, T., and Toyoda, A., and Itoh, T., and Kawagoe, C., and Watanabe, H., and Totoki, Y., and Taylor, T., and Weissenbach, J., and Heilig, R., and Saurin, W., and Artiguenave, F., and Brottier, P., and Bruls, T., and Pelletier, E., and Robert, C., and Wincker, P., and Smith, D.R., and Doucette-Stamm, L., and Rubenfield, M., and Weinstock, K., and Lee, H.M., and Dubois, J., and Rosenthal, A., and Platzer, M., and Nyakatura, G., and Taudien, S., and Rump, A., and Yang, H., and Yu, J., and Wang, J., and Huang, G., and Gu, J., and Hood, L., and Rowen, L., and Madan, A., and Qin, S., and Davis, R.W., and Federspiel, N.A., and Abola, A.P., and Proctor, M.J., and Myers, R.M., and Schmutz, J., and Dickson, M., and Grimwood, J., and Cox, D.R., and Olson, M.V., and Kaul, R., and Raymond, C., and Shimizu, N., and Kawasaki, K., and Minoshima, S., and Evans, G.A., and Athanasiou, M., and Schultz, R., and Roe, B.A., and Chen, F., and Pan, H., and Ramser, J., and Lehrach, H., and Reinhardt, R., and McCombie, W.R., and de la Bastide, M., and Dedhia, N., and Blocker, H., and Hornischer, K., and Nordsiek, G., and Agarwala, R., and Aravind, L., and Bailey, J.A., and Bateman, A., and Batzoglou, S., and Birney, E., and Bork, P., and Brown, D.G., and Burge, C.B., and Cerutti, L., and Chen, H.C., and Church, D., and Clamp, M., and Copley, R.R., and Doerks, T., and Eddy, S.R., and Eichler, E.E., and Furey, T.S., and Galagan, J., and Gilbert, J.G., and Harmon, C., and Hayashizaki, Y., and Haussler, D., and Hermjakob, H.,

- and Hokamp, K., and Jang, W., and Johnson, L.S., and Jones, T.A., and Kasif, S., and Kasprzyk, A., and Kennedy, S., and Kent, W.J., and Kitts, P., and Koonin, E.V., and Korf, I., and Kulp, D., and Lancet, D., and Lowe, T.M., and McLysaght, A., and Mikkelsen, T., and Moran, J.V., and Mulder, N., and Pollara, V.J., and Ponting, C.P., and Schuler, G., and Schultz, J., and Slater, G., and Smit, A.F., and Stupka, E., and Szustakowski, J., and Thierry-Mieg, D., and Thierry-Mieg, J., and Wagner, L., and Wallis, J., and Wheeler, R., and Williams, A., and Wolf, Y.I., and Wolfe, K.H., and Yang, S.P., and Yeh, R.F., and Collins, F., and Guyer, M.S., and Peterson, J., and Felsenfeld, A., and Wetterstrand, K.A., and Patrinos, A., and Morgan, M.J., and de Jong, P., and Catanese, J.J., and Osoegawa, K., and Shizuya, H., and Choi, S., and Chen, Y.J. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- Lazar, G., Zhang, H., and Goodman, H.M. (1993) The origin of the bifunctional dihydrofolate reductase-thymidylate synthase isogenes of *Arabidopsis thaliana*. *Plant J* **3**: 657-668.
- Leduc, D. (2004) Caractérisation par des études génétiques et biochimiques d'une nouvelle thymidylate synthase, ThyX. Vol. Thèse de doctorat Orsay: Université Paris XI, pp. 118.
- Leduc, D., Graziani, S., Lipowski, G., Marchand, C., Le Marechal, P., Liebl, U., and Myllykallio, H. (2004a) 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.
- Leduc, D., Graziani, S., Meslet-Cladiere, L., Sodolescu, A., Liebl, U., and Myllykallio, H. (2004b) Two distinct pathways for thymidylate (dTMP) synthesis in (hyper)thermophilic Bacteria and Archaea. *Biochem Soc Trans* **32**: 231-235.
- Liao, Z.Y., Sordet, O., Zhang, H.L., Kohlhagen, G., Antony, S., Gmeiner, W.H., and Pommier, Y. (2005) A novel polypyrimidine antitumor agent FdUMP[10] induces thymineless death with topoisomerase I-DNA complexes. *Cancer Res* **65**: 4844-4851.
- Liu, J., Schmitz, J.C., Lin, X., Tai, N., Yan, W., Farrell, M., Bailly, M., Chen, T., and Chu, E. (2002) Thymidylate synthase as a translational regulator of cellular gene expression. *Biochim Biophys Acta* **1587**: 174-182.
