



**HAL**  
open science

# ADN polymérasés et acides nucléiques endommagés chez les Archaea : maintenance génomique et Biotechnologies

Céline Ralec

## ► To cite this version:

Céline Ralec. ADN polymérasés et acides nucléiques endommagés chez les Archaea : maintenance génomique et Biotechnologies. Biochimie [q-bio.BM]. Université de Bretagne occidentale - Brest, 2013. Français. NNT : 2013BRES0057 . tel-02089275v2

**HAL Id: tel-02089275**

**<https://theses.hal.science/tel-02089275v2>**

Submitted on 5 Apr 2019

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

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



université de bretagne  
occidentale



**THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE**  
*sous le sceau de l'Université européenne de Bretagne*

pour obtenir le titre de  
**DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE**  
*Mention : Biochimie-Microbiologie*  
**École Doctorale des Sciences de la Mer**

présentée par

**Céline Ralec**

Préparée à l'UMR 6197, Ifremer, CNRS, UBO  
Laboratoire de Microbiologie des Environnements Extrêmes  
Equipe Maintenance génomique chez les Archaea

**Thèse soutenue le 26 Novembre 2013**

devant le jury composé de :

**Evelyne SAGE**

Directrice de recherche CNRS, Laboratoire de Biologie des Radiations,  
Institut Curie, Orsay / *Rapporteur*

**Jean-Luc RAVANAT**

Directeur de recherche CEA, Laboratoire des Lésions des Acides  
Nucléiques, Grenoble / *Rapporteur*

**Gwenaëlle LE BLAY-LALIBERTE**

Maître de conférences, Université de Bretagne Occidentale / *Examinatrice*

**Mirjam CZJZEK**

Directrice de recherche CNRS, Unité de recherche 7139, Végétaux Marins  
et Biomolécules, Station Biologique de Roscoff / *Examinatrice*

**Mohamed JEBBAR**

Professeur, Université de Bretagne Occidentale / *Examineur*

**Ghislaine HENNEKE**

Chargée de recherche, Ifremer, Brest / *Examinatrice*

"ADN polymérases et acides  
nucléiques endommagés  
chez les Archaea:  
Maintenance génomique et  
Biotechnologies."









## **REMERCIEMENTS**

---

*Cette thèse a été réalisée au sein du Laboratoire de Microbiologie des Environnements Extrêmes situé à Ifremer à Plouzané avec le soutien financier de la Région Bretagne, de l'Ifremer, de l'Université de Bretagne Occidentale, du CNRS et d'Europôle Mer.*

*Je remercie tout d'abord chaleureusement Ghislaine Henneke qui a assuré la direction de cette thèse et qui m'a encadrée tout au long de ce travail et fait découvrir toute la complexité des ADN polymérases, qui m'a soutenue et encouragée pendant ces trois années.*

*Je remercie les membres du jury pour avoir accepté de juger et d'évaluer l'ensemble de ce travail de thèse : Evelyne Sage, Directrice de recherche au CNRS ; Jean-Luc Ravanat, Directeur de recherche au CEA ; Mirjam Czjzek, Directrice de recherche au CNRS ; Gwenaëlle Le Blay-Laliberté, Maître de conférences à l'UBO et Mohamed Jebbar, Professeur à l'UBO.*

*Merci à Anne Godfroy, pour son accueil au sein du laboratoire LM2E et pour ses conseils en culture des micro-organismes hyperthermophiles et à Didier Flament pour son accueil au sein de l'équipe des « Réplicants ».*

*Merci également à Jérôme Gouge et Marc Delarue pour leur accueil de quelques semaines au sein de l'Unité de Dynamique des Macromolécules de l'Institut Pasteur et pour le partage de leur passion dans l'étude des protéines.*

*Je remercie Andrei Chabes pour son accueil chaleureux au sein de son équipe de recherche russe et suédoise à l'Université d'Umeå et Lisette Marjavaara pour la transmission de son savoir faire de mesure du « pool » de nucléotide.*

*Merci à Aline Perquis, Audrey Bossé et Sébastien Laurent pour leur aide dans la réalisation des expériences et à Tom Killelea pour les discussions techniques et philosophiques sur les ADN polymérases d'Archaea.*

*Je remercie également l'ensemble des membres des laboratoires LM2E, LEP et BMM de l'Ifremer.*

*Je n'oublie pas non plus Ulysse Pereira et Laurent Meijer qui ont joué un rôle déterminant dans mon initiation à la recherche scientifique.*

*Enfin, un grand merci à mes amis, ma famille et à San pour leur infini soutien.*

*Et, la Science vous remercie !*



# VALORISATION SCIENTIFIQUE

---

Publication dans un journal scientifique international à comité de lecture

---

- **Molecular recognition of canonical and deaminated bases by *Pyrococcus abyssi* family B DNA polymerase.** Jérôme Gouge, [Céline Ralec](#), Ghislaine Henneke and Marc Delarue. J Mol Biol, 2012, 26;423(3):315-36.
- **Unexpected DNA Synthesis with Ca<sup>2+</sup> as Cofactor by Archaeal Family B DNA Polymerase.** [Céline Ralec](#), Sébastien Laurent and Ghislaine Henneke. In preparation.
- **Anomalous electrophoretic migration of short oligodeoxynucleotides labelled with 5' terminal Cy5 dyes.** Tom Killelea, Christine Saint-Pierre, [Céline Ralec](#), Didier Gasparutto and Ghislaine Henneke. Submitted.

Participation à des congrès nationaux et internationaux

---

- 05/2011 : **9<sup>ème</sup> colloque des 3R – Réplication, Réparation, Recombinaison**, 2-5 Mai 2011, Presqu'île de Giens, France.

**Poster intitulé** : « Mécanisme de reconnaissance des bases en amont de la jonction matrice/amorce : Etude de la polymérase B de *Pyrococcus abyssi*. » Jérôme Gouge, [Céline Ralec](#), Ghislaine Henneke et Marc Delarue.

- 07/2012 : **Molecular Biology of Archaea III**, 2-4 Juillet 2012, Marburg, Allemagne.

**Poster intitulé** : « Molecular recognition of canonical and deaminated bases by *Pyrococcus abyssi* family B DNA polymerase. » [Céline Ralec](#), Jérôme Gouge, Ghislaine Henneke and Marc Delarue.

- 09/2012 : **Journées des doctorants de l'Ecole Doctorale des Sciences et de la Mer**, 24-26 Octobre 2012, Plouzané, France.

**Poster intitulé** : « Molecular recognition of canonical and deaminated bases by *Pyrococcus abyssi* family B DNA polymerase. » [Céline Ralec](#), Jérôme Gouge, Ghislaine Henneke and Marc Delarue.

**Prix du 2<sup>nd</sup> meilleur poster.**

- 06/2013 : **Colloque du groupe thématique des Archaea de la SFBBM**, 20-21 Juin 2013, Plouzané, France.

**Communication orale intitulée** : « Unexpected DNA synthesis with Ca<sup>2+</sup> as a cofactor by Archaeal Family B DNA polymerase. » [Céline Ralec](#) and Ghislaine Henneke.

- 07/2013 : **Gordon Research Conference** « Archaea : Ecology, Metabolism & Molecular Biology », 28 Juillet-2 Août 2013, Barga, Italie.

**Poster intitulé** : « Molecular recognition of canonical and deaminated bases by *Pyrococcus abyssi* family B DNA polymerase. » [Céline Ralec](#), Jérôme Gouge, Ghislaine Henneke and Marc Delarue.



## TABLE DES MATIERES

LISTE DES ABBREVIATIONS	13
LISTE DES FIGURES	15
LISTE DES TABLEAUX	16
INTRODUCTION GENERALE	17
<b>CHAPITRE I : ETUDE BIBLIOGRAPHIQUE</b>	<b>19</b>
<b>Première partie : L'origine des Archaea</b>	<b>21</b>
<b>I. La découverte des océans</b>	<b>21</b>
I.1. La découverte des abysses	21
I.2. Les sources hydrothermales profondes : lieu d'un nouvel écosystème	23
I.2.1. Formation des édifices hydrothermaux profonds	23
I.2.2. Niche écologique des abysses	23
<b>II. La découverte des Archaea</b>	<b>25</b>
II.1. La place des Archaea au sein de la classification phylogénétique du Vivant	25
II.2. Ecophysiologie des Archaea	29
II.3. La cellule archéenne	31
II.4. Caractéristiques moléculaires des Archaea	31
<b>III. La découverte de <i>Pyrococcus abyssi</i></b>	<b>32</b>
III.1. Mise en culture et étude microbiologique de <i>Pyrococcus abyssi</i>	32
III.2. Biologie moléculaire de <i>Pyrococcus abyssi</i>	33
III.3. <i>Pyrococcus abyssi</i> en tant que modèle biologique	34
<b>Deuxième partie : Les ADN polymérasés au cœur de la réplication de l'ADN</b>	<b>35</b>
<b>I. La réplication de l'ADN</b>	<b>38</b>
I.1. Le cycle catalytique des ADN polymérasés réplécatives	39
I.1.1. Liaison Pol :ADN	40
I.1.2. Liaison Pol :ADN :dNTP	40
I.1.3. Formation de la liaison phosphodiester	42
I.1.4. Libération du pyrophosphate et translocation ou décrochage de la pol	42
I.2. La fonction associée de correction des erreurs d'insertion de nucléotide	43
I.2.1. Transition des modes Pol-Exo et Exo-Pol	43
I.2.2. Mécanisme catalytique réactionnel de l'activité exonucléase	44
I.3. Le contrôle indirect de la fidélité par la concentration intracellulaire des désoxyribonucléotides	44
I.4. Le cas de l'incorporation de ribonucléotides	46
<b>II. Le contrôle des dommages à l'ADN lors de sa réplication</b>	<b>47</b>
II.1. Les différents types de dommage à l'ADN	47
II.2. La reconnaissance des bases désaminées par les ADN polymérasés réplécatives archéennes	48
II.3. La tolérance aux dommages à l'ADN par les ADN polymérasés de la famille Y	49
II.4. La réparation des dommages à l'ADN par les ADN polymérasés de la famille X	51
<b>Troisième partie : Les Archaea hyperthermophiles en tant que micro-organismes modèles de l'étude de la maintenance génomique</b>	<b>55</b>
<b>I. La réplication chez les Archaea</b>	<b>55</b>
<b>II. Implication des ADN polymérasés réplécatives archéennes dans la maintenance génomique s'effectuant à haute température</b>	<b>58</b>
II.1. La famille B chez les Archaea	58
II.2. La famille D chez les Archaea	59
II.3. Les ADN polymérasés archéennes face aux dommages à l'ADN	60
<b>III. Les ADN polymérasés en interaction avec les acteurs protéiques du réplisome</b>	<b>61</b>
<b>IV. Thermostabilité des ADN polymérasés d'archées et intérêt biotechnologique</b>	<b>62</b>

<b>CHAPITRE II : Etude biochimique comparée de l'ion Mg<sup>2+</sup> et de l'ion Ca<sup>2+</sup>, en tant que cofacteurs enzymatiques, sur les propriétés fonctionnelles de l'ADN polymérase B de <i>Pyrococcus abyssi</i></b>	<b>65</b>
<b>Première partie : Présentation de l'étude</b>	<b>67</b>
<b>Deuxième partie : Article en préparation</b>	<b>69</b>
<b>Troisième partie : Résultats supplémentaires</b>	<b>91</b>
<b>I. Matériels et méthodes supplémentaires</b>	<b>91</b>
I.1 Amplification de l'ADN génomique par PCR	91
<b>II. Résultats supplémentaires</b>	<b>91</b>
II.1. Effet du Ca <sup>2+</sup> en tant que cofacteur sur les propriétés intrinsèques de PabPolB	92
II.1.1. Liaison à différent type de structures ADN	92
II.2. Comportement de PabPol B face aux dommages à l'ADN en présence de Ca <sup>2+</sup>	93
II.2.1. Tolérance de PabPolB à quelques exemples de dommages	93
II.2.2. Influence de la variation de concentration de dNTPs sur la synthèse d'ADN erronée par PabPolB	95
II.3. Comportement de PabPolB en présence de ribonucléotides (NTPs)	96
II.3.1. Incorporation unique de NTP	96
II.3.2. Extension d'amorce en présence de NTPs libres et en fonction de la longueur de l'amorce	97
II.4. Effet du calcium sur l'activité de PabPolB au cours de la PCR	99
<b>CHAPITRE III : Etude structurale et biochimique de la reconnaissance de dommage à l'ADN (base nucléotidique désaminée) par l'ADN polymérase B de <i>Pyrococcus abyssi</i></b>	<b>101</b>
<b>Première partie : Présentation de l'étude</b>	<b>103</b>
<b>Deuxième partie : Article</b>	<b>105</b>
<b>Troisième partie : Résultats supplémentaires</b>	<b>133</b>
<b>I. Matériels et méthodes supplémentaires</b>	<b>133</b>
<b>II. Résultats supplémentaires</b>	<b>134</b>
II.1. Efficacité d'extension de PabPolB triple mutée	134
II.2. Tolérance du dommage uracile et contexte nucléotidique	135
II.3. Processivité de PabPolB triple mutée	136
II.4. Etude préliminaire de la sélectivité de PabPolB triple mutée	139
<b>III. Conclusion – Perspectives</b>	<b>140</b>
<b>CHAPITRE IV : Essais préliminaires pour la mesure de la concentration intracellulaire « pool » de désoxyribonucléotides et de ribonucléotides chez l'Archaea <i>Pyrococcus abyssi</i></b>	<b>143</b>
<b>Première partie: Présentation de l'étude</b>	<b>145</b>
<b>Deuxième partie : Matériels et méthodes</b>	<b>147</b>
<b>I. Mesure du pool intracellulaire des dNTPs et NTPs chez <i>Pyrococcus abyssi</i></b>	<b>147</b>
I.1. Culture de <i>Pyrococcus abyssi</i> , archée hyperthermophile, en fiole anaérobie	147
I.2. Suivi de croissance de <i>Pyrococcus abyssi</i>	147
I.3. Dosage des dNTPs et NTPs	147
I.3.1. Collecte des cellules de <i>Pyrococcus abyssi</i>	148
I.3.1.1. Collecte des cellules par filtration	148
I.3.1.2. Collecte des cellules par centrifugation	148
I.3.2. Extraction des nucléotides totaux par le fréon	148

I.3.3. Purification des désoxyribonucléotides sur colonne d'affinité boronate	148
I.3.4. Séparation et quantification des ribonucléotides et des déoxyribonucléotides par HPLC et détection UV	149
<b>II. Mesure du pool intracellulaire de dNTPs et NTPs de <i>Pyrococcus abyssi</i> ayant subi un stress physico-chimique</b>	<b>150</b>
II.1. Culture de <i>Pyrococcus abyssi</i> , en bioréacteur « gas-lift »	150
II.2. Application de stress en cours de croissance	150
II.3. Mesure de la viabilité cellulaire de <i>Pyrococcus abyssi</i> suite au stress pH	150
<b>Troisième partie: Résultats</b>	<b>151</b>
<b>I. Mesure du pool intracellulaire de dNTPs et NTPs chez <i>Pyrococcus abyssi</i></b>	<b>151</b>
I.1. Suivi de croissance de <i>Pyrococcus abyssi</i> en fiole anaérobie	151
I.2. Mesure du pool de dNTPs et NTPs chez <i>Pyrococcus abyssi</i>	154
I.2.1. Adaptation du protocole initial	154
I.2.2. Résultats obtenus après mises au point expérimentales	154
<b>II. Mesure du pool intracellulaire de dNTPs et NTPs chez <i>Pyrococcus abyssi</i> ayant subi un stress physico-chimique</b>	<b>157</b>
II.1. Culture de <i>Pyrococcus abyssi</i> , en bioréacteur « gas-lift »	157
II.2. Influence d'un stress acide sur la viabilité cellulaire de <i>Pyrococcus abyssi</i>	158
II.3. Dosage préliminaire du pool de dNTPs et de NTPs après un stress génotoxique	159
<b>III. Conclusion-Perspectives</b>	<b>160</b>
<b>CONCLUSION GENERALE ET PERSPECTIVES</b>	<b>161</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>165</b>
<b>ANNEXES</b>	<b>187</b>



## LISTE DES ABBREVIATIONS

---

%	pourcent
°C	degré Celsius
A	alanine
ADN	<u>a</u> cide <u>d</u> ésoxyribo <u>n</u> ucléique
Å	angström
ARN	<u>a</u> cide <u>r</u> ibonucléique
ARNm	ARN <u>m</u> essenger
ARNr 16S	petite sous-unité de l' <u>a</u> cide <u>r</u> ibonucléique <u>r</u> ibosomal
ATP	<u>a</u> dénosine <u>t</u> riphosphate
BSA	<u>b</u> ovine <u>s</u> erum <u>a</u> lbumine, (albumine de sérum bovin)
CaCl <sub>2</sub>	chlorure de calcium
D	acide aspartique
dNTP	<u>d</u> ésoxyribo <u>n</u> ucléotide <u>t</u> riphosphate
DTT	<u>d</u> ithio <u>t</u> hréitol
E	acide glutamique
EDTA	<u>a</u> cide <u>e</u> thylène <u>d</u> iamine <u>t</u> étracétique
EMSA	<u>e</u> lectrophoresis <u>m</u> obility- <u>s</u> hift <u>a</u> ssay, (essai de retard de migration sur gel)
Exo	exonucléase
g	gramme
h	heure
HCl	acide chlorhydrique
IPTG	<u>i</u> sopropyl-beta-D- <u>t</u> hiogalactopyranoside
kb	<u>k</u> ilo <u>b</u> ase
KCl	chlorure de potassium
kDa	<u>k</u> ilo <u>D</u> alton
L	litre
LM2E	Laboratoire de Microbiologie des Environnements Extrêmes
M	concentration molaire
m	mètre
Mbp	<u>m</u> illion de <u>p</u> aire de <u>b</u> ases
MCM	<u>m</u> inichromosome <u>m</u> aintenance, (complexe de maintenance minichromosomique)
MgCl <sub>2</sub>	chlorure de magnésium
min	minute
MMR	<u>m</u> is <u>m</u> atch <u>r</u> epair, (réparation des mésappariements)
NaCl	chlorure de sodium
NADPH	<u>n</u> icotinamide <u>a</u> dénine <u>d</u> inucléotide <u>p</u> hosphate
NaOH	hydroxyde de sodium
NER	<u>n</u> ucleotide <u>e</u> xcision <u>r</u> epair, (réparation par excision de nucléotide)
NHEJ	<u>n</u> on- <u>h</u> omologous <u>e</u> nd <u>j</u> oining, (recombinaison non homologue)
<i>Pab</i>	<i>Pyrococcus</i> <u>a</u> bys <i>si</i>

PabPolB	ADN <u>p</u> olymérase <u>B</u> de <i>Pyrococcus <u>a</u>byss</i>
PAD	<u>p</u> olymerase <u>a</u> ssociated <u>d</u> omain
pb	<u>p</u> aire de <u>b</u> ases
PCNA	<u>p</u> roliferating <u>c</u> ell <u>n</u> uclear <u>a</u> ntigen
PCR	<u>p</u> olymerase <u>c</u> hain <u>r</u> eaction, (réaction de polymérisation en chaîne)
PDB	<u>p</u> rotein <u>d</u> ata <u>b</u> ank, (banque de données des structures de protéines)
PHP	<u>p</u> olymerase/ <u>h</u> istidinol <u>p</u> hosphatase
PM	<u>p</u> oids <u>m</u> oléculaire
Pol	ADN <u>p</u> olymérase
RFC	<u>r</u> eplication <u>f</u> actor <u>C</u>
RPA	<u>r</u> eplication <u>p</u> rotein <u>A</u>
rpm	<u>r</u> otation <u>p</u> ar <u>m</u> inute
TFB	<u>f</u> acteur de <u>t</u> ranscription <u>B</u>
TFP	<u>T</u> AATA box binding <u>f</u> actor <u>p</u> rotein
Tris	<u>t</u> rihydroxyméthylaminométhane
UDG	<u>u</u> racil <u>D</u> NA <u>g</u> lycosylase
UV	<u>u</u> ltraviolet
V	<u>v</u> olt
W	<u>w</u> att