- Lundegaard, C., and Jensen, K.F. (1999) Kinetic mechanism of uracil phosphoribosyltransferase from *Escherichia coli* and catalytic importance of the conserved proline in the PRPP binding site. *Biochemistry* **38**: 3327-3334.
- Malinsky, J., Koberna, K., Stanek, D., Masata, M., Votruba, I., and Raska, I. (2001) The supply of exogenous deoxyribonucleotides accelerates the speed of the replication fork in early S-phase. *J Cell Sci* **114**: 747-750.
- Marais, A., Mendz, G.L., Hazell, S.L., and Megraud, F. (1999) Metabolism and genetics of *Helicobacter pylori*: the genome era. *Microbiol Mol Biol Rev* **63**: 642-674.
- Marashi, S.A., and Ghalanbor, Z. (2004) Correlations between genomic GC levels and optimal growth temperatures are not 'robust'. *Biochem Biophys Res Commun* **325**: 381-383.
- Mariani, B.D., Slate, D.L., and Schimke, R.T. (1981) S phase-specific synthesis of dihydrofolate reductase in Chinese hamster ovary cells. *Proc Natl Acad Sci U S A* **78**: 4985-4989.
- McEwan, C.E., Gatherer, D., and McEwan, N.R. (1998) Nitrogen-fixing aerobic bacteria have higher genomic GC content than non-fixing species within the same genus. *Hereditas* **128**: 173-178.
- Megraud, F. (2003) [When and how does *Helicobacter pylori* infection occur?]. *Gastroenterol Clin Biol* **27**: 374-379.
- Menard, R., Sansonetti, P.J., and Parsot, C. (1993) Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J Bacteriol* **175**: 5899-5906.
- Menz, G.L., Jimenez, B.M., Hazell, S.L., Gero, A.M., and O'Sullivan, W.J. (1994) De novo synthesis of pyrimidine nucleotides by *Helicobacter pylori*. *J Appl Bacteriol* **77**: 1-8.
- Merino, E., and Yanofsky, C. (2005) Transcription attenuation: a highly conserved regulatory strategy used by bacteria. *Trends Genet* **21**: 260-264.
- Meselson, M., and Stahl, F.W. (1958) The Replication Of Dna In *Escherichia Coli*. *Proc Natl Acad Sci U S A* **44**: 671-682.

- Mira, A., Ochman, H., and Moran, N.A. (2001) Deletional bias and the evolution of bacterial genomes. *Trends Genet* **17**: 589-596.
- Mitchell, D. (2007) GC content and genome length in Chargaff compliant genomes. *Biochem Biophys Res Commun* **353**: 207-210.
- Montfort, W.R., Perry, K.M., Fauman, E.B., Finer-Moore, J.S., Maley, G.F., Hardy, L., Maley, F., and Stroud, R.M. (1990) Structure, multiple site binding, and segmental accommodation in thymidylate synthase on binding dUMP and an anti-folate. *Biochemistry* **29**: 6964-6977.
- Moran, N.A. (2002) Microbial minimalism: genome reduction in bacterial pathogens. *Cell* **108**: 583-586.
- Moran, N.A., and Plague, G.R. (2004) Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* **14**: 627-633.
- Morganroth, P.A., and Hanawalt, P.C. (2006) Role of DNA replication and repair in thymineless death in *Escherichia coli*. *J Bacteriol* **188**: 5286-5288.
- Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Buillard, V., Cerutti, L., Copley, R., Courcelle, E., Das, U., Daugherty, L., Dibley, M., Finn, R., Fleischmann, W., Gough, J., Haft, D., Hulo, N., Hunter, S., Kahn, D., Kanapin, A., Kejariwal, A., Labarga, A., Langendijk-Genevaux, P.S., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Orengo, C., Petryszak, R., Selengut, J.D., Sigrist, C.J.A., Thomas, P.D., Valentin, F., Wilson, D., Wu, C.H., and Yeats, C. (2007) New developments in the InterPro database. Vol. 35, pp. D224-228.
- Munch, R., Hiller, K., Barg, H., Heldt, D., Linz, S., Wingender, E., and Jahn, D. (2003) PRODORIC: prokaryotic database of gene regulation. *Nucleic Acids Res* **31**: 266-269.
- Murzin, A.G. (2002) Biochemistry. DNA building block reinvented. *Science* **297**: 61-62.
- Musto, H., Naya, H., Zavala, A., Romero, H., Alvarez-Valin, F., and Bernardi, G. (2004) Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. *FEBS Lett* **573**: 73-77.
- Musto, H., Naya, H., Zavala, A., Romero, H., Alvarez-Valin, F., and Bernardi, G. (2005) The correlation between genomic G+C and optimal growth temperature of prokaryotes is robust: a reply to Marashi and Ghalanbor. *Biochem Biophys Res Commun* **330**: 357-360.
- Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**: 105-107.
- Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H.E., Moran, N.A., and Hattori, M. (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314**: 267.
- Naya, H., Romero, H., Zavala, A., Alvarez, B., and Musto, H. (2002) Aerobiosis increases the genomic guanine plus cytosine content (GC%) in prokaryotes. *J Mol Evol* **55**: 260-264.
- Nijhout, H.F., Reed, M.C., Budu, P., and Ulrich, C.M. (2004) A mathematical model of the folate cycle: new insights into folate homeostasis. *J Biol Chem* **279**: 55008-55016.
- Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F., and O'Neal, C. (1965) RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proc Natl Acad Sci U S A* **53**: 1161-1168.
- Ochman, H. (2002) Bacterial evolution: chromosome arithmetic and geometry. *Curr Biol* **12**: R427-428.
- Peters, G.J., Laurensse, E., Leyva, A., Lankelma, J., and Pinedo, H.M. (1986) Sensitivity of human, murine, and rat cells to 5-fluorouracil and 5'-deoxy-5-fluorouridine in relation to drug-metabolizing enzymes. *Cancer Res* **46**: 20-28.
- Peters, G.J., Backus, H.H., Freemantle, S., van Triest, B., Codacci-Pisanelli, G., van der Wilt, C.L., Smid, K., Lunec, J., Calvert, A.H., Marsh, S., McLeod, H.L., Bloemena, E., Meijer, S., Jansen, G., van Groenigen, C.J., and Pinedo, H.M. (2002) Induction of thymidylate synthase as a 5-fluorouracil resistance mechanism. *Biochim Biophys Acta* **1587**: 194-205.

- Powell, C.M., Rudge, T.L., Zhu, Q., Johnson, L.F., and Hansen, U. (2000) Inhibition of the mammalian transcription factor LSF induces S-phase-dependent apoptosis by downregulating thymidylate synthase expression. *Embo J* **19**: 4665-4675.
- Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**: 218-229.
- Rain, J.C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., Chemama, Y., Labigne, A., and Legrain, P. (2001) The protein-protein interaction map of *Helicobacter pylori*. *Nature* **409**: 211-215.
- Ranea, J.A., Grant, A., Thornton, J.M., and Orengo, C.A. (2005) Microeconomic principles explain an optimal genome size in bacteria. *Trends Genet* **21**: 21-25.
- Rengarajan, J., Sasseti, C.M., Naroditskaya, V., Sloutsky, A., Bloom, B.R., and Rubin, E.J. (2004) The folate pathway is a target for resistance to the drug para-aminosalicylic acid (PAS) in mycobacteria. *Mol Microbiol* **53**: 275-282.
- Reyes, P. (1969) The synthesis of 5-fluorouridine 5'-phosphate by a pyrimidine phosphoribosyltransferase of mammalian origin. I. Some properties of the enzyme from P1534J mouse leukemic cells. *Biochemistry* **8**: 2057-2062.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**: 1030-1032.
- Rocha, E.P., and Danchin, A. (2002) Base composition bias might result from competition for metabolic resources. *Trends Genet* **18**: 291-294.
- Sanford, R.A., Cole, J.R., and Tiedje, J.M. (2002) Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic myxobacterium. *Appl Environ Microbiol* **68**: 893-900.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-5467.
- Sat, B., Reches, M., and Engelberg-Kulka, H. (2003) The *Escherichia coli* mazEF suicide module mediates thymineless death. *J Bacteriol* **185**: 1803-1807.
- Scapin, G., Grubmeyer, C., and Sacchettini, J.C. (1994) Crystal structure of orotate phosphoribosyltransferase. *Biochemistry* **33**: 1287-1294.
- S eraphin, B., Puig, O., Bouveret, E., Rutz, B., and Caspary, F. (2001) Tandem Affinity Purification to Enhance Interacting Protein Identification. In *Protein-Protein Interactions: A Molecular Cloning Manual*. Golemis, E.a.A., FD. (ed): Cold Spring Harbor Laboratory Press, pp. 682.
- Shimosaka, M., Fukuda, Y., Murata, K., and Kimura, A. (1985) Purification and Properties of Orotate Phosphoribosyltransferases from *Escherichia coli* K-12, and Its Derivative Purine-Sensitive Mutant. *J Biochem* **98**: 1689-1697.