# LISTE DES FIGURES

Figure 1. Formation et répartition géographique des sources hydrothermales profondes.	24
Figure 2. Arbre phylogénétique non enraciné des trois règnes du Vivant.	26
Figure 3. Arbre phylogénétique non enraciné des phylums au sein du règne Archaea.	27
Figure 4. <i>Pyrococcus abyssi</i> .	33
Figure 5. Schéma du cycle catalytique des ADN polymérase.	39
Figure 6. Schéma des différentes structures amorce/matrice de l'ADN et leur utilisation par l'ADN polymérase.	40
Figure 7. Evènements moléculaires de l'ADN polymérase au cours de la synthèse d'ADN.	41
Figure 8. Biosynthèse des bases nucléotidiques.	44
Figure 9. Les différentes lésions de l'ADN et les voies de réparation spécifiques associées.	47
Figure 10. Positionnement de la base uracile dans la pochette de reconnaissance des bases désaminées de la Pol B de <i>P. abyssi</i> .	48
Figure 11. Activité spécialisées des différentes voies de la réparation sur l'ADN.	51
Figure 12. Schéma montrant la similitude de la machinerie de la réplication et la différence de mécanisme de division cellulaire entre les Archaea et les Eukarya.	55
Figure 13. Les grandes étapes successives de la réplication de l'ADN et l'organisation des acteurs protéiques à la fourche de réplication chez les Archaea.	57
Figure 14. Le mécanisme de transition structurale du mode exonucléase vers le mode polymérase du complexe Pol B-PCNA-ADN chez <i>Pyrococcus furiosus</i> .	62
Figure 15. Mécanisme réactionnel de l'activité catalytique exonucléase.	67
Figure 16. Liaison de PabPolB exo- à différents types d'ADN en présence de $Mg^{2+}$ ou de $Ca^{2+}$ .	93
Figure 17. Réaction d'extension d'amorce en présence d'ADN endommagé contenant une lésion ponctuelle par PabPolB exo- avec du $Mg^{2+}$ ou du $Ca^{2+}$ .	94
Figure 18. Influence de la concentration en dNTPs sur la réaction de polymérisation par PabPolB avec $Mg^{2+}$ ou $Ca^{2+}$ en présence d'une matrice ADN saine ou contenant le dommage uracile.	95
Figure 19. Réaction d'incorporation préférentielle d'un seul ribonucléotide face à une base guanine par PabPolB exo- et PabPolB exo+.	97
Figure 20. Extension d'amorce hybridée à une matrice de 87 nucléotides PabPolB exo- en fonction de la longueur de l'amorce ADN et en présence de $Mg^{2+}$ ou de $Ca^{2+}$ comme cofacteur.	98
Figure 21. Amplification d'ADN génomique de <i>P. abyssi</i> par PCR en présence de $CaCl_2$ .	99
Figure 22. Cinétique d'extension d'amorce hybridée à une matrice ADN de 7 250 nucléotides par PabPolB sauvage et PabPolB triple mutée.	134
Figure 23. Cinétique d'extension d'amorce hybridée à une matrice de 87 nucléotides par PabPolB exo-.	135
Figure 24. Réaction d'extension d'amorce hybridée au plasmide M13mp18 de 7250 nucléotides par PabPolB exo- et PabPolB triple mutée en présence de quantités décroissantes d'héparine.	136
Figure 25. Réactions de processivité de PabPolB exo- et PabPolB triple mutée.	137-138
Figure 26. Réaction d'incorporation préférentielle d'un seul nucléotide face à une base cytosine ou une base uracile par PabPolB exo- et triple mutée.	140
Figure 27. Influence du pool de dNTPs sur la réaction d'extension d'amorce par PabPolB exo+.	146
Figure 28. Schéma montrant les différentes étapes du protocole de mesure des concentrations intracellulaires des NTPs d'une part et des dNTPs d'autre part.	149
Figure 29. Suivi de la croissance de <i>P. abyssi</i> dans trois différents milieux de culture en fioles pénicilline de 20 mL et en bioréacteur « gas-lift » de 2 L.	152
Figure 30. Suivi de croissance de <i>Pab</i> en fiole de 800 mL dans le milieu 2216 + solution C.	153
Figure 31. Concentration des dNTPs et NTPs mesurées chez <i>Pab</i> .	155
Figure 32. Calcul du ratio rNTPs/dNTPs pour <i>Pab</i> et comparaison avec ceux de l'eucaryote <i>S. cerevisiae</i> .	156
Figure 33. Courbe de suivi de croissance de <i>Pab</i> en bioréacteur « gas-lift ».	157
Figure 34. Suivi de la viabilité cellulaire de <i>Pab</i> pendant un stress acide.	158
Figure 35. Quantité de chaque NTPs dans les cellules de <i>Pab</i> après exposition à l'agent génotoxique 4-NQO.	159

## LISTE DES TABLEAUX

---

Tableau 1. Diversité d'habitat et paramètres optimaux de croissances d'une sélection d'espèces d'Archaea.	30
Tableau 2. Les différentes familles d'ADN polymérase, leur fonctions enzymatiques associées et leurs structures tridimensionnelles variées.	36-37
Tableau 3. Valeurs de chaque dNTPs, NTPs et de leurs ratios mesurés chez différents organismes du Vivant.	45

# INTRODUCTION GENERALE

---

Le maintien des espèces vivantes nécessite un équilibre étroit entre, d'une part, l'apparition de mutations au niveau du génome, qui sont essentielles à l'évolution et à l'adaptation des espèces à un environnement fluctuant et, d'autre part, la fidélité de la réplication pour une transmission fiable de l'information génétique d'une descendance à l'autre. Au sein de toutes cellules, des acteurs moléculaires participent à la préservation du patrimoine génétique et interviennent dans la maintenance génomique.

Le matériel génétique héréditaire d'une cellule est stocké dans la molécule universelle du Vivant : l'acide désoxyribonucléique (ADN). Une partie seulement de l'ADN correspond à des gènes et chaque gène porte les instructions nécessaires à la synthèse de molécules organiques par la cellule, les protéines. Ils régissent la synthèse des protéines et de ce fait assurent la régulation de la plus part des activités métaboliques cellulaires. Au cours de la division cellulaire, le transfert de l'information génétique est assuré par la copie, à l'identique, de la séquence entière du génome. La double hélice d'ADN se déroule et les deux brins servent de matrice sur lesquelles une seconde chaîne d'ADN est formée selon un principe universel qui suit le modèle décrit en 1953 par James Watson et Francis Crick. Les deux chaînes, disposées de façon anti-parallèle, sont composées chacune d'une alternance de groupements phosphates et de désoxyriboses et sont reliées entre elles par des liaisons hydrogènes impliquant les bases azotées. Les bases azotées pointent vers l'intérieur de la double hélice et s'associent par des liaisons hydrogènes selon un appariement universel, en effet, l'adénine est associée à la thymine (A:T) et la cytosine à la guanine (C:G).

Tout règne du vivant confondu, et même dans les environnements exotiques, dans une cellule donnée, une protéine enzymatique est confrontée directement dans le choix de l'assemblage des bases A, T, C et G de l'ADN selon le brin d'ADN complémentaire à répliquer ; il s'agit de l'ADN polymérase. Ce mécanisme sera vu dans les détails moléculaires dans la suite de ce manuscrit. Le Dogme actuel de la biologie moléculaire, proposé par Francis Crick est le suivant : le programme génétique stocké dans l'ADN s'exprime par sa transcription en ARN messager (ARNm), puis, l'ARNm, grâce à la participation des ARN de transfert (ARNt), est traduit en protéine (trois bases de l'ARNm dénommé codon se lient spécifiquement à un ARNt donné qui sélectionne les acides aminés un à un). La chaîne d'acides aminés constituera alors une protéine avec une structure donnée et une fonction associée.

L'ADN est une molécule sensible. En effet, sous l'action d'agents chimiques et physiques de nombreuses lésions se forment sur l'ADN. L'exploration assez récente d'environnements qualifiés d'extrêmes, montre que la vie est possible dans des milieux riches en sel, très chauds ou très froids, sans oxygène (favorable à la molécule d'ADN) et très acides, mettant en évidence de nouveaux types de voies métaboliques. Parmi les trois domaines du Vivant identifiés à ce jour, le règne des Archaea colonise de nombreux habitats exotiques. Ainsi, la question de la maintenance génomique de ces organismes se pose.



# CHAPITRE I

## ÉTUDE BIBLIOGRAPHIQUE

---

L'origine des Archaea

Les ADN polymérase au cœur de la réplication  
de l'ADN

Les Archaea hyperthermophiles en tant que  
micro-organismes modèles de l'étude de la  
maintenance génomique



# Première partie : L'origine des Archaea

---

## **I. La découverte des océans**

Au commencement, il y a 4,5 milliards d'années, la Terre était une boule fumante et incandescente entourée d'une masse nuageuse contenant de la vapeur chargée en eau. Quatre cents millions d'années plus tard la météorologie des pluies s'installe, l'eau sort massivement de son état gazeux. Il pleut alors pendant des millions d'années, un déluge s'abat à la surface du globe et entraîne son refroidissement progressif. Ce sont ces pluies diluviennes, qui sont à l'origine des océans, qui réguleront le climat et permettront l'installation de la vie il y a 4 milliards d'années. Aujourd'hui, la mer représente 71 pour cent (%) de la surface de la Terre. Le volume des eaux océaniques totalise 1 370 milliard de km<sup>3</sup> et la profondeur moyenne est estimée à 3 800 mètres (m).

Dès la plus haute Antiquité, la mer fut considérée par des philosophes tels que Thalès de Millet (VII-VI<sup>e</sup> siècle av. JC) et ses disciples comme la source de toute essence organique et comme le milieu initial pour tous les êtres vivants sur notre planète. L'exploration et la découverte des océans se sont réellement initiés au cours du Moyen-âge et à la Renaissance, cependant les organismes peuplant les océans étaient encore peu décrits et relevaient de récits mythologiques. Il faudra attendre le Siècle des Lumières pour que l'on voit les premiers naturalistes embarquer à bords des navires de commerces et effectuer quelques prélèvements de la faune et la flore maritime. Ainsi, Louis-Antoine de Bougainville (1729-1811) fut le premier grand navigateur à accueillir à bord de ses navires de conquêtes et de commerce international des savants, dont le plus célèbre d'entre eux fut le naturaliste Philibert Commerson ; il constitua d'importantes collections de poissons. Les premières collectes de faune et flore furent réalisées dans de simples bocaux et permis d'effectuer les premières descriptions de nouvelles espèces zoologies marines. Ainsi, les expéditions de Louis-Antoine de Bougainville, de James Cook, du comte de La Pérouse et de Nicolas Baudin dans le Pacifique, furent des modèles dont les grandes expéditions ultérieures, strictement scientifiques du XIX<sup>e</sup> et XX<sup>e</sup> siècles allaient s'inspirer. Des salles de laboratoire seront ainsi peu à peu implantées à bord de certains grands navires à voiles.

### **I.1. La découverte des abysses**

A la fin du XIX<sup>e</sup> siècle, l'exploration des océans entre dans une phase véritablement scientifique. Les premiers sondages en profondeur vont alors être effectués jusqu'à 1 900 m avec une pelle métallique. Les animaux vivants en profondeur furent réellement étudiés par Michael Sars qui publia en 1851 une liste d'espèces récoltées à 550 m et, simultanément, Thomas Henry Huxley décrivit les foraminifères et autres protozoaires remontés de fonds abyssaux. Des éponges, des

concombres de mer, des crustacés, des mollusques et des vers marins furent remontés de dragages effectués à 2 500 m proche du Groenland par Otto Torell en 1861. Charles Wyville Thomson interprètera l'ensemble de ces recherches comme les premières preuves de l'existence d'une vie animale en profondeur.

En 1859, Charles Darwin décrit, dans son essai sur la théorie de l'évolution *De l'origine des espèces au moyen de la sélection naturelle* (Darwin, 1859), que les espèces se modifient de façon graduelle et que les organismes qui se ressemblent sont apparentés, ils possèdent donc un ancêtre commun. Ce courant d'idée de la théorie de l'évolution va porter l'engouement de la recherche de fossiles vivants, c'est-à-dire d'organismes présentant des caractères ancestraux. L'hypothèse de les trouver dans les grands fonds marins avait alors été émise. Pour la première fois, des laboratoires scientifiques vont être installés à bord des navires. Thomson et Carpenter entament alors ce qui sera la plus grande mission océanographique au monde : la circumnavigation du H.M.S Challenger de 1872 à 1876. Le caractère pluridisciplinaire (géophysique, topographie, météorologie, zoologie, botanique) permettra d'adopter une vue globale des océans et donnera naissance à l'océanographie moderne. Une vingtaine d'années sera nécessaire pour publier l'ensemble des résultats de cette expédition sous la forme d'une encyclopédie. Ainsi, environ 5 000 nouvelles espèces furent décrites ; la plus grande profondeur fut décelée près des îles Mariannes à 8 200 m.

Par la suite, de nombreuses expéditions océanographiques seront mises en place. Au début du XIX<sup>e</sup> siècle, la création des stations marines sur le littoral équipées de laboratoires et d'aquarium donnera une toute autre importance à la biologie marine. Des expéditions spécialisées dites « campagnes océanographiques » vont progressivement se multiplier. En parallèle, les instruments de récolte depuis la surface se développent : bennes, carottiers, dragues, chalut, filet à plancton. Cependant, les organismes fixés aux roches et dissimulés dans les fissures restèrent impossibles à capturer. Il faudra attendre l'avènement des technologies modernes après la Seconde Guerre mondiale pour que l'Homme pénètre en profondeur dans l'océan. C'est en 1960, suite à l'élaboration du premier Batyscaphe nommé le Trieste en 1953 par Auguste Piccard, que Jacques Piccard descendit dans la fosse des Mariannes à un record de 10 916 m de profondeur. Arrivé sur le fond, il aperçut un poisson plat ressemblant à une sole ; cette observation mettait fin à l'hypothèse de la « zone azoïque » (qualifie des zones de la terre dépourvue d'animaux vivants). A partir de cette époque, les engins d'exploration très spécialisés se multiplièrent (sous-marin de poche, caméra des grands fonds). En 1977, les géologues américains Jack Corliss et John Edmon, embarquent à bord du submersible Alvin, par 2 500 m de profondeur, sur la ride océanique des Galápagos, à la recherche de phénomènes hydrothermaux. Au cours de cette plongée ils découvrirent comme prévu des sources hydrothermales mais également autour de ces dernières des communautés animales supérieures comportant des crevettes, des crabes, des palourdes, des vers géants, des poulpes, des poissons zoarcidés etc, constituant un écosystème nouveau pour la science. Depuis lors, ces communautés animales ont été retrouvées dans tous les océans, là où le magma entre en surface du plancher océanique et plus de 500 espèces animales ont pu être dénombrées.

## **I.2. Les sources hydrothermales profondes : lieu d'un nouvel écosystème**

Les sources hydrothermales sont l'une des caractéristiques les plus spectaculaires des fonds marins. Sous le plancher océanique, les roches en fusion sont le moteur qui entraîne la formation de sources hydrothermales. Ainsi, les événements hydrothermaux se trouvent uniquement dans les zones où le magma, généré par l'activité volcanique, est assez proche de la surface pour permettre un échauffement suffisant des liquides. La plupart des événements ont été découverts le long de la dorsale médio-océanique (Figure 1B) et également au niveau des zones de subduction. Ils peuvent se former à n'importe quelle profondeur, autant à 3 600 m qu'à 30 m et aussi au niveau de la terre, comme les geysers du parc national de Yellowstone aux Etats-Unis.

### **I.2.1. Formation des édifices hydrothermaux profonds**

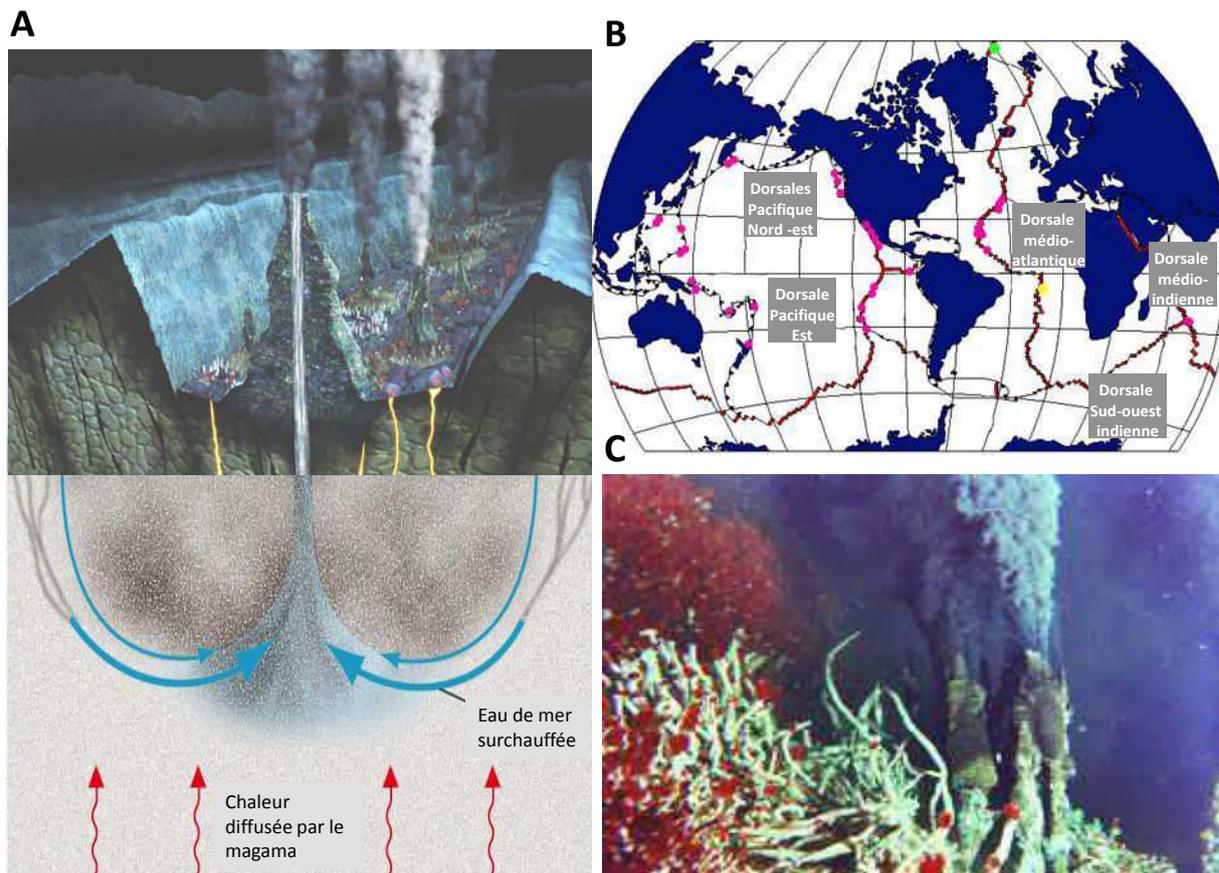
Des phénomènes géologiques et physico-chimiques sont à la base de la formation des sources hydrothermales. Voici la chronologie de leur formation (Figure 1A) : l'eau de mer froide (2°C) s'infiltré en profondeur à travers les fissures du plancher océanique, entre en contact avec le magma de roches en fusion, s'échauffe jusqu'à 350-400 °C et, par réaction chimique, se modifie au contact de la roche. Ainsi, l'oxygène est dissipé, l'acidité augmente, les métaux comme le fer, le cuivre et le zinc se dissolvent et du sulfure d'hydrogène se forme. Les liquides chauds sont moins denses et plus légers que les froids, par conséquent, le fluide hydrothermal, chargé de métaux dissous, remonte dans l'océan, jaillit par des ouvertures de ventilation et se mélange à l'eau froide. Les métaux formés dans les fluides se combinent avec le soufre et forment des sulfures minéraux noirs ce qui donne au fluide hydrothermal un aspect de fumée noire. Cette réaction est déclenchée par les températures froides et la présence d'oxygène. Dans le cas des fumeurs blancs, le mélange du fluide hydrothermal à l'eau de mer a lieu sous le plancher océanique. D'autres types de composés chimiques, comme la silice, reste dans le fluide. Lorsque le fluide sort de la cheminée, la silice précipite. De l'anhydrite, minéral blanc est également formé.

L'ensemble de ces échanges physico-chimiques participent à la création de forts gradients de température, de pH et de concentration en métaux qui s'étendent depuis la source hydrothermale vers l'eau de mer environnante. Ces écosystèmes sont en fait une source variée et riche en éléments chimiques, où la vie est rendue possible grâce à la chimiosynthèse.

### **I.2.2. Niche écologique des abysses**

En raison de l'absence de lumière au fond des océans, la photosynthèse et donc la présence de plantes y est impossible. Initialement, on pensait que les végétaux étaient à la base de toutes chaînes alimentaires connues. Il a été émis l'hypothèse que les animaux vivants sur le plancher océanique se nourrissaient des déchets organiques tombants à la surface des océans. Mais dans les écosystèmes abyssaux de ce type, l'énergie vient d'une toute autre source. En effet, les micro-organismes puisent leur énergie en décomposant les différents composés chimiques dissous dans les fluides hydrothermaux. Les composés sulfurés réduits, le méthane et l'hydrogène gazeux sont les trois sources d'énergie chimique, actuellement connues dans cet écosystème. Les bactéries tirent leur

énergie de l'oxydation de l'hydrogène sulfuré ( $H_2S$ ) ou du méthane ( $CH_4$ ), et utilisent cette énergie produite sous forme d'adénosine triphosphate (ATP) ou de nicotinamide adénine dinucléotide phosphate réduit (NADPH) pour la chimiosynthèse [incorporation du carbone du  $CO_2$  et de l'azote sous ses différentes formes (azote moléculaire, nitrates ou ammoniac) dans des molécules organiques (hydrates de carbones, acides aminés)]. Ces études ont permis de comprendre et d'identifier un nouveau type d'écosystème non fondé sur la photosynthèse. Bien que les métabolismes soient très variés, la richesse exceptionnelle (Figure 1C) des peuplements de ces zones s'explique par la présence, à la base de la chaîne alimentaire, de producteurs primaires bactériens.