- Singer, C.E., and Ames, B.N. (1970) Sunlight ultraviolet and bacterial DNA base ratios. *Science* **170**: 822-825.
- Sinha, S.C., and Smith, J.L. (2001) The PRT protein family. *Curr Opin Struct Biol* **11**: 733-739.
- Skouloubris, S., Thiberge, J.M., Labigne, A., and De Reuse, H. (1998) The *Helicobacter pylori* Urel protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infect Immun* **66**: 4517-4521.
- Sueoka, N. (1962) On the genetic basis of variation and heterogeneity of DNA base composition. *Proc Natl Acad Sci U S A* **48**: 582-592.
- Testerman, T.L., McGee, D.J., and Mobley, H.L. (2001) *Helicobacter pylori* growth and urease detection in the chemically defined medium Ham's F-12 nutrient mixture. *J Clin Microbiol* **39**: 3842-3850.
- Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K.,

- Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Waththey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M., and Venter, J.C. (1997) The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**: 539-547.
- Ulmer, J.E., Boum, Y., Thouvenel, C.D., Myllykallio, H., and Sibley, C.H. (2008) Functional analysis of the *Mycobacterium tuberculosis* FAD-dependent thymidylate synthase, ThyX, reveals new amino acid residues contributing to an extended ThyX motif. *J Bacteriol* **190**: 2056-2064.
- Vellai, T., Takacs, K., and Vida, G. (1998) A new aspect to the origin and evolution of eukaryotes. *J Mol Evol* **46**: 499-507.
- Vellai, T., and Vida, G. (1999) The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells. *Proc Biol Sci* **266**: 1571-1577.
- Venter, J.C., and Adams, M.D., and Myers, E.W., and Li, P.W., and Mural, R.J., and Sutton, G.G., and Smith, H.O., and Yandell, M., and Evans, C.A., and Holt, R.A., and Gocayne, J.D., and Amanatides, P., and Ballew, R.M., and Huson, D.H., and Wortman, J.R., and Zhang, Q., and Kodira, C.D., and Zheng, X.H., and Chen, L., and Skupski, M., and Subramanian, G., and Thomas, P.D., and Zhang, J., and Gabor Miklos, G.L., and Nelson, C., and Broder, S., and Clark, A.G., and Nadeau, J., and McKusick, V.A., and Zinder, N., and Levine, A.J., and Roberts, R.J., and Simon, M., and Slayman, C., and Hunkapiller, M., and Bolanos, R., and Delcher, A., and Dew, I., and Fasulo, D., and Flanigan, M., and Florea, L., and Halpern, A., and Hannenhalli, S., and Kravitz, S., and Levy, S., and Mobarry, C., and Reinert, K., and Remington, K., and Abu-Threideh, J., and Beasley, E., and Biddick, K., and Bonazzi, V., and Brandon, R., and Cargill, M., and Chandramouliswaran, I., and Charlab, R., and Chaturvedi, K., and Deng, Z., and Di Francesco, V., and Dunn, P., and Eilbeck, K., and Evangelista, C., and Gabrielian, A.E., and Gan, W., and Ge, W., and Gong, F., and Gu, Z., and Guan, P., and Heiman, T.J., and Higgins, M.E., and Ji, R.R., and Ke, Z., and Ketchum, K.A., and Lai, Z., and Lei, Y., and Li, Z., and Li, J., and Liang, Y., and Lin, X., and Lu, F., and Merkulov, G.V., and Milshina, N., and Moore, H.M., and Naik, A.K., and Narayan, V.A., and Neelam, B., and Nuskern, D., and Rusch, D.B., and Salzberg, S., and Shao, W., and Shue, B., and Sun, J., and Wang, Z., and Wang, A., and Wang, X., and Wang, J., and Wei, M., and Wides, R., and Xiao, C., and Yan, C., and Yao, A., and Ye, J., and Zhan, M., and Zhang, W., and Zhang, H., and Zhao, Q., and Zheng, L., and Zhong, F., and Zhong, W., and Zhu, S., and Zhao, S., and Gilbert, D., and Baumhueter, S., and Spier, G., and Carter, C., and Cravchik, A., and Woodage, T., and Ali, F., and An, H., and Awe, A., and Baldwin, D., and Baden, H., and Barnstead, M., and Barrow, I., and Beeson, K., and Busam, D., and Carver, A., and Center, A., and Cheng, M.L., and Curry, L., and Danaher, S., and Davenport, L., and Desilets, R., and Dietz, S., and Dodson, K., and Doup, L., and Ferreira, S., and