**Figure 1. Formation et répartition géographique des sources hydrothermales profondes.** Elles se forment par processus géothermique (A) et sont localisées à la lisière des dorsales océaniques de la Terre (B). Photo de la faune se développant à proximité des fumeurs noirs et blancs (C).

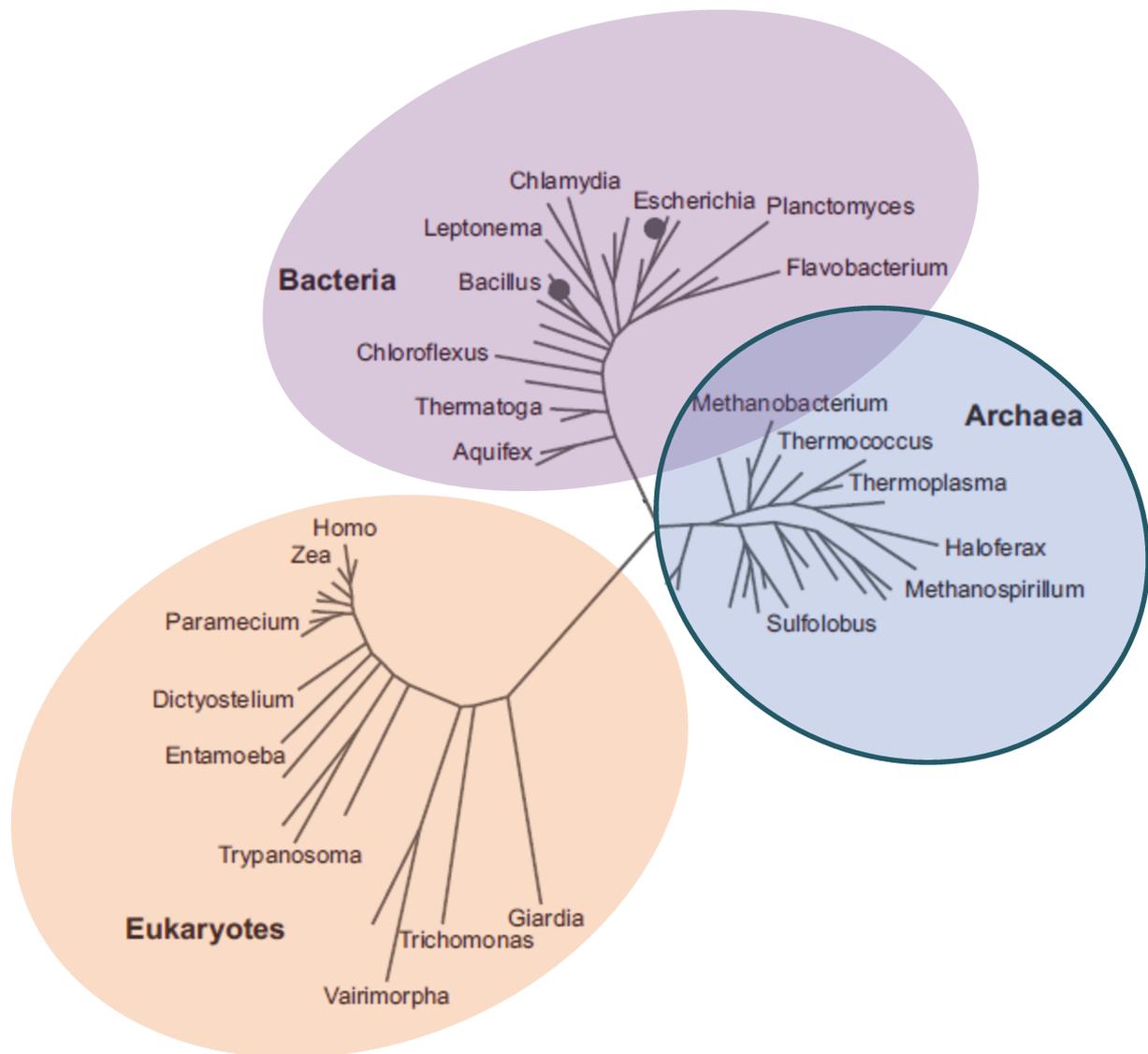
En 1981, Meridith Jones (Jones, 1981) observe dans les vers tubicoles géants *Riftia pachyptila*, la présence de sulfobactéries chimiosynthétiques vivant en étroite association (endosymbiose) au sein des organes. Les bactéries produisent la matière organique (leurs propres produits de dégradation métabolique) qui constituera la nourriture du vers. Depuis, des symbioses de ce type ont été mises en évidence dans d'autres organismes abyssaux comme les palourdes, les moules géantes, les crevettes (Zbinden and Cambon-Bonavita, 2003). A l'heure actuelle, les sulfobactéries représente la seule source de nourriture identifiée chez les organismes animaux des sources hydrothermales.

## **II. La découverte des Archaea**

Les premières traces de vie sur Terre dateraient la présence du plus ancien organisme archéen entre 3,5 et 3,9 milliards d'années. Des traces ont été trouvées sur des roches de cheminées hydrothermales profondes datant du Précambrien (Schopf and Packer, 1987 ; Brasier et al., 2002). Il y a seulement quarante ans que le premier organisme hyperthermophile a été isolé (Bott and Brock, 1969 ; Stetter et al., 1981 ; Zillig et al., 1981 ; Stetter, 1982). Plus de 1 500 souches d'Archaea ont à ce jour été isolées, notamment par le pionnier de l'hyperthermophilie Karl O. Stetter, mais la question de la survie à 121°C avec le cas de *Pyrolobus fumarii* qui résiste à l'autoclavage, est toujours sans réponse (Blöchl et al., 1997 ; Kashefi and Lovley, 2003). La résistance des Archaea à un large spectre de conditions physico-chimiques (par exemple, à la fois haute et basse températures, pH acide et pH basique, pression hydrostatique forte) pose des questions fascinantes sur l'origine de la vie. Leur métabolisme à partir d'éléments simples et qui étaient présents au début de la création du monde anoxique ainsi que leur position au sein de l'arbre phylogénétique du vivant fournit une preuve d'un ancêtre commun qui serait hyperthermophile.

### **II.1. La place des Archaea au sein de la classification phylogénétique du Vivant**

Il est possible de classer les organismes vivants selon différents critères. Au cours de l'histoire de la biologie, de nombreux critères déterminants la taxonomie des espèces ont été employés selon les outils technologiques disponibles. A l'heure actuelle, la classification communément admise utilise la biologie moléculaire et repose sur l'analyse des homologies de séquence de la petite sous-unité (16S) de l'acide ribonucléique ribosomal (ARNr 16S). Cette classification, fondée sur les divergences de la séquence de l'ARNr 16S, entre les organismes vivants, a été proposée par Woese et Fox en 1977 (Woese and Fox, 1977). Elle repose sur l'idée fondamentale que les structures et séquences moléculaires sont généralement plus révélatrices des relations d'évolution que les phénotypes classiques (morphologie, métabolisme cellulaire), et en particulier pour la classification des microorganismes (Woese, 1990). Ainsi, un nouvel arbre phylogénétique du Vivant composé de trois domaines : Bacteria, Archaea et Eukarya a été proposé (Woese and Kandler et al., 1990) (Figure 2).



**Figure 2. Arbre phylogénétique non enraciné** construit sur la base de l'analyse des séquences de l'ARNr 16S montrant les espèces représentatives des trois règnes du Vivant, adapté de Pace, 1997 et de Joyce, 2012. Une distance élevée entre deux branches témoigne d'une grande distance évolutive entre les espèces.

Cette classification est particulièrement révolutionnaire puisqu'elle discerne un nouveau règne, les Archaea, jusqu'alors considérées comme des bactéries ancestrales (archaeobactéries) et les place en tant que troisième règne du Vivant. Cette classification paraît plus juste étant donnée la divergence de composition des membranes cellulaires et des processus en charge de la réplication de l'ADN mise en évidence entre les bactéries et les archées. L'enracinement profond de l'arbre phylogénétique du vivant suggère une divergence cellulaire très ancienne. La séparation de la branche des Bacteria et de la branche que se partagent les Archaea et les Eukarya représenterait le point d'embranchement le plus ancien. Les Archaea ont une grande diversité phylogénétique. En effet, au sein même des Archaea, six phylum (ou embranchements) ont été établis selon l'analyse des séquences d'ARNr 16S ; il s'agit des phylum des Crenarchaeota, des Euryarchaeota, des Korarchaeota, des Nanoarchaeota, des Thaumarchaeota et des Aigarchaeota (Figure 3). Cependant, cette classification est encore à l'heure actuelle discutée, certains auteurs proposent le regroupement des Thaumarchaeota, Aigarchaeota, Crenarchaeota et Korarchaeota en un seul et même groupe : le « TACK » superphylum (Guy and Ettema, 2011).

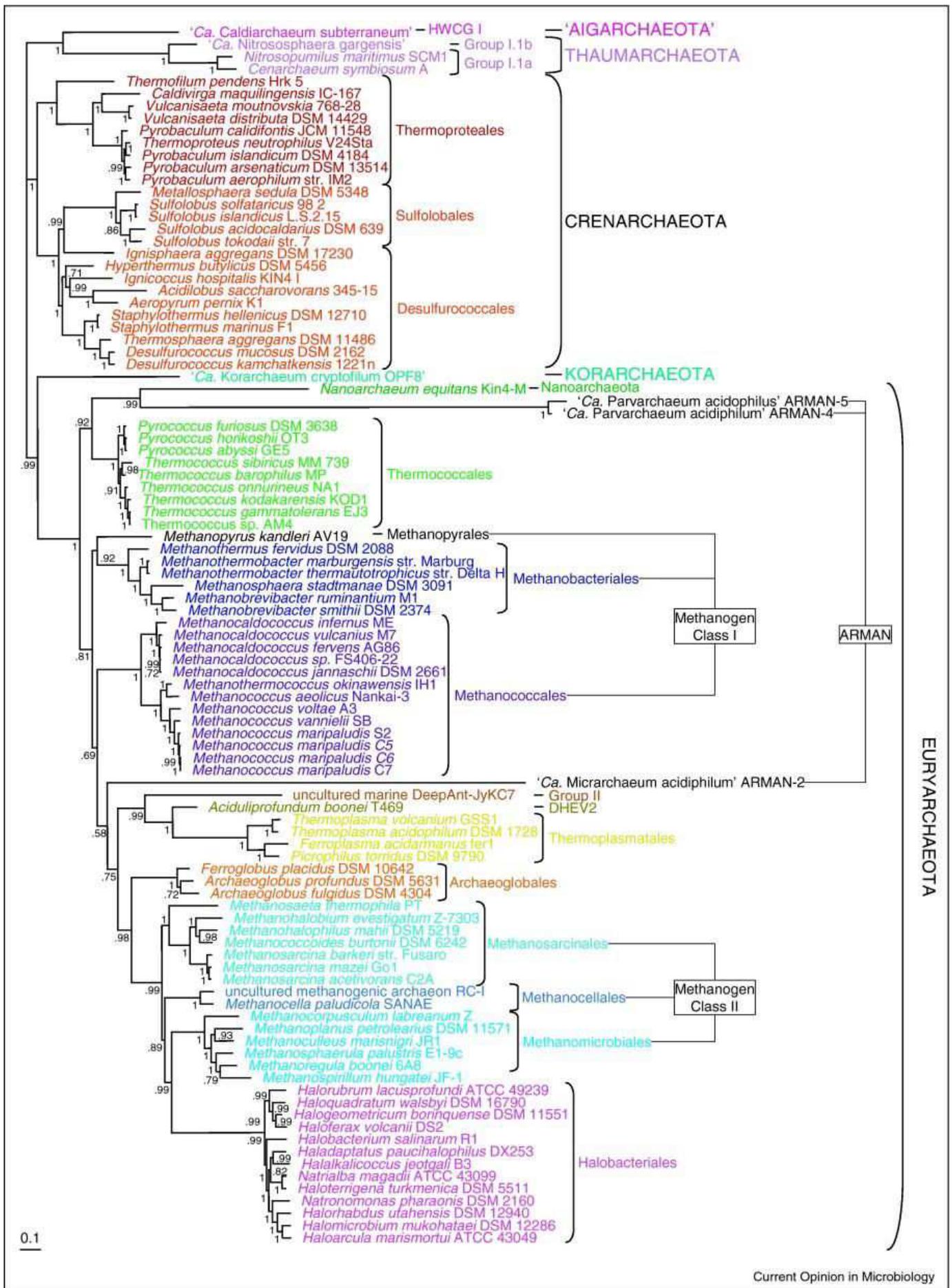


Figure 3. Arbre phylogénétique non enraciné des phylums au sein du règne Archaea, d'après Brochiet-Armanet et al., 2011.

Le phylum des Euryarchaeota rassemble un groupe d'organismes au métabolisme et à la physiologie variés, d'où leur nom « Eurus » signifiant large. Il comprend des hyperthermophiles, des thermophiles et des mésophiles, des méthanogènes, des thermococcales et des halobactériales. C'est le phylum qui compte le plus de représentants cultivés. De ce fait, certaines espèces sont utilisées comme modèles biologiques archéens pour la compréhension des mécanismes physiologiques et enzymatiques qui régissent la vie dans des contextes de milieu de vie extrême. C'est le cas, entre autre, des espèces *Haloferax volcanii*, *Thermococcus kodakaraensis* et *Pyrococcus furiosus*.

Les Crenarchaeota contiennent exclusivement des groupes de thermophiles et d'hyperthermophiles, qui métabolisent le soufre, d'où leur nom « Crene » signifiant « source », « fontaine », en grec, en référence aux sources chaudes d'où elles ont été initialement isolées. Le phylum des Korarchaeota, décrit en 1996 (Barns et al., 1996), aurait divergé très tôt au cours de l'évolution des Archaea (Elkins et al., 2008). Présentement, un seul représentant a pu être physiquement isolé d'une source chaude terrestre, il s'agit de *Candidatus Korarchaeum cryptofilum*. Son incapacité à synthétiser les cofacteurs essentiels (vitamines, purines) suggère qu'il vivrait en association symbiotique avec un organisme hôte (Nealson et al., 2008). Le phylum Nanoarchaeota représente actuellement une seule espèce (*Nanoarchaeum equitans*). La cellule, de 400 nm de diamètre, la forme de vie la plus petite décrite actuellement, de plus, son génome de 491 kb est le plus petit des génomes non viraux à avoir été séquencé (Brochier et al., 2005). Sa culture n'est rendue possible qu'en présence d'un hôte, également archéen (Crenarchaea), *Ignicoccus hospitalis* (Huber et al., 2002). Le phylum Thaumarchaeota a été proposé en 2008 (Brochier-Armanet et al., 2008) sur la base de la séquence de l'ARNr 16S et en raison de la présence d'une ADN topoisomérase de type I, enzyme jusqu'à présent considérée comme uniquement présente au sein du règne eucaryote. Aujourd'hui, seuls deux représentants des Thaumarchaeota sont cultivés : *Nitrosopumilus maritimus* et *Cenarchaeum symbiosum*, ce dernier étant un symbiote d'éponge, il n'a pas pu être cultivé en dehors de ce symbiote. Le phylum des Thaumarchaeota regroupe des organismes essentiellement mésophiles, mais aussi thermophiles, et sont très abondants dans les sédiments, le sol terrestre et les eaux marines. Sur la base de l'analyse du marqueur moléculaire de type ARNr 16S, les Thaumarchaeota apparaissent au sein des Crenarchaeota, en tant que groupe qui aurait divergé précocement. Cependant, cette vision a été revisitée grâce à l'analyse d'autres marqueurs moléculaires et de caractéristiques génomiques, rendue possible par le séquençage des génomes de *N. maritimus* et *C. symbiosum* (deux archées qui oxydent l'ammoniac). Les gènes ribosomiaux et le fait que les sous unités  $\alpha$  et  $\beta$  de l'ARN polymérase ne soient pas scindées, les distinguent des deux autres phyla d'Archaea. De plus, les Thaumarchaeota ont à la fois des caractères de Crenarchaeota (système de division cellulaire, présence de certaines nucléases), et des caractères d'Euryarchaeota (présence d'une ADN polymérase de la famille D et d'un seul homologue de PCNA) permettant de conforter l'existence de ce phylum (Spang et al., 2010). Un groupe d'organismes non cultivés fut appelé Hot Water Crenarchaeotic Group I (HWCG I) et plus récemment, le premier génome appartenant à ce groupe a été séquencé puis étudié. A partir des résultats de génomique et, notamment, de la présence de gènes de modification protéiques basés sur l'ubiquitine,

certain auteurs ont proposé l'existence d'un nouveau phylum : Aigarchaeota (Nunoura et al., 2011). Cependant ce phylum est incertain car il possède des caractéristiques typiques des Thaumarchaeota et que la distance qui le sépare des Thaumarchaeota est faible.

Aujourd'hui encore, des questions autour de l'arbre phylogénétique du Vivant se posent, notamment au sujet de la position de la racine et de la nature de l'ancêtre cellulaire commun. De plus, son mode de construction est sujet à débat du fait que selon certains auteurs, les transferts horizontaux de gènes entre les espèces perturberaient le signal phylogénétique (Doolittle and Baptiste, 2007).

## **II.2. Ecophysiologie des Archaea**

Les paramètres physico-chimiques conditionnent les habitats écologiques. En effet, la température, le pH, la pression hydrostatique et les radiations ionisantes sont des paramètres fluctuants qui limitent le fonctionnement des mécanismes biologiques des êtres vivants. Depuis la découverte des Archaea, les échantillonnages et les identifications de nouvelles souches dans les environnements dits extrêmes se sont intensifiés et ont révélé leur haut degré d'adaptation à la survie et à la multiplication cellulaire aux limites de la zone dite abiotique. La grande variabilité génétique des Archaea explique leur capacité à coloniser un large type de niches écologiques (Chaban et al., 2006), ce sont des organismes ubiquitaires. Le tableau 1 présente une liste non exhaustive d'espèces d'Archaea montre la diversité des milieux colonisés par des espèces d'Archaea ainsi que leur large et extrême spectre de conditions physico-chimiques optimales et/ou tolérées. La présence d'Archaea a aussi été constatée dans le tube digestif des animaux et de l'homme (Pochart et al., 1992) au niveau des muqueuses intestinales et orales (Dridi et al., 2011). Aucune pathogénicité associée n'a pour le moment été signalée.

Carateristiques métaboliques	T°C	NaCl (%)	pH	Pression hydro-statique	Rayonne-ments ionisants	Habitats	Organismes	Références
Hyperthermophilie	106	1,7	5,5	+	+	Source hydrothermale profonde (- 3 650 m)	<i>Pyrolobus fumarii</i>	Blöchl <i>et al.</i> , 1997
Hyperthermophilie Acidophilie	80		2-3			Fumeroles soufrés	<i>Sulfolobus solfataricus</i>	Brock <i>et al.</i> , 1972
Thermophilie Acidophilie	60		0,6			Solfatares arides	<i>Picrophilus torridus</i>	Schleper <i>et al.</i> , 1995
Acidophilie Ferro-oxydation	40		1- 1,7			Mine de fer	<i>Ferroplasma acidarmanus</i>	Dopson <i>et al.</i> , 2004
Halophilie	42	20-25	7			Mer Morte	<i>Haloferax volcanii</i>	Torreblanca <i>et al.</i> , 1986
Alkalophilie	85		9			Vulcano	<i>Thermococcus acidamonivorans</i>	Dirmeier <i>et al.</i> , 1998
	70		6,8- 7,5	+	+	Réservoir géologique de pétrole (- 1 700 m, - 2 500 m)	<i>Thermotogales</i>	Slobodkin <i>et al.</i> , 1999
Psychrophilie	-2,5 +29		7- 7,8			Lac de l'Antarctique	<i>Methanococcoïdes burtonii</i>	Franzmann <i>et al.</i> , 1992
Hyperthermophile	88		6,5	+	20 KGy	Source hydrothermale profonde (- 2 600 m, - 3 650 m)	<i>Thermococcus marinus</i>	Jolivet <i>et al.</i> , 2004
Hyperthermophile Piezophilie	98	3,5	7,5- 8	52 MPa	+	Source hydrothermale profonde (- 4 100 m)	<i>Pyrococcus CH1</i>	Zeng <i>et al.</i> , 2009

**Tableau 1. Diversité d'habitat et paramètres optimaux de croissance d'une sélection d'espèces d'Archaea** et mise en valeurs des paramètres optimaux peu conventionnels. Seules les valeurs des paramètres physico-chimiques caractéristiques sont indiquées pour chaque organisme. Les valeurs correspondent aux optima de croissance définis lors de la culture en laboratoire.

Les Archaea se composent d'une abondante variété d'espèces et représenteraient environ 20 % de la biomasse sur Terre (DeLong and Pace, 2001). Leur abondance et leur distribution dans les océans et dans les sédiments semblent contribuer aux cycles biogéochimiques énergétiques de l'azote, du soufre et du carbone de la planète (Wuchter et al., 2006 ; Madsen et al., 2011 ; Offre et al., 2013). Les Archaea seraient donc des acteurs importants au sein de la biosphère. L'exploration récente du microbiome humain (Grice and Segre, 2012) a mis en évidence que toutes les communautés écologiques de la flore commensale, symbiotique et pathogène seraient pour l'Homme une source de diversité génétique, de modulateur de pathologies, et influenceraient le métabolisme donc l'activité des traitements médicamenteux. Tout récemment, les Archaea ont été retrouvées sur la peau humaine, elles représenteraient 4 % du microbiome de la peau. D'après la signature de l'ARNr 16S, il s'agirait de Thaumarchaeota (Probst, 2013). On comprend donc toute l'importance de l'étude de ces organismes, d'un point de vue à la fois environnemental et humain.

### II.3. La cellule archéenne

Les Archaea sont des organismes unicellulaires. Leur matériel génétique est libre dans la cellule, en ce sens, il n'est pas compartimenté dans un noyau. Les Archaea sont, de ce fait, tout comme les bactéries, des procaryotes. Leur taille varie de 0,6  $\mu\text{m}$  à 8  $\mu\text{m}$  et leur forme est diverse. Les formes les plus communes des organismes cultivés sont des formes de coques ou de bâtonnets, tout comme les bactéries. Certaines espèces d'Archaea se distinguent par des formes qui leur sont propres, telles que les branches de Thermoproteus (Zillig et al., 1981) et les polygones de quelques halophiles (Bolhuis et al., 2004).

La structure de leur paroi est fondamentalement différente de celles des bactéries (Albers and Meyer, 2011). En effet, à la différence des bactéries, elles sont dépourvues de peptidoglycanes mais comportent une couche de surface paracrystalline composée de protéines et d'acide N-acetylmuramique nommée S-layer, elle forme la membrane interne des Archaea (Kandler and König, 1998 ; Lombard et al., 2012). Une autre caractéristique différente est la composition chimique de la membrane. Alors que les membranes des bactéries et des eucaryotes sont composées de lipides glycol-ester, celles des Archaea sont composées de lipides glycol-ether et cette différence jouerait un rôle primordial dans l'adaptation des Archaea à l'hyperthermophilie et à l'acidité (Boyd et al., 2013). Les Archaea possèdent des structures flagellaires qui les rendent mobiles (Albers and Pohlschröder, 2009) ainsi que tout un ensemble de protéines qui s'organisent en un complexe structural et fonctionnel dénommé archaellum (Jarrell and Albers, 2012). Les Archaea peuvent former des biofilms mono-espèces, ou avec d'autres communautés, par la sécrétion d'exopolysaccharides en réponse à des stress métaboliques. Cela leur permet, entre autre, de s'adapter aux changements de conditions physico-chimiques de leur environnement (Fröls, 2013).

### II.4. Caractéristiques moléculaires des Archaea

Etant donnée la place des Archaea au sein de l'arbre phylogénétique du Vivant, il existe de grandes similitudes, d'un point de vue génétique, entre les Archaea et les eucaryotes d'une part, et les Archaea et les procaryotes d'autre part. Les protéines impliquées dans les processus d'énergie et de la communication cellulaire semblent être distincts dans les trois règnes du Vivant (Andrade et al., 1999). Les principaux facteurs de transcription tel que le TATA box binding factor (TFP) et le facteur de transcription B (TFB) ont été caractérisés chez les Archaea (Bartlett, 2005). Ainsi, le type de mécanisme de régulation de la transcription des Archaea se rapprocherait de celui adopté par les eucaryotes (Bell and Jackson 2001 ; Bell, 2005). Le séquençage complet de plusieurs génomes archéens a permis de mettre en évidence une similitude d'homologie de séquence avec les protéines de la réplication présentes chez les eucaryotes. Tout comme pour les eucaryotes, le promoteur des Archaea contient l'élément TATA, localisé approximativement à 26 nucléotides en amont du site d'initiation de la transcription (Reiter et al., 1990). Les processus informationnels que sont la réplication, la transcription et la traduction sont assurés par des homologues de protéines eucaryotes tandis que les processus dits opérationnels (métabolisme cellulaire) impliquent des homologues de protéines bactériens (Olsen and Woese, 1996 ; Kelman and White 2005 ; Makarova et al., 2010).

La cohabitation parfois étroite des Archaea avec les eucaryotes et avec les bactéries a permis l'établissement d'échange de matériel génétique par transfert horizontal (Jain et al., 2002). Notamment, des mécanismes d'échanges d'ADN entre les bactéries et les archées hyperthermophiles ont été découverts et joueraient un rôle dans la survie et l'adaptation de la vie à haute température (van Wolferen, 2013). Au cours d'une étude, des échanges de gènes ont pu être induits par les rayonnements ultra-violet (UV) et l'analyse a montré que les pili permettaient de véhiculer les gènes (Ajon et al., 2011). Récemment, le séquençage de cinq plasmides nouvellement identifiés chez plusieurs souches de *Thermococcus* révèle qu'un échange d'éléments génétiques est possible au niveau des sources hydrothermales. En effet, les plasmides, ont été attribués à deux familles, pTN2-like et pEXT9a-like et des homologies de séquences entre le plasmide pTN2-like et le virus PAV1 de *Pyrococcus abyssi* (Geslin et al., 2003) ont été mises en évidence (Krupovic et al., 2013).

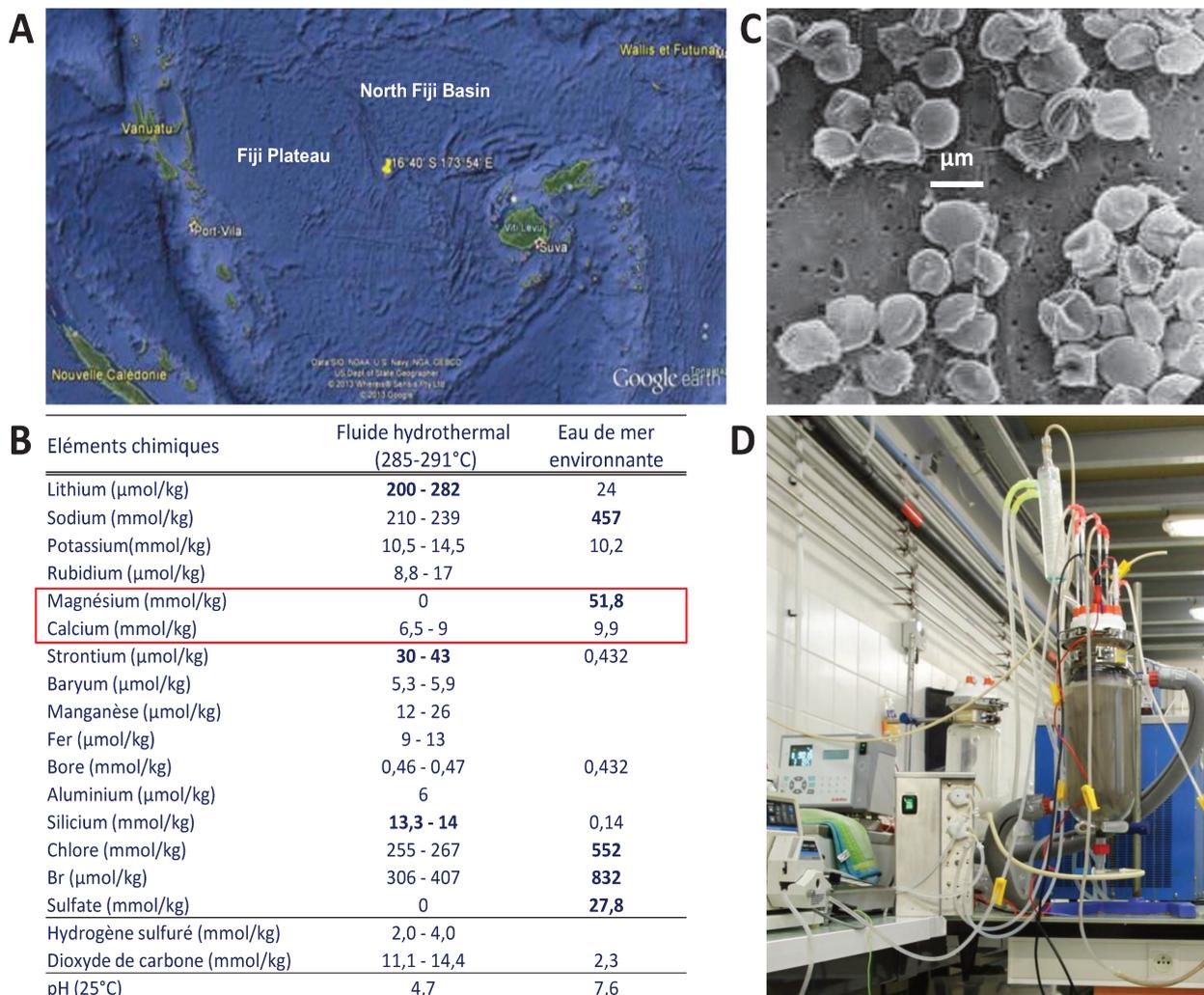
Pour conclure, trois caractères principaux distinguent les Archaea des bactéries et des eucaryotes : la stéréochimie du glycérol phosphate impliqué dans les phospholipides qui composent la membrane cellulaire des Archaea (Boucher et al., 2007) ; le caractère non pathogène des Archaea (à ce jour non démontré) (Cavicchioli et al., 2003), la méthanogénèse (Ferryard and Kestead, 2007) et l'oxydation anaérobie du méthane (DeLong, 2000 ; Thauer and Shima, 2006).

### **III. La découverte de *Pyrococcus abyssi***

*Pyrococcus abyssi* (*P. abyssi*) est une Euryarchaea affiliée à l'ordre des Thermococcales. La souche GE5 a été échantillonnée lors de la campagne à la mer Starmer (juin 1989-juillet 1989), dont la mission, était destinée à effectuer la description des communautés d'organismes associés aux sources hydrothermales dans l'océan Pacifique, au nord du bassin fidjien dans l'eau de mer environnant une source hydrothermale à environ 2 000 mètres de profondeur (Figure 4A) (Auzende et al., 1991). Une analyse de la composition en éléments chimiques de l'eau émise par les fumeurs noirs et de l'eau à proximité des sources a été effectuée (Figure 4B) (Ishibashi et al., 1994).

#### **III.1. Mise en culture et étude microbiologique de *Pyrococcus abyssi***

*P. abyssi* a été cultivée et caractérisée (Erauso et al., 1993). Les cellules sont Gram négatives et sont en forme de coques de 0,8 à 2 µm, à membrane irrégulière (Figure 4C), des flagelles lui confère une certaine mobilité. La souche est une Archaea hyperthermophile qui croit entre 68 et 101,5 °C et montre une tendance à la barophilie (Marteinsson et al., 1997). C'est un organisme au métabolisme anaérobie strict. Hétérotrophe, elle utilise comme source de carbone principale des peptides et des acides aminés. Elle puise son énergie dans la réduction du soufre inorganique et produit de l'hydrogène sulfuré ainsi que de l'acétate, de l'isovalérate, du propionate, de l'hydrogène et du dioxyde de carbone. Au laboratoire, la culture continue en fermenteur « gas-lift », dans les conditions physico-chimiques optimales de croissance définies de la souche *P. abyssi* a été mise au point (Figure 4D) (Godfroy et al., 2000).



**Figure 4.** *Pyrococcus abyssi* a été échantillonné dans l'océan Pacifique au niveau du bassin fidjien (A), l'eau de mer issue du fluide hydrothermal est enrichie en métaux (tableau adapté de Ishibashi et al., 1994) (B). (C) Observation de la forme et de la taille des cellules de *P. abyssi* au microscope électronique à balayage, (photo Ifremer). (D) Culture de *P. abyssi* en bioréacteur « gas-lift » en anaérobiose et à 96°C au laboratoire LM2E (photo Ifremer C. Ralec).

### III.2. Biologie moléculaire de *Pyrococcus abyssi*

Le génome de *P. abyssi* (souche GE5) a été entièrement séquencé en 1998 au génoscope d'Evry (France) et son annotation complète, en 2003, a permis d'identifier de nombreux acteurs moléculaires régissant son fonctionnement (Cohen et al., 2003). Ainsi, la séquence génomique comprend un chromosome de 1,7 million de bases (Mb), riche en GC de 44,7% et un plasmide pGT5 de 3,4 kilobases (kb) qui a pu être modifié afin de constituer un outil de génétique *in vivo* chez *P. abyssi* (Erauso et al., 1996 ; Lucas et al., 2002). La séquence comporte 1 764 cadres de lecture ouverts codants pour des protéines dont environ 30% à ce jour n'ont pas de fonction identifiée. Récemment, une nouvelle annotation du génome a permis d'identifier la fonction de 200 nouvelles protéines (Gao and Wang, 2012).

### III.3. *Pyrococcus abyssi* en tant que modèle biologique

Les Thermococcales sont un des ordres des Euryarchaea particulièrement étudié dans des problématiques abordant l'hyperthermophilie et la piézophilie. En effet, chez les organismes hyperthermophiles qui survient et se multiplie à plus de 100°C la question palpitante de la résistance des protéines et de l'ADN à la dénaturation et des mutations à l'ADN se pose. De façon intéressante, des études ont montré que la gyrase inverse est une protéine spécifique des organismes hyperthermophiles (thèse Isabelle Goulet 1985). Cette protéine a la capacité d'augmenter la torsion de la double-hélice d'ADN, ce qui rend plus difficile la séparation des deux brins à haute température (Bouthier de la Tour et al., 1990 ; Bouthier de la Tour et al., 1991). De plus, la mise en évidence de plasmides relâchés et surenroulés positivement chez les Archaea hyperthermophiles pourrait expliquer la résistance de l'ADN à la chaleur (Forterre et al., 1992, Charbonnier et al., 1992 ; Charbonnier et al., 1994 ; Lopez-Garcia et al., 1997). D'autant plus que chez les mésophiles, les plasmides d'Archaea sont surenroulés négativement (Charbonnier et al., 1994).

L'étude directe de la stabilité de l'ADN à très haute température a notamment été abordée par Patrick Forterre (Forterre et al., 1992 ; Marguet and Forterre, 1994). Ils ont observé que l'ADN topologiquement fermé ne se dénaturait pas sous l'effet de la chaleur au moins jusqu'à 107 °C. Et que les sels (MgCl<sub>2</sub>, KCl) ralentissaient considérablement la thermodégradation de l'ADN et protégeaient l'ADN de la dépurination. Ainsi sur le plan physiologique, compte tenu de la concentration intracellulaire très élevée en potassium chez nombre d'entre eux, les sels stabiliseraient les acides nucléiques chez les organismes hyperthermophiles (Mojica et al., 1994).

Les Archaea constituent des modèles d'étude intéressants d'autant plus que leur mise en culture au laboratoire a fait l'objet de nombreuses optimisations expérimentales et leur croissance en est ainsi aisée. Nous verrons dans la suite de cette étude bibliographique que les protéines impliquées dans la maintenance génomique des Archaea sont de type eucaryote. De plus, suite au séquençage du génome de *P. abyssi*, de nombreuses protéines ont pu être purifiées au laboratoire et, notamment, des protéines impliquées dans la réplication et la réparation de l'ADN. Nous avons vu précédemment que les Archaea isolées des sources hydrothermales profondes sont soumises à des conditions physico-chimiques drastiques qui peuvent porter atteinte à leur génome. Pour toutes ces raisons, *P. abyssi* constitue un modèle biologique d'intérêt dans la compréhension des mécanismes biologiques ayant un rôle dans la maintenance génomique.

## Deuxième partie : Les ADN polymérase au cœur de la réplication de l'ADN et de la maintenance génomique

---

En 1959, Arthur Kornberg reçut le Prix Nobel pour ses travaux sur le principe de duplication de l'ADN par les ADN polymérase (pols). Par la suite, il a été confirmé dans chacun des trois règnes du Vivant que les pols utilisent une matrice ADN simple brin pour synthétiser un brin complémentaire, et qu'en revanche elles n'ont pas la capacité de synthétiser l'ADN *de novo*, c'est-à-dire en l'absence de matrice. En effet, les pols dépendent d'une autre protéine appelée primase qui est responsable de l'assemblage de courts fragments d'ARN composés d'extrémités 3'-OH et pouvant être étendue par une ADN polymérase répliquative. A la fourche de réplication, l'ADN polymérase agit de façon coordonnée avec plusieurs partenaires protéiques pour mener à bien la réplication de l'ADN. Ces facteurs intervenant en amont de la réplication proprement-dite, vont rendre l'ADN à répliquer accessible à l'ADN polymérase et faciliter son activité, on parle de réplisome. Les ADN polymérase sont présentes dans toutes les cellules et dans tous les règnes du Vivant y compris au niveau des éléments génétiques mobiles (virus, plasmides) et des organites cellulaires (mitochondries, chloroplastes). Il s'agit donc très probablement de l'une des premières machineries enzymatiques à être apparue sur Terre conjointement à la molécule d'ADN. Depuis la première purification à partir d'un extrait cellulaire d'*Escherichia coli* d'une enzyme capable d'incorporer des désoxyribonucléotides à l'ADN par Kornberg et ses collaborateurs en 1956 (Kornberg et al., 1956), de nombreuses ADN polymérase ont été découvertes dans chaque règne du Vivant et caractérisées biochimiquement. De plus, de multiples structures cristallographiques des pols seules ou en complexes avec l'ADN et des désoxyribonucléotides ont été résolues. Elles ont permis d'observer au niveau atomique [(de l'ordre de l'angström (Å)] la préhension des pols à leurs substrats et notamment les interactions mises en jeu entre les acides aminés et les substrats. Ainsi, sur la base de la séquence primaire des pols et de leur structure tridimensionnelle, un classement en famille a été établi. A ce jour sept familles de pols ont pu être identifiées, il s'agit des familles A, B, C, D, X, Y, et RT (transcriptase inverse) (Tableau 2).

Familles	Règnes	ADN polymérases	Fonctions principales	Fonctions additionnelles	
A	Bactériophage	T3 T5 T7	Réplication Réplication Réplication		
	Bacteria	Pol I	Maturation des fragments d'Okazaki NER	5'→3' exonucléase 3'→5' exonucléase	
	Eukarya	Pol γ	Maintenance de la mitochondrie	3'→5' exonucléase	
		Pol θ Pol ν	Translésion Translésion	dRP lyase dRP lyase	
B	Bactériophage /virus	HSV	Réplication		
		RB69 T4 T6	Réplication Réplication Réplication		
	Bacteria	Pol II	Translésion	3'→5' exonucléase	
	Archaea	Crenarchaea	Pol B1 Pol B2 Pol B3	Réplication Réplication Réplication	
		Euryarchaea	Pol B1	Réplication	3'→5' exonucléase
	Eukarya		Pol ε Pol δ	Réplication du brin avancé Réplication du brin retardé NER BER	3'→5' exonucléase 3'→5' exonucléase
			Pol α Pol ζ	Amorçage de brin Translésion	Synthèse ARN
C	Bacteria	Pol III	Réplication	3'→5' exonucléase (SU ε)	
D	Archaea	Euryarchaea	Pol D	Réplication	3'→5' exonucléase
X	Virus	ASFV X	Réplication		
		Pol β	BER Translésion Réparation cassures double brin	dRP/AP lyase terminale transférase	
	Eukarya	Pol λ	NHEJ BER Recombinaison V(D)J	dRP lyase terminale transférase polynucléotide synthase terminale transférase	
		Pol μ	NHEJ Recombinaison V(D)J		
		Pol σ	Cohésion des chromatides soeurs		
Bacteria	TdT Pol X	Recombinaison V(D)J Gap filling			
Y	Bacteria	Pol IV Pol V	Translésion Translésion		
	Archaea	Crenarchaea	Dpo4/Dbh	Translésion	
		Eukarya	Pol η Pol ι	Translésion Translésion	
		Pol κ	BER Translésion		
		Pol REV1	NER Translésion		
RT	Virus	Reverse transcriptase	Translésion	dCTP transférase	
	Eukarya	Téломérase	Réplication des extrémités des chromosomes		

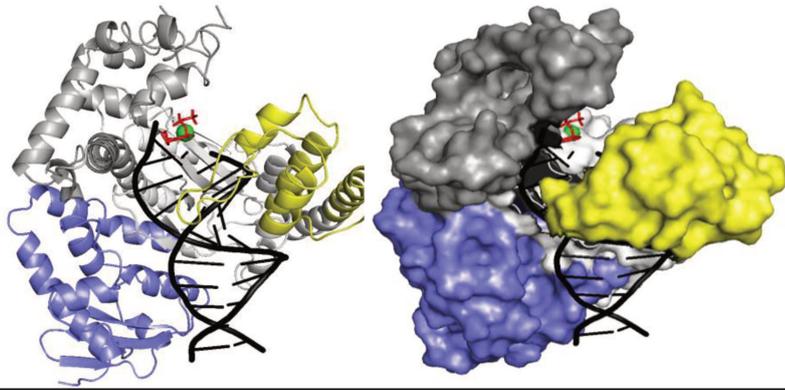
Tableau 2. Les différentes familles d'ADN polymérases, leur fonctions enzymatiques associées et leurs structures tridimensionnelles variées (ci-après).