Garg, N., and Gluecksmann, A., and Hart, B., and Haynes, J., and Haynes, C., and Heiner, C., and Hladun, S., and Hostin, D., and Houck, J., and Howland, T., and Ibegwam, C., and Johnson, J., and Kalush, F., and Kline, L., and Koduru, S., and Love, A., and Mann, F., and May, D., and McCawley, S., and McIntosh, T., and McMullen, I., and Moy, M., and Moy, L., and Murphy, B., and Nelson, K., and Pfannkoch, C., and Pratts, E., and Puri, V., and Qureshi, H., and Reardon, M., and Rodriguez, R., and Rogers, Y.H., and Romblad, D., and Ruhfel, B., and Scott, R., and Sitter, C., and Smallwood, M., and Stewart, E., and Strong, R., and Suh, E., and Thomas, R., and Tint, N.N., and Tse, S., and Vech, C., and Wang, G., and Wetter, J., and Williams, S., and Williams, M., and Windsor, S., and Winn-Deen, E., and Wolfe, K., and Zaveri, J., and Zaveri, K., and Abril, J.F., and Guigo, R., and Campbell, M.J., and Sjolander, K.V., and Karlak, B., and Kejariwal, A., and Mi, H., and Lazareva, B., and Hatton, T., and Narechania, A., and Diemer, K., and Muruganujan, A., and Guo, N., and Sato, S., and Bafna, V., and Istrail, S., and Lippert, R., and Schwartz, R., and Walenz, B., and Yooseph, S., and Allen, D., and Basu, A., and Baxendale, J., and Blick, L., and Caminha, M., and Carnes-Stine, J., and Caulk, P., and Chiang, Y.H., and Coyne, M., and Dahlke, C., and Mays, A., and Dombroski, M., and Donnelly,



- M., and Ely, D., and Esparham, S., and Fosler, C., and Gire, H., and Glanowski, S., and Glasser, K., and Glodek, A., and Gorokhov, M., and Graham, K., and Gropman, B., and Harris, M., and Heil, J., and Henderson, S., and Hoover, J., and Jennings, D., and Jordan, C., and Jordan, J., and Kasha, J., and Kagan, L., and Kraft, C., and Levitsky, A., and Lewis, M., and Liu, X., and Lopez, J., and Ma, D., and Majoros, W., and McDaniel, J., and Murphy, S., and Newman, M., and Nguyen, T., and Nguyen, N., and Nodell, M., and Pan, S., and Peck, J., and Peterson, M., and Rowe, W., and Sanders, R., and Scott, J., and Simpson, M., and Smith, T., and Sprague, A., and Stockwell, T., and Turner, R., and Venter, E., and Wang, M., and Wen, M., and Wu, D., and Wu, M., and Xia, A., and Zandieh, A., and Zhu, X. (2001) The sequence of the human genome. *Science* **291**: 1304-1351.
- Vernis, L., Piskur, J., and Diffley, J.F. (2003) Reconstitution of an efficient thymidine salvage pathway in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **31**: e120.
- Walther, A., and Wold, M. (2001) Eukaryotic Replication Fork. In *ENCYCLOPEDIA OF LIFE SCIENCES*. John Wiley & Sons, L.C.h.w.e.n. (ed), pp. 1-8.
- Wang, H.C., Susko, E., and Roger, A.J. (2006) On the correlation between genomic G+C content and optimal growth temperature in prokaryotes: data quality and confounding factors. *Biochem Biophys Res Commun* **342**: 681-684.
- Wang, X., Ma, C., Wang, X., and Xu, P. (2007) Orotate Phosphoribosyltransferase from *Corynebacterium ammoniagenes* Lacking a Conserved Lysine. *J. Bacteriol.* **189**: 9030-9036.
- Watson, J.D., and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**: 737-738.
- Wilquet, V., Van de Castele, M., Gigot, D., Legrain, C., and Glansdorff, N. (2004) Dihydropteridine reductase as an alternative to dihydrofolate reductase for synthesis of tetrahydrofolate in *Thermus thermophilus*. *J Bacteriol* **186**: 351-355.
- Yablonski, M.J., Pasek, D.A., Han, B.D., Jones, M.E., and Traut, T.W. (1996) Intrinsic activity and stability of bifunctional human UMP synthase and its two separate catalytic domains, orotate phosphoribosyltransferase and orotidine-5'-phosphate decarboxylase. *J Biol Chem* **271**: 10704-10708.
- Zhao, X., Zhang, Z., Yan, J., and Yu, J. (2007) GC content variability of eubacteria is governed by the pol III alpha subunit. *Biochem Biophys Res Commun* **356**: 20-25.