Fragment de Klenow d'*E. coli*

en triplex avec ADN 12/13 et dCTP  
1 ion  $Mg^{2+}$  dans le site polymérase  
Résolution : 2,3 Å

publié par Li et al., 1998  
PDB 2KTQ

**FAMILLE A**



D  
O  
M  
A  
I  
N  
E  
S

Doigt

Paume

Pouce

N-terminal

Exonucléase

PHP

$\beta$ -binding

Petit doigt

dRP 8 kDa

ADN

$Mg^{2+}$

$Ca^{2+}$

Autres ions

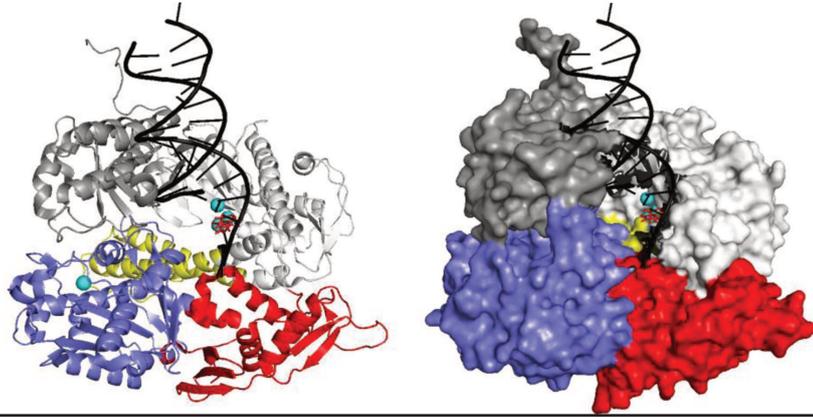
Nucléotide entrant

Pol  $\alpha$  du phage RB69

en triplex avec ADN 14/18 et TTP  
2 ions  $Ca^{2+}$  dans le site polymérase  
1 ion  $Ca^{2+}$  dans le site exonucléase  
Résolution : 2,6 Å

publié par Franklin et al., 2001  
PDB 1IG9

**FAMILLE B**

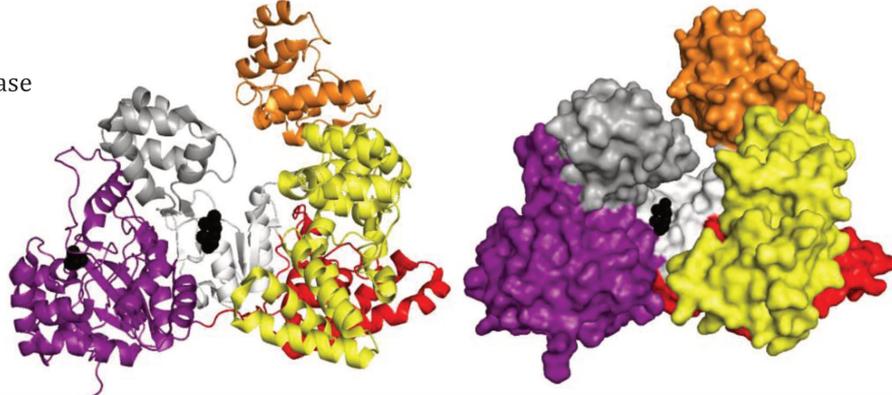


Pol III (sous-unité  $\alpha$ ) d'*E. coli*

2 ions  $PO_4^{2-}$  dans le site polymérase  
Résolution : 2,3 Å

publié par Lamers et al., 2006  
PDB 2HNH

**FAMILLE C**



Pol D (sous-unité DP2) de *P. horikoshii*

Résolution : 2,2 Å

publié par Matsui et al., 2011  
PDB 3O59

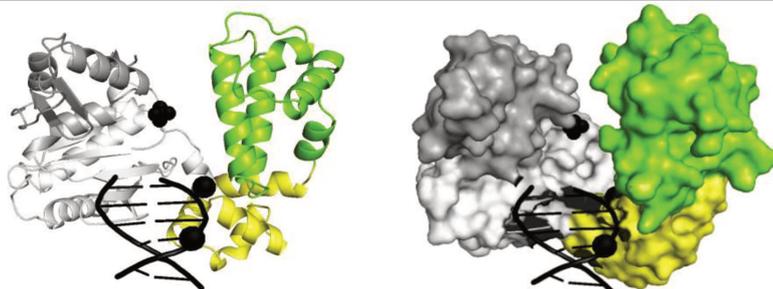
**FAMILLE D**



Pol  $\beta$  humaine

en duplex avec ADN 6/7  
Résolution : 2,6 Å  
publié par Pelletier et al., 1996  
PDB 9ICW

**FAMILLE X**

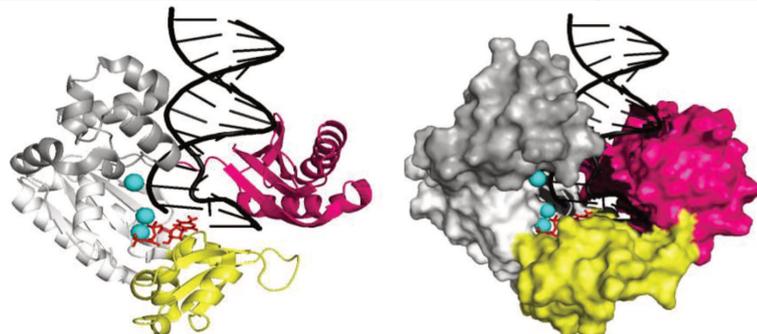


Pol Dpo4 de *S. solfataricus*

en triplex avec ADN 13/19 et dCTP  
3 ions  $Ca^{2+}$  dans le site polymérase  
Résolution : 2,0 Å

publié par Wu et al., 2011  
PDB 3QZ7

**FAMILLE Y**



La représentation des familles de pol au sein des règnes du Vivant est hétérogène. En effet, jusqu'à présent, trois types d'ADN polymérases ont été identifiées chez les Archaea, cinq chez la bactérie *E. coli*, neuf chez l'eucaryote unicellulaire *S. cerevisiae* et seize chez l'Homme.

Nous verrons, dans la suite de ce chapitre, que cette classification des ADN polymérases établie sur la structure est en lien avec leur fonction principale et leur rôle soit dans la réplication de l'ADN soit dans la prise en charge des lésions de l'ADN. Les pols ont une structure tridimensionnelle diversifiée. Certains domaines structuraux sont spécifiques à une famille (Tableau 2), il s'agit du domaine N-terminal (rouge), du domaine exonucléase (bleu), du domaine polymérase/histidinol phosphatase (PHP) (orangé), du domaine  $\beta$ -binding (violet), du domaine « petit-doigt » (rose) et du domaine dRP (vert). Ces domaines confèrent des propriétés catalytiques intrinsèques bien particulières aux pols, ils modulent par exemple la sélectivité au substrat et la fidélité intrinsèque de la pol. Cependant, la structure du cœur catalytique des pols constituée des domaines fonctionnels de même que la réaction d'incorporation de nucléotide à l'extrémité du 3'-OH du brin néosynthétisé est conservé chez toutes les familles de pol. Les ADN polymérases partagent une base structurale identique. Celle-ci se réfère à une main droite, où, trois domaines, communément appelés « doigt », « paume » et « pouce », agissent de concert dans le positionnement de la double hélice au sein du site catalytique et dans la réaction d'insertion de nucléotides (Tableau 2, respectivement, domaines jaune, blanc et gris).

La partie qui va suivre consiste à exposer le déroulement du mécanisme de la réplication par les pols, d'un point de vu moléculaire, dans le but de mieux comprendre comment les événements au sein des pols participent à la haute fidélité de la réplication de l'ADN. Nous verrons que le contrôle de la fidélité est effectué à plusieurs niveaux lors de la réplication de l'ADN. En effet, ces contrôles s'opèrent lors de la sélection des désoxyribonucléotides triphosphates (dNTPs) libres à insérer par la pol (fidélité intrinsèque de la pol), de l'extension de l'ADN (activité de relecture par le domaine exonucléase de la pol), de la modulation de la concentration intracellulaire du pool de dNTP et aussi lors de l'intervention d'enzymes spécialisées dans la réparation de l'ADN (contrôle *a posteriori*).

## **I. La réplication de l'ADN (incorporation de base canonique)**

La polymérisation de l'ADN s'effectue selon deux grands principes. Premièrement, elle consiste en la formation d'une liaison phosphodiester entre le groupement 5'-phosphate d'un nucléotide libre et le groupement 3'-hydroxyle de l'extrémité du brin d'ADN néosynthétisé (le brin amorce). Et deuxièmement, chaque désoxyribonucléotide ajouté au niveau de l'amorce est sélectionné par appariement de base complémentaire selon l'instruction du brin matrice. Les pols effectuent la synthèse d'un brin d'ADN amorce hybridé à une matrice, dans le sens 5'→3' ; ainsi,  $ADN_n + dNTP \rightarrow ADN_{(n+1)} + Ppi$ .

L'insertion des nucléotides un à un le long du brin amorcé est un phénomène très rapide, par exemple, l'ADN polymérase B répliquative de l'Archaea *P. abyssi* insert 255 nucléotides par secondes (Dietrich et al., 2002). Cependant, les techniques de cristallisation des protéines (Henzler-Wildman

and Kern, 2007) et la résolution des structures apo (pol seule), binaire (pol :ADN), et ternaire (pol :ADN :dNTP), associées aux études enzymatiques ont permis de mieux comprendre le phénomène moléculaire catalytique et de le découpler. Ce chapitre traitera de l'activité des pols d'un point de vu mécanistique moléculaire et insistera sur le contrôle de la fidélité de la réplication. L'une des propriétés principales des pols répliquatives est la fidélité. Nous verrons que la taille du site actif, la flexibilité du site actif et l'accommodation de l'ADN contribuent à la fidélité intrinsèque de l'enzyme.

### I.1. Le cycle catalytique des ADN polymérases répliquatives (voir schéma général en Figure 5)

Le cycle catalytique des pols répliquatives consiste en une succession d'étapes réactionnelles qui sont liées à un changement de conformation de l'enzyme. De façon séquentielle, l'ADN polymérase lie une structure particulière d'ADN (1), puis sélectionne le nucléotide approprié complémentaire à la base instructrice du brin matrice (2). Ensuite, la réaction de catalyse permet d'effectuer la synthèse de la liaison phosphodiester (3) et (4). Enfin, un pyrophosphate est relargué (5), puis, soit un nouveau cycle d'extension se met en place, la synthèse est alors dite processive (6), soit la pol se décroche de l'ADN, la synthèse est alors dite distributive (7).

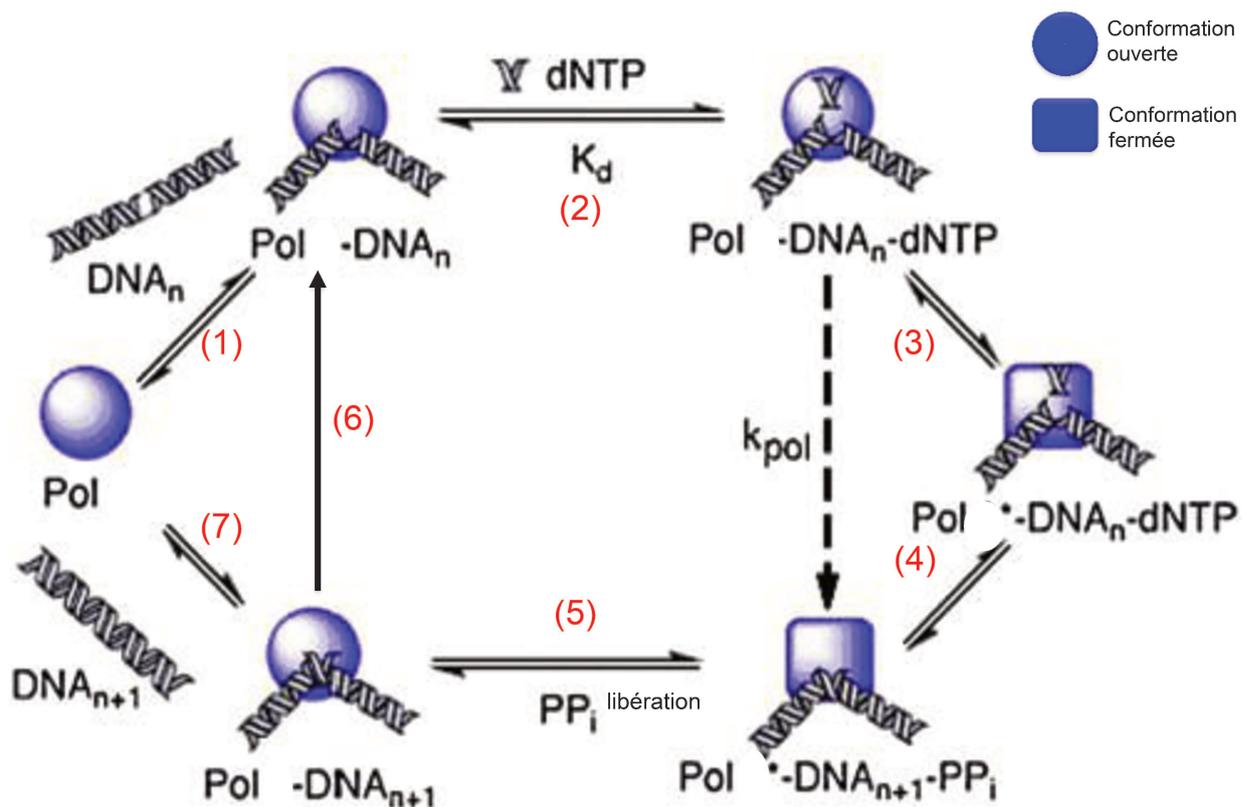


Figure 5. Schéma du cycle catalytique des ADN polymérases (en bleu), adapté de Yamtich and Sweasy, 2010.

**I.1.1. Liaison Pol :ADN [étape (1)]**

L'ADN peut adopter différents intermédiaires structuraux. Cinq catégories principales existent (Figure 6) ; les pols sont capables de les distinguer les une des autres selon le contexte physiologique. L'ADN polymérase se lie à l'ADN de façon à positionner le brin matrice et l'extrémité 3'-OH du brin à étendre (aussi appelée amorce) dans le site catalytique [Figure 5 (1)]. La capacité de l'ADN polymérase à former ce complexe binaire (Pol :ADN) se caractérise par son affinité. Ainsi, l'information sur le nucléotide à insérer est assurée.

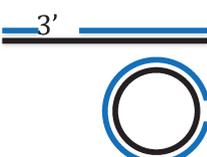
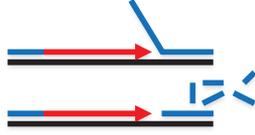
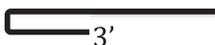
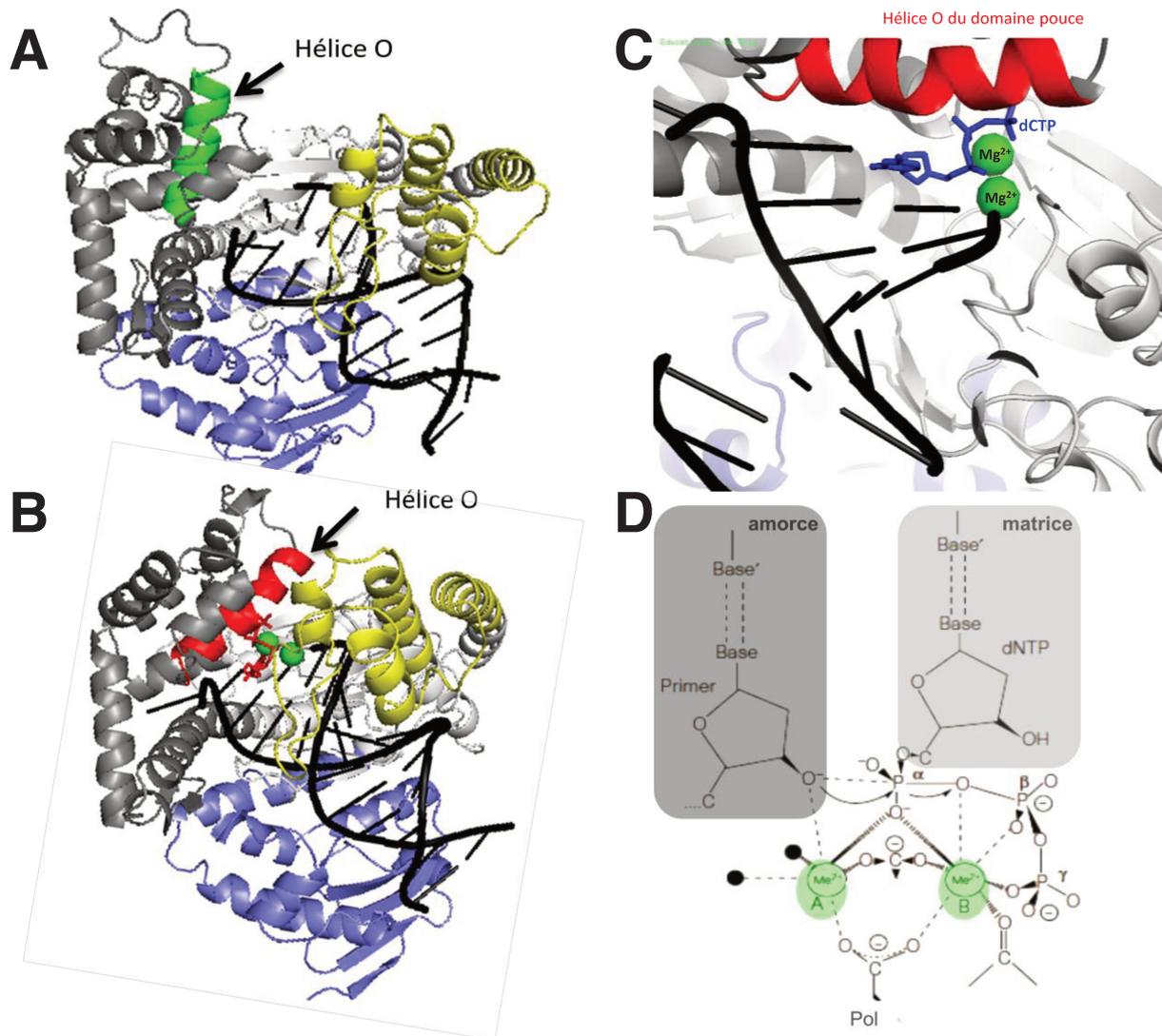
Structures secondaires de l'hybride ADN amorce/matrice	Type de réaction	Produit de la réaction Fait par Pol I de <i>E. coli</i>
Double brin 	∅	
Simple brin 	∅	
« Nick » 	Déplacement de brin ou « nick translation »	
« Gap » 	« Gap filling »	
Simple brin 	Extension d'amorce	
Simple brin amorcé 	Extension d'amorce	
	Extension d'amorce	

Figure 6. Schéma des différentes structures amorce/matrice de l'ADN et leur utilisation par l'ADN polymérase.

**I.1.2. Liaison Pol :ADN :dNTP [étapes (2)et (3)]**

Le complexe binaire Pol:ADN change de conformation et un désoxyribonucléotide triphosphate libre (dNTP) est sélectionné selon les lois de complémentarité des bases Watson-Crick [Figure 5 (2)] (notion de spécificité et de sélectivité du substrat), s'il est complémentaire à la base nucléotidique instructrice de la matrice alors le site actif de l'ADN polymérase adopte une géométrie particulière de conformation fermée [Figure 5 (3)]. Cet événement moléculaire s'opère au niveau du

domaine « doigt ». En effet, l'hélice O effectue un mouvement de 45° ce qui permet l'accessibilité du site d'insertion et l'entrée du dNTP libre (Figure 7A et 7B). Récemment, chez Klenoq1 (le fragment de Klenow de l'ADN polymérase I de *Thermus aquaticus*) un équilibre entre trois états structuraux distincts de complexe ternaire a été mis en évidence. En effet, après liaison du nucléotide, si l'appariement Watson-Crick est légitime, l'ADN polymérase adopte une conformation fermée alors que si l'appariement est illégitime, l'ADN polymérase reste en conformation ouverte et le nucléotide est libéré (Xia et al., 2013 ; Rothwell et al., 2013).



**Figure 7. Evènements moléculaires de l'ADN polymérase au cours de la synthèse d'ADN.** Structures cristallographiques du fragment de Klenow (famille A) montrant le basculement de l'hélice O du domaine « pouce » (gris) de 45° lorsqu'un dNTP canonique (rouge) est sélectionné et qui induit le passage de la pol d'une conformation ouverte (hélice O en vert) (A) (PDB 4KTQ) à une conformation fermée (hélice O en rouge) (B) (PDB 3KTQ). (C) Zoom au sein du site catalytique polymérase de la conformation fermée: les deux ions magnésium sont représentés par des sphères vertes et sont en contact avec le dCTP entrant en bleu, l'amorce ADN est en noir et l'hélice O en rouge. (D) Mécanisme réactionnel de la formation de la liaison phosphodiester dans le site catalytique polymérase, adapté de Brautigam and Steitz, 1998.

Ces caractéristiques structurales des pols de famille A, B et C et probablement celle de la famille D, permettent d'expliquer leur haut degré de sélectivité car, en effet, elles effectuent relativement très peu d'erreur dans le choix du dNTP entrant (une erreur toutes les  $10^6$  insertions) (Kunkel and Bebenek, 1988 ; Drake, 1991). En revanche, les pols de la famille Y sont connues pour effectuer la réplication de l'ADN avec un taux d'erreur élevé (une erreur toutes les 10 insertions). Le mécanisme structural lié aux taux d'erreur élevés au sein de cette famille de pol est mal connu. Une étude cristallographique de l'ADN polymérase  $\gamma$  humaine, (famille Y), a démontré récemment que les erreurs d'incorporation, préférentiellement en face de la base T, mettraient en cause le domaine « doigt » dans la faible spécificité relative aux pyrimidines (Kirouac and Ling, 2009).

### **1.1.3. Formation de la liaison phosphodiester [étape (4)]**

Au niveau du domaine « paume », qui est le domaine le plus hautement conservé chez toutes les ADN polymérases, deux ou trois acides aspartiques invariables coordonnent deux ions métalliques divalents. Le domaine « doigt » adjacent au domaine « paume », entre en contact avec le dNTP entrant et adopte une conformation fermée lorsque le nucléotide est à proximité du métal, de la matrice et du 3' du brin naissant (Figure 7A et B) (Yang et al., 1999 ; Franklin et al., 2001). L'un des ions métallique, appelé ion A, fait le lien entre l'amorce et le nucléotide sélectionné (Figure 7C), il entraîne la diminution du pKa du groupe hydroxyle 3' (3'-OH) ce qui l'active, et ainsi, induit une attaque nucléophile sur le phosphate  $\alpha$  du désoxyribonucleotide triphosphate entrant dans le site catalytique ; il y a alors formation de la liaison phosphodiester (Figure 7D). Le métal B, pour sa part, stabilise l'état de transition et assure la libération du pyrophosphate. Il s'agit du mécanisme « two-metal catalytic mechanism » (Steitz and Steitz, 1993) qui est conservé chez toutes les pols répliquatives. En structure, un troisième métal a été observé dans le site actif de l'ADN polymérase  $\eta$  humaine lors de l'insertion du nucléotide au moment de la synthèse de la liaison phosphodiester, il permettrait de stabiliser l'état intermédiaire (Nakamura et al., 2012). Egalement, dans une structure de Pol  $\beta$ , un troisième métal a été observé lors de l'insertion d'un dNTP canonique, en revanche lorsqu'une base non canonique est insérée celui-ci disparaît (Freudenthal et al., 2013). Les auteurs proposent un nouveau modèle, celui du « three-metal catalytic mechanism ».

### **1.1.4. Libération du pyrophosphate et translocation ou décrochage de l'ADN polymérase [étapes (5), (6) et (7)]**

Un changement conformationnel s'opère [Figure 5 (5)] et un pyrophosphate, produit de la réaction, est libéré. A cette étape, soit une translocation de la pol sur l'ADN s'effectue favorisant ainsi l'insertion d'un second dNTP [Figure 5 (6)], la synthèse est alors qualifiée de processive, soit l'ADN polymérase se détache de l'ADN [Figure 5 (7)], conférant une synthèse distributive. A la fourche de réplication, les ADN polymérases répliquatives coopèrent pour répliquer les deux brins d'ADN. En effet, du fait de l'antiparallélisme des brins et du sens unidirectionnel de polymérisation des pols, la réplication apparaît bidirectionnelle. A la fourche de réplication, deux types de brins se forment : le brin avancé (continu) et le brin retardé (discontinu) synthétisé dans la direction opposée au sens de l'avancement global de l'œil de réplication. De ce fait, ce dernier est constitué d'une répétition de

courts fragments, les fragments d'Okazaki. Une étape de maturation est ensuite nécessaire pour éliminer les ARN et rendre continu le brin synthétisé. Généralement, la synthèse avancée et la synthèse retardée sont coordonnées, mais le mécanisme varie selon les organismes vivants. Une étude des pils Dpo1 (pils répliquatives de *S. solfataricus*) a montré que l'association des pils en trimère pouvait augmenter à la fois le taux de synthèse d'ADN et la processivité de la réplication (Mikheikin et al., 2009).

## I.2. La fonction associée de correction des erreurs d'insertion de nucléotide

Malgré la haute sélectivité nucléotidique des pils répliquatives, celle-ci peuvent réaliser des erreurs d'incorporation de dNTP. Une fonction exonucléase existe chez certaines pils répliquatives et leur confère une haute fidélité (Tableau 2, domaine bleu), elle vient s'ajouter à la fidélité intrinsèque (elle augmente de 100 à 1 000 fois la fidélité de la pil). De plus, elle constitue une barrière à la synthèse translésionnelle (Khare and Eckert, 2002). Elle consiste en la dégradation du brin néoformé, par l'extrémité 3'-OH, lorsqu'une mauvaise insertion de dNTP est détectée (Shevelev and Hübscher, 2002). Les structures cristallographiques des pils révèlent que le site exonucléase peut être situé à distance variable du site polymérase sur une même sous-unité ou sur une sous-unité associée (Patel and Loeb, 2001 ; Kunkel, 2004 ; Mc Henry, 2011). Dans les deux cas, la distance qui sépare le site exonucléase du site polymérase nécessite qu'au moins trois nucléotides du brin d'ADN néoformé soient séparés du brin matrice (Shamoo and Steitz, 1999 ; Hogg et al., 2004 ; Subuddhi et al., 2008). De façon intéressante, il a été montré que les sites polymérase et exonucléase sont distants de 30-40 Å facilitant ainsi une coordination efficace et rapide des fonctions exonucléase et polymérase (Freemont et al., 1988 ; Wang et al., 1997 ; Kamtekar et al., 2004).

### I.2.1. Transition des modes Pol-Exo et Exo-Pol

Lorsqu'un nucléotide non canonique est inséré dans le brin néoformé au cours de l'activité polymérasique, la transition de conformation de la pil ouverte/fermée est défavorisée et la géométrie du site catalytique polymérase est imparfaite. Cela constitue un signal moléculaire pour l'activation du mode exonucléase. Le passage de la conformation du mode pil au mode exo est encore mal connu d'un point de vue moléculaire, cependant une coordination étroite entre les deux modes existe et aurait une origine structurale (Komori and Ishino, 2000 ; Kuroita et al., 2005 ; Swan et al., 2009). Le domaine « pouce », qui interagit avec le duplex d'ADN, en amont du site polymérase, est responsable, chez la famille A et la famille B, d'influencer la transition conformationnelle du mode exonucléase au mode polymérase (Doublé and Ellenberger, 1998 ; Franklin et al., 2001). Récemment, une étude par FRET (Fluorescence resonance energy transfer) a montré que la conformation fermée de l'ADN polymérase B1 de l'Archaea *S. solfataricus* n'est favorisée qu'en présence d'un dNTP canonique, et qu'un mésappariement au niveau du 3'-OH du brin néoformé favorise le positionnement de l'ADN dans le site exonucléase (Mawell and Zucchi, 2013). La vitesse de transition du brin néosynthétisé du site exonucléase vers le site polymérase est très rapide, par exemple, elle est de  $500 \text{ s}^{-1}$  chez l'ADN polymérase T4 (Fidalgo da Silva and Reha-Krantz, 2007). De façon intéressante, une étude structurale de l'interaction entre Pol B et le PCNA décrit un nouveau domaine d'interaction qui aurait

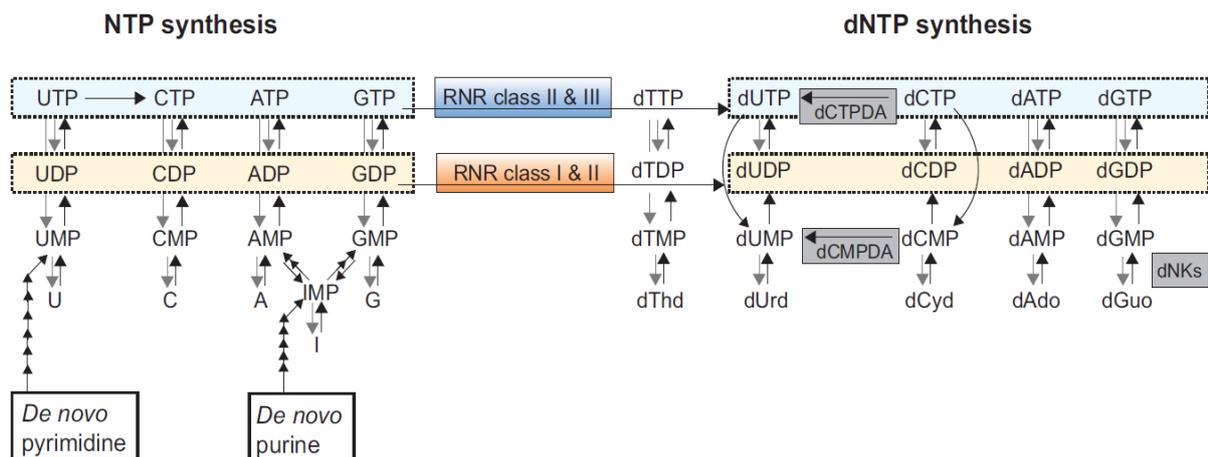
un rôle crucial dans la transition structurale des deux modes, induisant une configuration du complexe PCNA-polymérase à l'origine de la dégradation de l'amorce (Nishida et al., 2009).

### I.2.2. Mécanisme catalytique réactionnel de l'activité exonucléase

L'activité 3'→5' exonucléase s'effectue dans le domaine catalytique exonucléase des polys réplicatives. Chez la Pol A du fragment de Klenow, le mécanisme moléculaire est bien décrit. En effet, un ion métallique A et un ion métallique B sont logés dans le site catalytique et liés à la protéine par l'intermédiaire de deux ou trois acides aminés. Le métal A lie un des atomes d'oxygène de la liaison phosphodiester à rompre ainsi que l'extrémité 3'-OH de la base nucléotidique. C'est dans ce contexte qu'une molécule d'eau, ou un hydroxyle, effectue une attaque nucléophile du phosphate impliqué dans la liaison phosphodiester. Les deux métaux stabilisent l'intermédiaire réactionnel au cours de l'attaque nucléophile. Tout comme pour le site catalytique polymérase, deux ions métalliques divalents sont indispensables à l'activité catalytique exonucléase (Beese et Steitz, 1991). La mutation d'un seul des deux acides aminés responsable de la liaison au métal est d'ailleurs une stratégie de mutagenèse pour abolir efficacement l'activité catalytique exonucléase (Evans et al., 2000). Par exemple, chez *P. abyssi*, PabPolB D215A est un mutant de la fonction exonucléasique.

### I.3. Le contrôle indirect de la fidélité par la concentration intracellulaire des désoxyribonucléotides

Dans les voies de biosynthèse des nucléotides puriques ou pyrimidiques, les ribonucléotide réductases sont les enzymes qui catalysent la réduction des ribonucléotides (ADP, GDP, CDP et UDP) en désoxyribonucléotides triphosphates (dNTPs) au niveau du groupement 2'-hydroxyle. Elles participent activement à la régulation de la concentration intracellulaire des dNTPs (pool de dNTPs) par allostérie. De façon générale : NDP → dNDP → dNTP (Figure 8).



**Figure 8. Biosynthèse des bases nucléotidiques**, d'après Hofer et al., 2012.

La classe de RNR utilisée, la présence de différents désoxynucleosides kinases (dNKs) varient entre les espèces. Les abréviations d'une lettre dans la synthèse des NTPs (U, C, A, G et I) représentent à la fois la base et son ribonucléoside correspondant (I représente hypoxanthine et inosine). Les voies de dégradation ne sont généralement pas incluses à l'exception des événements de déphosphorylation, qui sont représentés par des flèches grises sauf s'ils sont directement impliqués dans la synthèse de dNTPs.

La ribonucléotide réductase et une kinase sont liées à l'ADN polymérase et permettent de diriger les dNTPs vers le site actif de la polymérase. Il en existe plusieurs isoformes. Ces enzymes sont présentes dans les trois règnes du Vivant (Lundin et al., 2010). La concentration de chaque dNTP dans la cellule ainsi que la qualité de la matrice ADN influence le taux d'erreur d'incorporation (Kunkel, 2011) et n'est pas seulement imputable à l'ADN polymérase. Il est possible de mesurer la concentration intracellulaire des nucléotides, plusieurs méthodes existent, ainsi, une réaction enzymatique utilisant le fragment de Klenow et le marquage radioactif a initialement été mise au point (Sherman and Fyfe, 1989) puis améliorée (Ferraro et al., 2010). Aujourd'hui, la méthode de choix utilisée est un dosage par chromatographie liquide à haute performance (HPLC) suivie d'une quantification dans l'ultra-violet (UV).

Le tableau 3 répertorie les valeurs en dNTPs et NTPs mesurés chez différentes espèces du monde Vivant. De façon reproductible, le dTTP est le désoxyribonucléoside triphosphate qui apparaît être le plus abondant dans les cellules eucaryotes.

Domaine	Bacteria		Eukarya	
Organisme	<i>Escherichia coli</i>		<i>Saccharomyces cerevisiae</i>	Fibroblastes humains
dNTPs	dATP	42 (pmol/10 <sup>8</sup> cellule)	16 (μM)	24 (μM)
	dTTP	117 (pmol/10 <sup>6</sup> cellule)	30 (μM)	37 (μM)
	dCTP	250 (pmol/10 <sup>6</sup> cellule)	14 (μM)	29 (μM)
	dGTP	33 (pmol/10 <sup>6</sup> cellule)	12 (μM)	5.2 (μM)
NTPs	ATP	Non déterminé (ND)	3000 (μM)	3150 (μM)
	UTP	ND	1700 (μM)	570 (μM)
	CTP	ND	500 (μM)	280 (μM)
	GTP	ND	700 (μM)	470 (μM)
ratio NTPs/dNTPs	A	ND	188	130
	T/U	ND	57	15
	C	ND	36	14
	G	ND	58	90
Références	Gon et al., 2011		Nick McElhinny et al., 2010	Traut, 1994

Tableau 3. Valeurs de chaque dNTPs, NTPs et de leurs ratios mesurés chez différents organismes du Vivant.

La fidélité de la réplication de l'ADN peut être influencée par la quantité relative des quatre dNTPs. Par exemple, une sur-production de dTTP associée à une sous-production de dCTP peut favoriser l'insertion d'un dTTP en face d'une base G et augmenter ainsi le taux d'erreur d'incorporation par l'ADN polymérase. Chez la levure *Saccharomyces cerevisiae*, il a été montré qu'en réponse à un dommage à l'ADN, la concentration du pool de dNTP augmente (Chabes and Stillman, 2007) suite à une sur-régulation des RNRs (Håkansson et al., 2006). Ceci influence directement l'activité polymérase de pol réplivative en favorisant le passage d'une lésion à l'ADN (Sabouri et al., 2008). Chez *E. coli*, cette augmentation du pool de dNTPs en réponse à un dommage a aussi été mesurée et jouerait un rôle dans la mutagenèse spontanée et induite (Gon et al., 2011). Chez *S. cerevisiae*, il a

été montré que la concentration en dNTPs augmente de 6 à 8 fois en réponse à un dommage à l'ADN et que cette augmentation est associée à une tolérance au dommage à l'ADN suggérant que la synthèse translésionnelle d'ADN est plus efficace lorsque la concentration en dNTPs est élevée (Chabes et al., 2003).

#### I.4. Le cas de l'incorporation de ribonucléotides

En raison de leur faible spécificité de substrat, il n'est pas rare que les Pols X, impliquées dans la réparation de l'ADN, incorporent des ribonucléotides (NTPs). Ainsi, Pol  $\mu$  humaine est incapable de discriminer le groupement 2'-OH du ribose des nucléotides et par conséquent elle insert des NTPs autant que des dNTPs. Un seul acide aminé semble être responsable de cette discrimination, il s'agit de Gly433, *in vitro*, sa mutation engendre une augmentation de la discrimination dNTP/NTP par Pol  $\mu$  (Ruiz et al., 2003).

Il n'y a que récemment qu'une attention particulière a été portée sur l'impact des NTPs dans l'influence de la fidélité des pols répliquatives eucaryotes. En effet, le site actif des ADN polymérases adopte une conformation stérique permettant de discriminer les dNTPs des NTPs et leur confère une haute spécificité de substrat (Joyce, 1997). Elles exercent un contrôle strict géométrique et par conséquent détectent la nature chimique du nucléotide (hydroxyle en position 2') par un seul acide aminé nommé « the steric-gate » (Brown and Suo, 2011). Chez les organismes Vivants, compte tenu de la concentration intracellulaire relativement faible des dNTPs par rapport aux NTPs il est possible que les NTPs soient incorporés par les pols. En effet, il a été montré *in vitro* que l'ADN polymérase  $\alpha$ , l'ADN polymérase  $\delta$  et l'ADN polymérase  $\epsilon$  pouvaient incorporer des NTPs avec un taux d'extension élevé pendant la synthèse d'ADN (Nick McElhinny et al., 2010a). De plus, *in vivo*, chez *S. cerevisiae* et *S. pombe*, ces trois pols incorporent des dNTPs en face de NTPs, provoquant des délétions génétiques de 2 à 5 nucléotides (Nick McElhinny et al., 2010b ; Miyabe et al., 2011). Le mécanisme exact lié à l'erreur chez *S. cerevisiae* n'est pas connu mais Pol  $\alpha$ , Pol  $\delta$  et Pol  $\epsilon$  effectuent des arrêts de polymérisation lorsqu'un ribonucléotide est rencontré dans la matrice d'ADN et ceci pourrait conduire à des délétions dans les séquences répétées (Watt et al., 2011). Chez l'Homme, la Pol  $\epsilon$  de la famille B est également capable d'incorporer des NTPs (Göksenin et al., 2012).

De façon très intéressante, récemment, une étude a montré chez la Pol B archéenne *Tgo* que la mutation E664K dans le site actif polymérase conjointe à la mutation de la « steric-gate » (Y409G) transforme l'ADN polymérase en une ARN polymérase capable de synthétiser des fragments d'ARN de plus de 1,7 kb (Cozens et al., 2012). Ainsi se posent des questions extrêmement intéressantes relatives à l'évolution adaptative existant entre les ARN polymérases et les ADN polymérases en lien avec leur spécificité de substrat.

## II. Le contrôle des dommages à l'ADN lors de sa réplication

### II.1. Les différents types de dommage à l'ADN

Des lésions à l'ADN surviennent en permanence suite à l'action de diverses sources chimiques (radicaux libres de l'oxygène, agents alkylants) et physiques (rayonnements ionisants, agitation thermique) mais aussi dans les conditions physiologiques normales par la voie spontanée. Les lésions générées sont de natures très différentes et apparaissent de l'échelle moléculaire à l'échelle chromosomique. Les lésions peuvent être des modifications chimiques des bases azotées de l'ADN, des cassures simple brin de l'ADN, des pontages intrabrans et interbrans, des pontages ADN protéines et finalement des cassures double brin de l'ADN qui perturbent l'intégrité du chromosome entier (Figure 9).

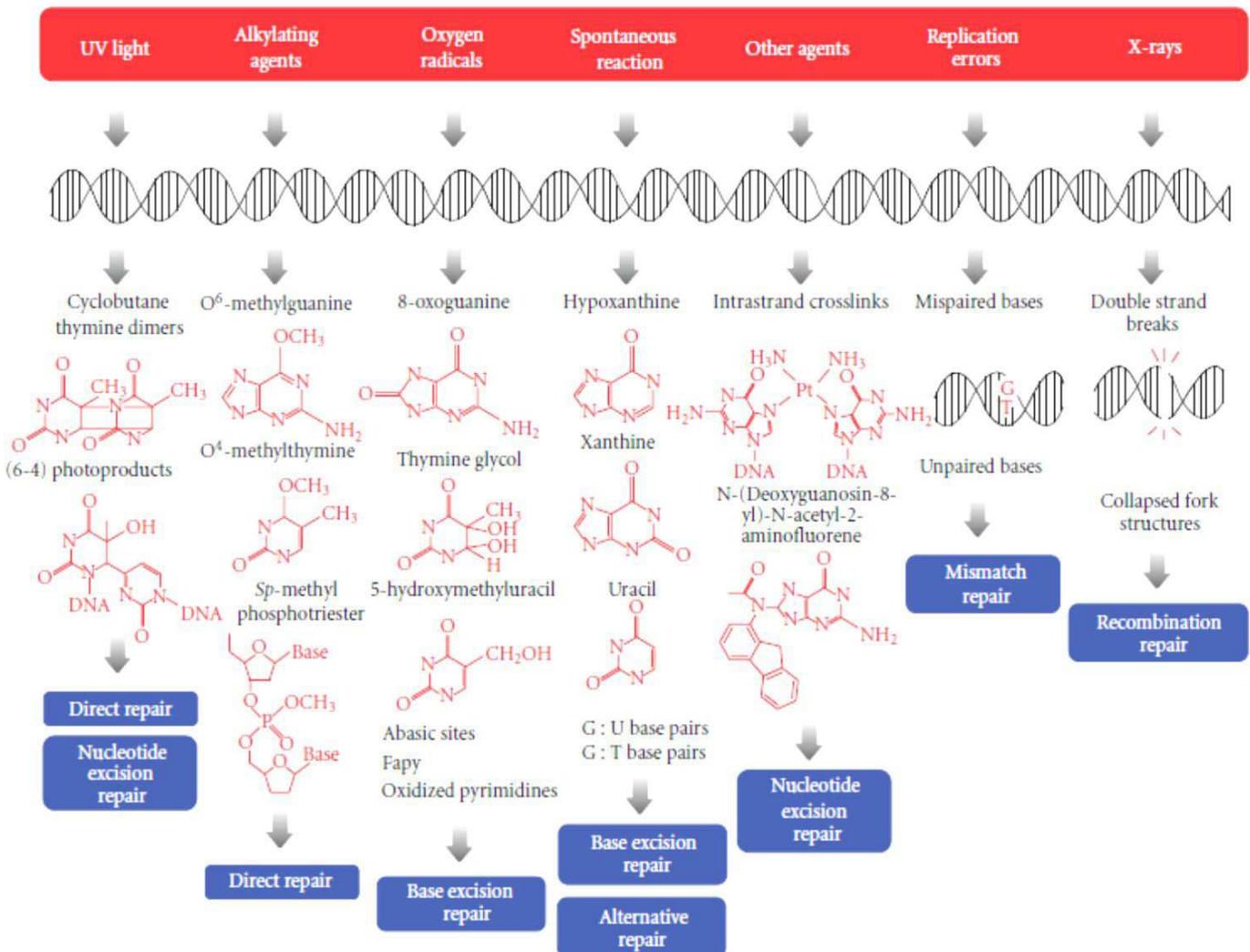
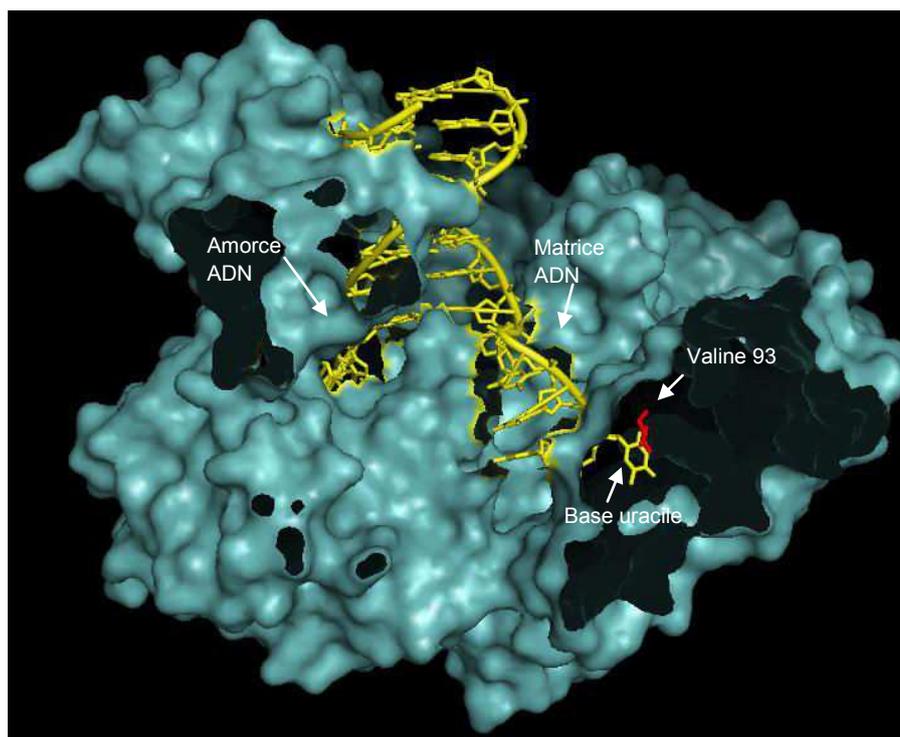


Figure 9. Les différentes lésions de l'ADN et les voies de réparation spécifiques associées, d'après Morita et al., 2010.

## II.2. La reconnaissance des bases désaminées par les ADN polymérases répliquatives archéennes

Les polymérases répliquatives de la famille B et uniquement de type archéennes (Wardle et al., 2008) possède un domaine N-terminal en amont du domaine exonucléase dont la fonction est de détecter la présence d'une base uracile (produit de la désamination de la base cytosine) ou d'une base hypoxanthine (produit de la désamination de la base adénosine) dans le brin d'ADN matrice qui aurait échappé au mécanisme de réparation par l'UDG (uracil-N-glycosylase) avant qu'il n'atteigne le site polymérase (Greagg et al., 1999 ; Fogg et al., 2002). Lorsqu'une base désaminée est reconnue, l'ADN polymérase effectue une pause dans la réplication exactement quatre bases en amont de la lésion. Plusieurs structures de cristaux de Pols B archéennes en complexe avec une amorce d'ADN hybridée à une matrice contenant une base uracile ont été obtenues (Firbank et al., 2008 ; Killelea et al., 2010 ; Gouge et al., 2012). Les bases désaminées sont enfouies dans une poche située dans leur domaine amino-terminal (Figure 10) et l'ADN adopte une position proche de celle observée lorsque la pol est en mode exonucléase.



**Figure 10. Positionnement de la base uracile dans la poche de reconnaissance des bases désaminées de la Pol B de *P. abyssi* (PDB 4FM1).** Le duplex d'ADN (jaune) comporte un mésappariement de trois bases, ce qui favorise la transition structurale de la pol en mode exonucléase. Le brin amorce se positionne dans le domaine exonucléase (à gauche) et la matrice se situe dans le domaine N-terminal (spécifique des pols répliquatives) (à droite) et plus particulièrement, la base uracile effectue une torsion et se positionne dans la poche de reconnaissance. L'acide aminé V93 (rouge), est impliqué dans le mécanisme de reconnaissance, sa mutation induit une exclusion stérique de l'uracile de la poche de reconnaissance (Firbank et al., 2008).

De plus, des études biochimiques ont montré que lorsque le site catalytique polymérase de la pol se rapproche de la base désaminée situé en position +3 +2 +1 voir même 0, l'activité exonucléase est fortement stimulée, ainsi la polymérase ne peut pas répliquer au-delà de la base désaminée (Russell et al., 2009). Une étude souligne un déroulement prononcé des brins amorce/matrice lorsque la pol rencontre un uracile, or la séparation des deux brins est nécessaire à l'activité exonucléase ; un acide aminé particulier (Y261) semble être responsable de la séparation des deux brins d'ADN (Richardson et al., 2013b). Ce mécanisme spécifique des Archaea, initialement caractérisé chez les Pol B d'Euryarchaea a aussi été identifié chez les Pols B1 de Crenarchaea (Greagg et al, 1999 ; Grüz et al., 2003 ; Savino et al., 2004 ; Asami et al., 2006). Très récemment, une inhibition de la réplication d'un ADN contenant une base uracile a été mise en évidence chez des pols de la famille D archéenne, les détails du mécanisme ne sont pas encore établis mais le mécanisme d'inhibition serait différent de celui observé chez la famille B (Richardson et al., 2013a).

### II.3. La tolérance aux dommages à l'ADN par les ADN polymérases de la famille Y

Face à des dommages sur les bases de l'ADN, la majorité des pol répliquatives, en raison de leur haute sélectivité et fidélité sont bloquées. Il en résulte un arrêt de la réplication. En revanche, les Pols Y sont capables de franchir des lésions à l'ADN en incorporant un nucléotide en face de la base lésée, il s'agit de la synthèse translésionnelle. De plus, elles peuvent incorporer des dNTPs endommagés (Shimizu et al., 2003). En plus du cœur catalytique, composé des domaines « doigt », « paume », « pouce », présent chez toutes les familles de pol, les Pols Y possèdent un domaine qui leur est spécifique, le domaine PAD (polymerase associated domain) (Ling et al., 2001; Silvan et al., 2001; Trincao et al., 2001) dit aussi « domaine petit doigt » qui aurait un rôle dans la modulation de leur spécificité de substrat (Boudsocq et al., 2004). Les Pols Y ne possèdent pas d'activité de relecture exonucléase ce qui explique qu'elles soient peu fidèles lors de la synthèse d'ADN canonique (Kunkel, 2004). De plus, contrairement aux pols des autres familles, elles ne possèdent pas la poche de liaison à l'ADN pour accommoder les paires de bases complémentaires aux niveau de l'extrémité du brin néosynthétisé (Ling et al., 2001). Compte tenu de l'ensemble de ces caractéristiques, les Pols Y peuvent accommoder des structures erronées d'ADN et donc répliquer de l'ADN contenant des dommages. Le principal rôle des Pols Y réside dans la voie de tolérance aux lésions de l'ADN car en effet, si la cellule échoue dans la réparation des dommages de son ADN, les Pols Y peuvent éviter ces lésions en effectuant une synthèse d'ADN en face du site endommagé par un processus appelé synthèse translésionnelle (Prakash et al., 2005). Après l'insertion de quelques nucléotides, leur forte distributivité les restreint dans leur capacité d'extension, et de ce fait évite l'accumulation d'erreur dans la réplication. De plus, le caractère distributif des Pols Y permet de faciliter le transfert pol translésionnelle/pol répliquative à la fourche de réplication.

Elles sont présentes dans les trois règnes du Vivant (voir Tableau 2 pages 36-37). Chez les eucaryotes il y a quatre famille Y : Pol  $\eta$ , Pol  $\iota$ , Pol  $\kappa$  et Rev1 (Ohmori et al., 2001 ; Shcherbakova and Fijalkowska, 2006). Il a été montré *in vitro* que chacune de ces pols a une capacité de translésion spécifique d'un type particulier de lésion. En effet, alors que Pol  $\eta$  effectue de la synthèse

translésionnelle en face des dimères de thymine (McCulloch et al., 2004), Pol  $\iota$  a plus de difficulté à contourner ces dommages (Johnson et al., 2000 ; Tissier et al., 2000) et semble être plus efficace dans la translésion d'adduits à l'ADN (Perrino et al., 2005; Wolffe et al., 2005). De nombreuses études suggèrent que ces pols sont impliquées dans la mutagenèse spontanée ou induite par les dommages (Bavoux et al., 2005 ; Abdulovic and Jinks-Robertson, 2006 ; Dumstorf et al., 2006). Une étude très récente a montré que la pol humaine Pol  $\kappa$  est impliquée dans la synthèse de courts brin d'ADN lors de l'arrêt d'une fourche de réplication et faciliterait la mise en place des points de contrôle de la division cellulaire, ce mécanisme participant à la maintenance génomique (Bétous et al., 2013). De plus, cette propriété de la famille Y, chez les mammifères, est utilisée dans le processus d'hypermuation somatique, phénomène de mutation ciblée qui s'opère au niveau des gènes codant pour les immunoglobulines et qui contribue à la production d'anticorps hautement affins (Seki et al., 2005). Chez les procaryotes, les Pols Y qui sont Pol IV et Pol V ont clairement été décrites comme impliquées dans la mutagenèse spontanée ou induite par les agents mutagènes (Tang et al., 2000 ; Fuchs et al., 2004). Le taux d'erreur de mésincorporation est relativement élevé chez les Pols Y eucaryotes et bactériennes, de  $10^{-3}$  à  $10^{-4}$  par nucléotides (Goodman, 2002).

Au sein des Archaea, une seule famille de Pols Y a été identifiée et caractérisée dans l'embranchement des Crenarchaea (Tableau 2). Il s'agit de la pol Dpo4 de *Sulfolobus solfataricus* (Boudsocq et al., 2001) et de la pol Dbh de *Solfataricus acidocaldarius* (Kulaeva et al., 1996). Elles ont également un rôle dans la translésion (Potapova et al., 2002) et effectue du « slippage » qui est à l'origine de la création de délétion de base (Wu et al., 2011). Le taux d'erreur de mésincorporation est relativement plus faible que leurs homologues bactériens et eucaryotes, de  $10^{-1}$  à  $10^{-3}$  par nucléotides (Kunkel et al., 2004). Une étude portant sur la délétion du gène codant pour Dbh chez *Sulfolobus solfataricus* a montré que l'absence de cette pol ne perturbait pas la croissance de la souche ni sa sensibilité aux agents mutagènes. En revanche, la substitution des paires de bases G :C  $\rightarrow$  T :A était favorisée chez les mutants (Sakofsky et al., 2012).

Il est important de noter que d'autres familles de pols, et notamment les pols répliquatives, peuvent aussi effectuer de la synthèse translésionnelle. En effet, il a été montré *in vitro* que la Pol B répliquative de *Pyrococcus abyssi* effectuait de la synthèse translésionnelle lorsqu'elle rencontrait des sites apuriques/aprimidique (Palud et al., 2008). Le dommage uracile peut également être toléré par cette pol (Gouge et al., 2012) et par la Pol D (Richardson et al., 2013a). Chez la Pol B de *Sulfolobus solfataricus*, cette propriété a aussi été mise en évidence (Grúz et al., 2003). L'ADN polymérase répliquative humaine Pol  $\delta$  est un hétérotétramère, suite à l'activité d'agents génotoxiques comme les UV, Pol  $\delta$  est convertie en trimère suite à la dégradation de la petite sous unité p12. *In vitro*, le trimère montre des propriétés biochimiques modifiées, il effectue moins de synthèse translésionnelle des sites AP, en revanche, il possède une meilleure activité exonucléase et effectue moins d'erreur lors de l'insertion de nucléotides. La conversion du tétramère en trimère consiste finalement à une réponse cellulaire aux dommages à l'ADN (Meng et al., 2008). Même si cette propriété des pols répliquatives peut engendrer des mutations au niveau du génome, elle permet en outre d'éviter le blocage et l'arrêt de la réplication de l'ADN.

## II.4. La réparation des dommages à l'ADN par les ADN polymérases de la famille X

La cellule possède toute une batterie de systèmes plus ou moins complexes qui lui permet de détecter les dommages à l'ADN, de les réparer ou de s'en accommoder. Ainsi, on recense six principaux systèmes de réparation au sein des trois règnes du Vivant. Il s'agit de la réparation directe de la lésion (photolyase pour les dimères de thymine, méthyltransférases pour m6G, m1A, m3C), la réparation par excision de base ou « *base excision repair* » (BER), la réparation par excision de nucléotides ou « *nucleotide excision repair* » (NER), la réparation des mésappariements ou « *mismatch repair* » (MR), le « *non-homologous end-joining* » (NHEJ) et la réparation par recombinaison homologue (Figure 11). La première catégorie comprend des mécanismes *ad hoc*, spécifiques d'un type de lésion donnée. Les cinq derniers systèmes sont généralistes et sont chacun capable de réparer un ensemble de lésions diverses.

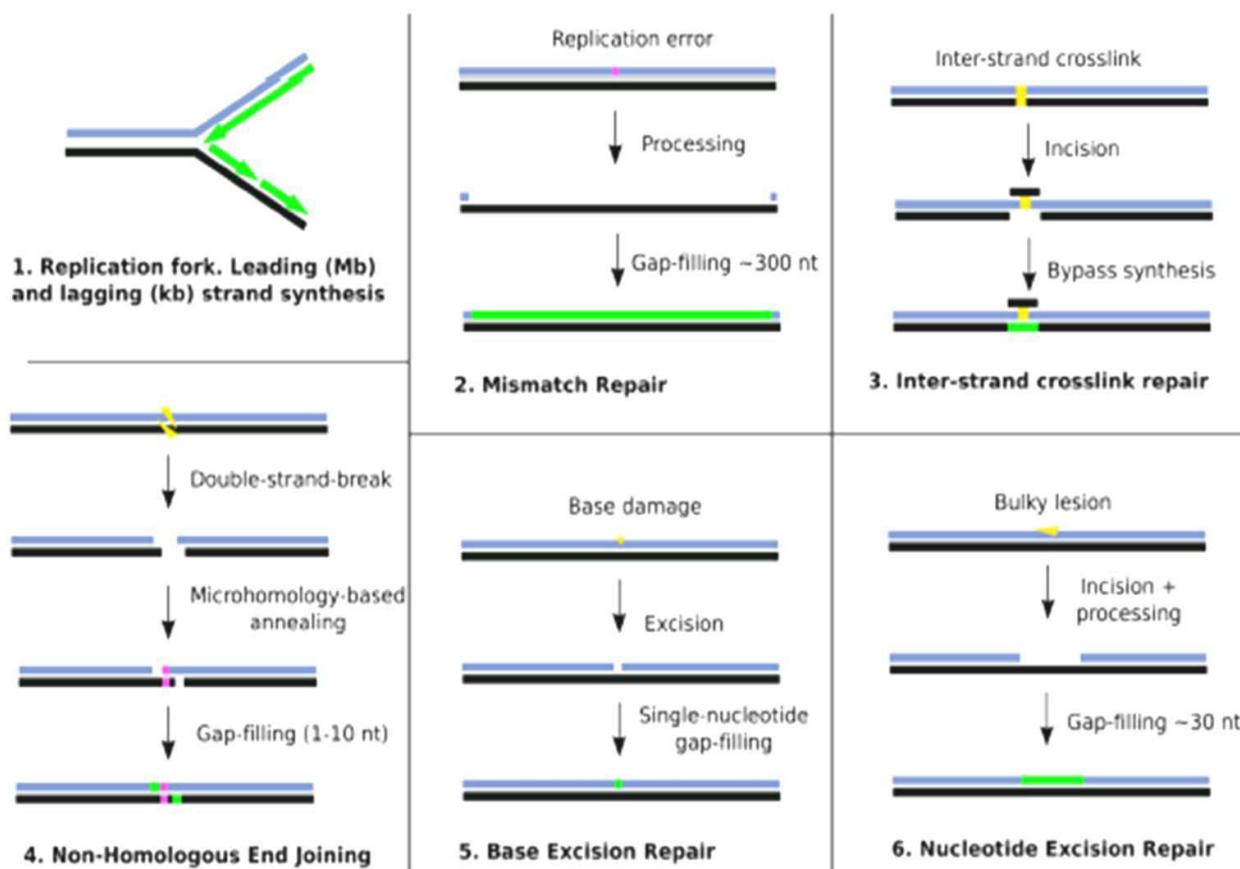


Figure 11. Activités spécialisées des différentes voies de la réparation sur l'ADN, d'après Garcia-Diaz and Bebenek, 2007.

La forte diversité des pols impliquées dans les voies de la réparation de l'ADN explique leur place au sein des diverses voies de réparation de l'ADN. La famille X est présente chez les

eucaryotes, chez les procaryotes ainsi que chez certains virus, (Garcia-Diaz et al., 2000 ; Garcia-Diaz et al., 2005). Chez les Archaea, aucune pol de cette famille n'a pour le moment été identifiée. Les membres de la famille X [Pol  $\beta$ , pol  $\lambda$ , Pol  $\mu$ , Pol  $\sigma$  et la terminal deoxynucleotidyl transferase (TdT)] ont des structures tridimensionnelles similaires mais des propriétés biochimiques différentes. La particularité structurale de la plus par des membres de cette famille est la présence d'un domaine de liaison à l'ADN à l'extrémité N-terminal de 8 kDa qui est impliqué dans la liaison aux substrats d'ADN « gapped » (Beard et al., 2006) et serait impliqué dans le mécanisme de remplissage de brin. Elles se distinguent dans la diversité des interactions avec l'ADN, de ce fait, elles interviennent dans différentes voies de réparation à l'ADN.

Durant la première phase du BER, des bases intermédiaires apurinique/apirimidique (AP) hautement mutagènes sont produites. Les endonucléase AP jouent alors un rôle important, dans le sens où elles reconnaissent les résidus abasiques et hydrolysent la liaison phosphodiester créant une structure intermédiaire avec une extrémité 3'-OH et une extrémité 5'-dRP. Les membres de la famille X sont alors capables d'éliminer le résidu dRP et d'incorporer le nucléotide manquant.

La pol  $\beta$  et la pol  $\lambda$  ont une activité spécifique 5'-désoxyribose phosphate (dRP) lyase portée par un domaine accessoire en N-terminal. Leur activité polymérase se situe dans le domaine C-terminal. Pol  $\beta$  insère préférentiellement un seul nucléotide au cours du BER (Wilson et al., 2000) alors que Pol  $\lambda$  et Pol  $\mu$  interviennent dans la réparation des cassures doubles brins. Au cours de la voie du BER, Pol  $\beta$  hydrolyse le 5'désoxyribose résultant du clivage par l'endonucléase de la liaison phosphodiester en 5' du ribose endommagé ; elle comble la brèche et la ligature s'effectue par une ADN ligase. Pol  $\beta$  peut répliquer les lésions induites par les UV comme les dimères de thymine, cela suggère une implication de cette pol dans l'instabilité génétique induite par les UV (Servant et al., 2002). Pol  $\beta$  est la principale dRPLYase impliquée dans la réparation des lésions oxydatives telle que la 8-oxoguanine et la 5,6-dihydrouracil (Allinson et al., 2001 ; Batra et al., 2012).

Pol  $\lambda$  et Pol  $\mu$  ont la propriété intrinsèque de synthétiser l'ADN de façon discontinue sur les extrémités terminales des ADN appariés ou non (Domínguez et al., 2000), cette propriété, due à la présence d'un domaine BRCT les rend aptes à la participation à la voie du NHEJ. De plus, Pol  $\lambda$  est impliquée dans la voie de réparation des lésions oxydatives (8oxoG notamment) et agit de concert avec la glycosylase MutY (Markkanen et al., 2012). Les pols TdT effectuent de la synthèse d'ADN indépendamment du brin matrice et participent au NHEJ qui a lieu dans les cellules immunitaires au niveau du gène récepteur de l'antigène (Nick McElhinny and Ramsden, 2004 ; Bertocci et al., 2006). Les cellules eucaryotes portent également un autre membre de la famille X, la Pol  $\sigma$ . Elle serait impliquée dans la cohésion des chromatides chez *S. cerevisiae* (Wang et al., 2000).

A l'échelle de l'organisme, chez les mammifères, les pols de la famille X ont une spécificité cellulaire et ne s'expriment pas dans tous les tissus. Aussi, des études de délétions de gènes ont permis de montrer des dysfonctionnements du système immunitaire en l'absence de Pol  $\lambda$  ou de Pol  $\mu$

chez la souris, alors que l'absence de Pol  $\beta$  engendre une mort rapide et met en évidence son rôle essentiel à la survie des mammifères.

De façon intéressante, une Pol X a été identifiée chez le procaryote *Bacillus subtilis* (PolX<sub>Bs</sub>), elle effectue le remplissage de petites brèches à l'ADN et elle pourrait jouer un rôle dans la réparation de l'ADN (Baños et al., 2008). D'autres Pols X bactériennes, celles de *Thermus thermophilus* (ttPolX) (Nakane et al., 2009) et *Deinococcus radiodurans* PolX(Dr) (Lecointe et al., 2004 ; Blasius et al., 2006) ont été caractérisées. Ces Pols X bactériennes présenteraient une activité AP-endonucléase intrinsèque qui, en coordination avec l'activité polymérase, leur permettraient de reconnaître, exciser et réparer les sites AP (Baños et al., 2010). Par ailleurs, une délétion de Pol X(Dr) chez *D. radiodurans*, exposé aux rayons gamma dans le but d'induire des cassures doubles brins à l'ADN, entraîne la diminution de la survie des bactéries mutées (Lecointe et al., 2004). Le mécanisme de gap-filling d'une seule base a été caractérisé chez ttPol X et de façon curieuse cette pol semblerait lier le substrat dNTP avant de se lier à l'ADN endommagé (Nakane et al., 2012). Les Pols X bactériennes possèdent des similarités de séquence avec celles des Pols X eucaryotes, en effet, elles possèdent un domaine additionnel polyhistidinol phosphatase (PHP) dans la partie C-terminale. Ce domaine est également présent chez les pols répliquatives de la famille C.

Même si chaque membres des familles de pols répliquatives, translésionnelles et de la réparation ont des activités catalytiques très spécifiques et sont impliquées dans des voies enzymatiques distinctes, une certaine dynamique et collaboration s'effectue entre les pols. Par exemple, au cours de la réplication lorsqu'un dommage est détecté, une dynamique d'interaction entre Pol B et Pol Y a été mise en évidence chez *Sulfolobus solfataricus* (De Felice et al., 2007).



# Troisième partie :

## Les Archaea hyperthermophiles en tant que micro-organismes modèles de l'étude de la maintenance génomique

### I. La réplication chez les Archaea

La réplication chez les Archaea s'effectue selon un mode de réplication de type bactérien mais implique des homologues de protéines de la réplication de types eucaryotes (Figure 12) (Myllykallio et al., 2000 ; Grabowski and Kelman, 2003 ; Barry and Bell, 2006 ; Cann, 2008).

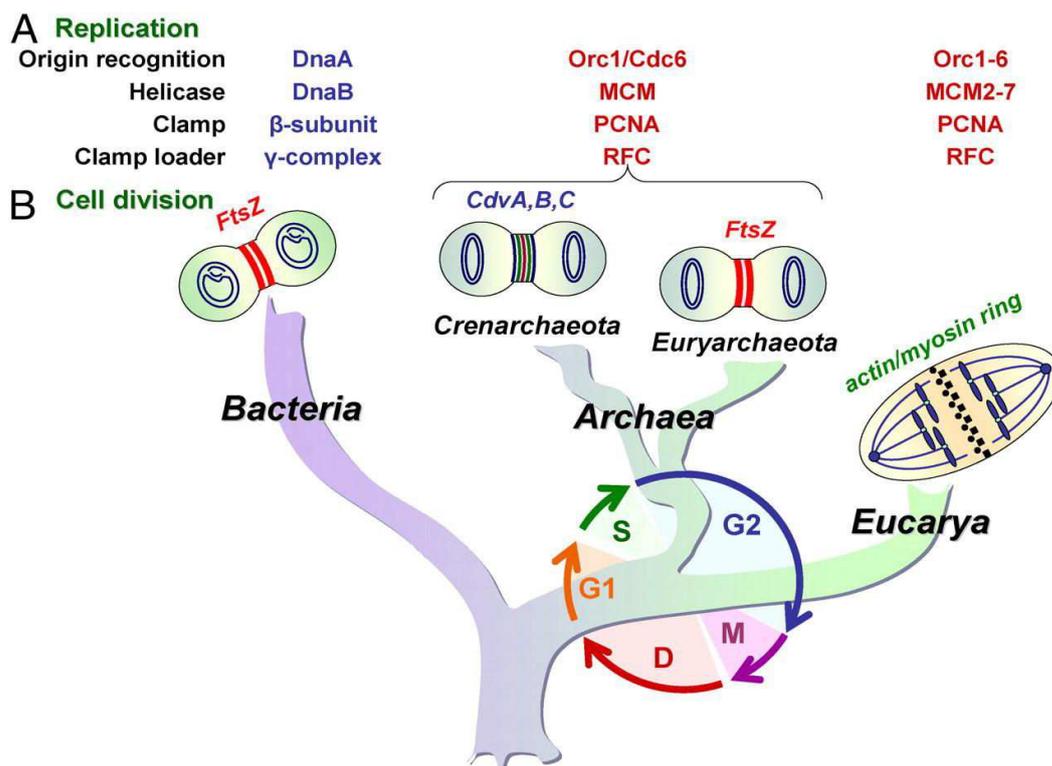
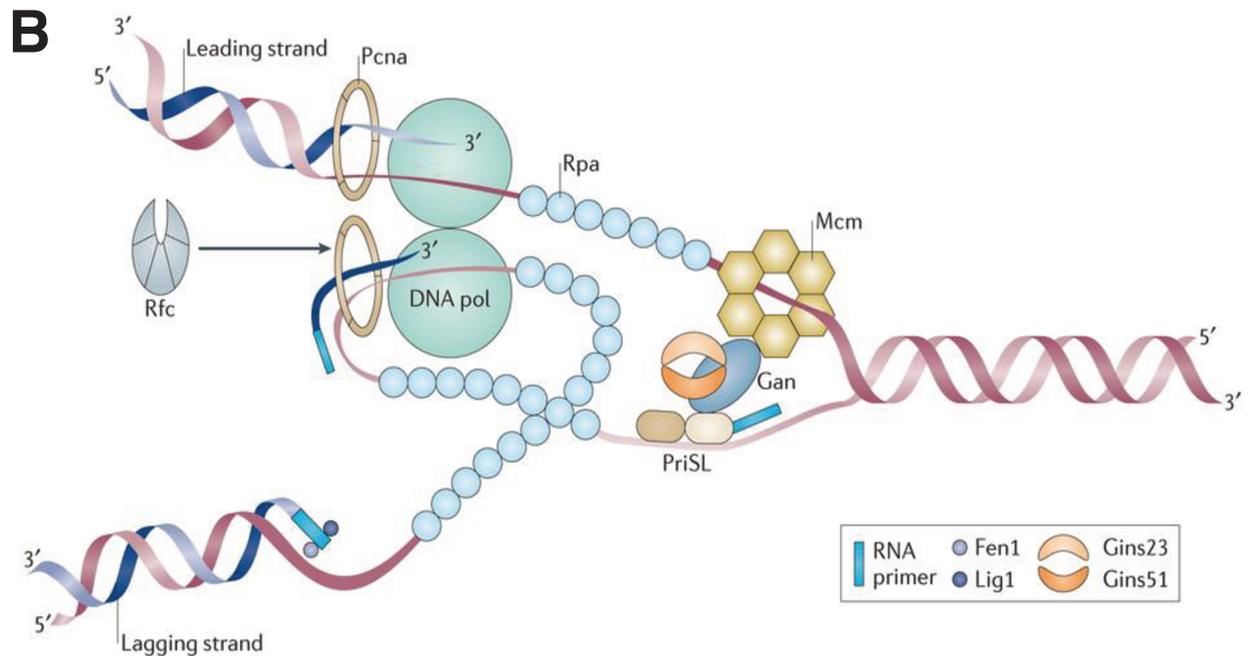
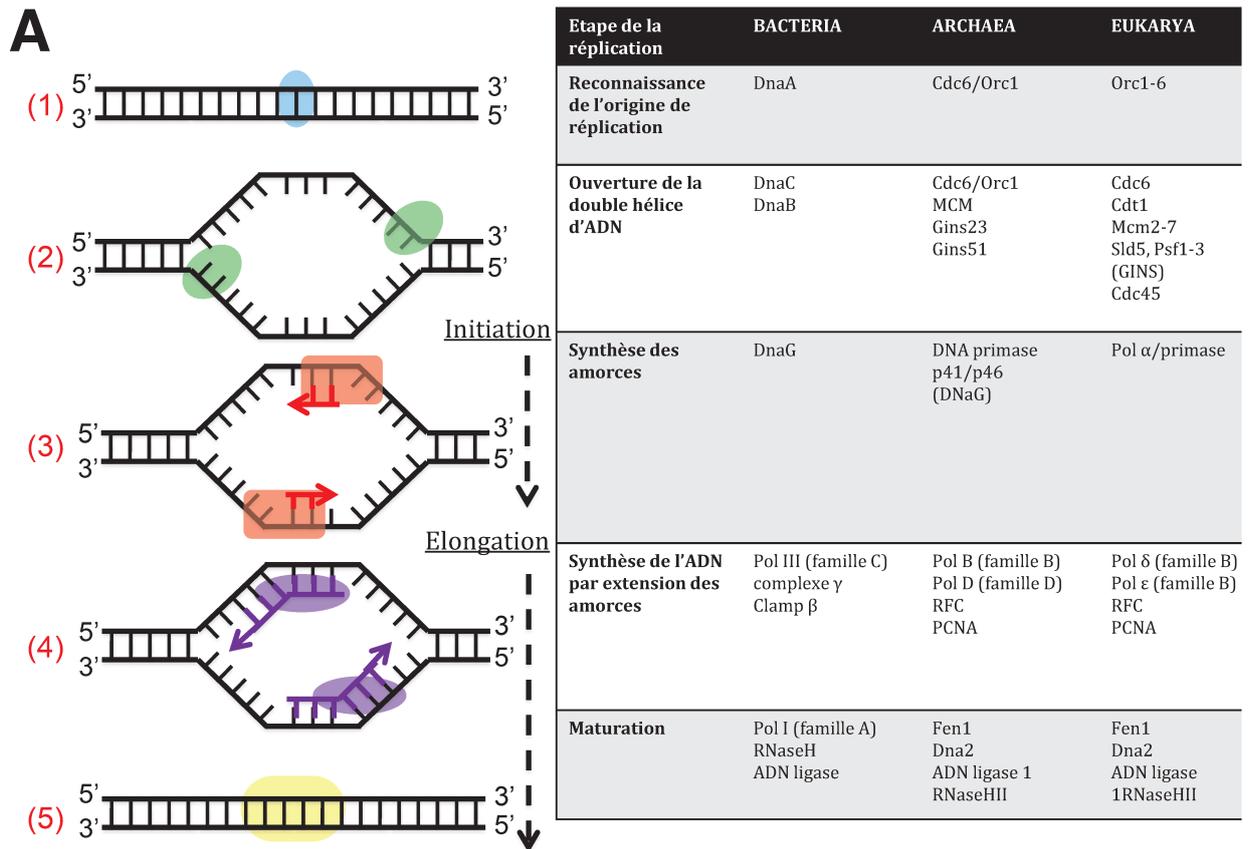


Figure 12. Schéma montrant la similitude de la machinerie de la réplication entre les Archaea et les Eucarya (A) et la différence de mécanisme de division cellulaire entre les Archaea et les Eucarya (B), d'après Cann, 2008.

En effet, le chromosome est circulaire tout comme chez les bactéries et la taille moyenne est d'environ 1,5 Mbp. Chez les Archaea, l'ADN est répliqué de façon bidirectionnelle, à grande vitesse, à partir d'une origine de réplication, c'est le cas de *Pyrococcus abyssi* (Myllykallio et al., 2000) ; ou de plusieurs origines de réplication, c'est le cas par exemple de la Crenarchaea *Sulfolobus solfataricus*

(Lundgren et al., 2004), de l'Euryarchaea *Haloferax volcanii* (Norais et al., 2007), et de la Crenarchaea *Pyrobaculum calidifontis* (Pelve et al., 2012). L'organisation des opérons et le motif de Shine-Dalgarno nécessaire à l'initiation de la traduction sont communs aux Archaea et aux bactéries (Karlin et al., 2005). De plus, les fragments d'Okazaki caractérisés chez les Archaea ont des tailles de 150 pb tout comme chez les eucaryotes, chez les bactéries ils sont de plus grande taille, de 1 à 2 kb (Matsunaga et al., 2003). Le RPA (replication protein A) des Archaea comporte trois sous-unités, tout comme chez les eucaryotes (Komori and Ishino, 2001). Cependant, des différences au sein même des Archaea, en fonction de l'embranchement, ont été mises en évidence. En effet, chez les Crenarchaea, il existe trois isoformes de PCNA (De Felice, 1999 ; Kawarabayasi, 1999, Kawai et al., 2011), deux ou trois ADN polymérases répliquatives de la famille B (Pisani et al., 1992 ; Edgell et al., 1997 ; Cann et al., 1999a), et une ADN polymérase de la famille Y (Kulaeva et al., 1996, Boudsocq et al., 2001). Chez les Euryarchaea, à l'heure actuelle, un PCNA (Cann et al., 1999b), une ADN polymérase de la famille B (Cann et al., 1999b) et une ADN polymérase de la famille D (Uemori et al., 1997 ; Cann et al., 1998 ; Gueguen et al., 2001) ont été identifiés. Il existe en fait différents réplisomes, un pour chaque domaine du Vivant, cependant ils ont tous les trois un mécanisme réactionnel commun menés par des acteurs protéiques fonctionnellement et structurellement conservés.

Voici les différentes étapes de la réplication de l'ADN au niveau de la fourche de réplication et l'intervention séquentielle des acteurs protéiques dans ce mécanisme pour le cas du réplisome archaéen (Figure 13A et 13B). Il y a reconnaissance de l'origine de réplication par les protéines OBP (Origin Binding Proteins), Cdc6/Orc1 s'y lie, ce qui provoque une torsion de la double hélice d'ADN et son déroulement local rendant ainsi l'origine de réplication accessible aux acteurs protéiques de la réplication [Figure 13A (1)]. Ensuite, l'hélicase répliquative appelé aussi le complexe de maintenance minichromosomique (MCM) ouvre la double hélice d'ADN [Figure 13A (2)] et les protéines de liaison à l'ADN simple brin RPA se lient à l'ADN déroulé en vue de le protéger. Gins23 et Gins51 interagissent avec la primase et avec le MCM par l'intermédiaire d'une nucléase associée aux Gins (Gan) et forment ainsi la base structurale du réplisome (Figure 13B). Ensuite, au niveau du brin retardé et du brin avancé, l'étape d'initiation s'effectue : la primase (p41/p46) synthétise les fragments d'ARN [Figure 13A (3)] ; puis, il s'agit de l'étape d'élongation : tout d'abord le facteur de processivité (PCNA) trimérique, est chargé sur l'ADN à l'aide du facteur de réplication C (RFC) pentamérique de façon ATP-dépendante, le PCNA encercle alors l'ADN et positionne les ADN polymérases B ou D au niveau de la jonction amorce/matrice. Ainsi, les ADN polymérases effectuent la synthèse de l'ADN par extension des amorces [Figure 13A (4)]. Enfin, au cours de l'étape de maturation des fragments d'Okazaki, les amorces d'ARN sont enlevées par Fen1 ou par RNaseHII/Fen1 et les fragments d'Okazaki sont liés par l'ADN ligase ATP-dépendante (Lig1) [Figure 13A (5)].



**Figure 13. (A) Les grandes étapes successives de la réplication de l'ADN et la comparaison des acteurs protéiques associés à ces étapes dans les trois règnes du Vivant, schéma adapté de Ishino and Ishino, 2012. (B) L'organisation des acteurs protéiques à la fourche de réplication chez les Archaea, d'après Lindàs and Bernander, 2013.**

## **II. Implication des ADN polymérases répliquatives archéennes dans la maintenance génomique s'effectuant à haute température**

### **II.1. La famille B chez les Archaea**

De un à trois membres de la famille B ont été identifiés chez les Archaea à l'heure actuelle. Les Crenarchaea possèdent plusieurs pols de la famille B avec des rôles distincts bien identifiés dans la réplication du chromosome. Ces protéines ont des similarités dans leur séquence en acides aminés, dans leur structure tridimensionnelle et possèdent toutes une activité de relecture 3' → 5' exonucléase. Tout comme les Pols B des organismes eucaryotes et procaryotes, les propriétés biochimiques de l'enzyme [liaison à l'ADN,  $K_m$  pour les dNTPs, processivité, sensibilité à des inhibiteurs comme l'antibiotique aphidicoline (inhibiteur historiquement spécifique de Pol  $\alpha$  de la famille B)], varie considérablement d'une espèce à l'autre (Perler et al., 1996 ; Kelman, 2000). Les différences de conditions environnementales dans lesquelles se développent ces organismes pourraient en partie expliquer ces caractéristiques biochimiques (Chaban et al., 2006).

Beaucoup de membre de Pol B d'Euryarchaea comportent entre une et trois intéines (Perler et al., 1996 ; Ishino and Cann, 1998 ; MacNeill, 2001). Elles se situent toutes dans des régions hautement conservées au sein de l'unité catalytique polymérase. Un autre caractère, actuellement identifié uniquement chez Pol B de *M. thermotrophicus*, concerne la sous-unité catalytique qui est séparée en deux polypeptides, ainsi, l'enzyme n'est active que lorsque les deux sous-unités sont associées (Kelman et al., 1999).

Comme exposé dans la deuxième partie de ce chapitre, les ADN polymérases sont indispensables à la réplication mais ont également un rôle dans la recombinaison et la réparation de l'ADN. Chez les Archaea, la question se pose de savoir si Pol B est la pol répliquative. Chez les Crenarchaea, cela semble être le cas car seules des pols de famille B ont été identifiées, même si certaines espèces possèdent également une pol de la réparation de la famille Y, ce qui est le cas de *Sulfolobus solfataricus* (Boudsocq et al., 2001) et de *Methanosarcina acetivorans* (Lin et al., 2010). Chez les Euryarchaea, le rôle de Pol B comme pol répliquative a été suggéré notamment chez *Pyrococcus abyssi* (Henneke et al., 2005), toutefois, une étude génétique récente sur l'Euryarchaea *Thermococcus kodakaraensis*, a montré que la Pol B ne serait pas essentielle à la survie de l'organisme. La fonction de pol répliquative pour la Pol B archéenne peut être admise car :

i) toutes les répliquases identifiées chez les eucaryotes (Pol  $\alpha$ , Pol  $\delta$ , Pol  $\epsilon$ ) appartiennent à la famille B,

ii) une activité de relecture 3' → 5' a été détectée chez toutes les Pol B archéennes qui se trouve être aussi une propriété catalytique des Pols B bactériennes et eucaryotes, iii) chez les bactéries et les eucaryotes, l'activité polymérase des pols répliquatives est connue pour être stimulée par les protéines accessoires de la réplication tels que le PCNA/clamp et la protéine de chargement (RCF/clamp loader) (Hingorani and O'Donnell, 2000), ce qui est aussi le cas des Pols B archéennes.

Chez les Euryarchaea, en plus de Pol B une nouvelle famille dans le règne du Vivant, a été découverte, il s'agit d'une pol de la famille D. L'ensemble de ces observations appuie l'idée que chez les Euryarchaea, Pol B soit la seule enzyme répliquative et pourrait agir de concert avec Pol D à la fourche de réplication.

La structure tridimensionnelle des Pols B étudiées chez plusieurs Archaea ont révélé une architecture similaire. Six structures de Pol B issues d'organismes d'Euryarchaea ont été résolues chez *Desulfurococcus Tok* (Zhao et al., 1999), *Thermococcus gorgonarius* (Hopfner et al., 1999), *Pyrococcus kodakaraensis* (Hashimoto et al., 2001), *Thermococcus sp. 9°N-7* (Rodriguez et al., 2000), *Pyrococcus furiosus* (Kim et al., 2008) et *Pyrococcus abyssi* (Gouge et al., 2012). Leur topologie est également similaire à la structure de la pol gp43, pol de la famille B du bactériophage RB69 (Wang et al., 1997). Un domaine de liaison à l'ARN a été identifié dans la structure des Pols B archéennes et est également présent chez les pols répliquatives eucaryotes, Pol  $\epsilon$  et Pol  $\delta$  (Zhao et al. 1999).

## II.2. La famille D chez les Archaea

Pol D, quant à elle, est une ADN polymérase spécifique des Euryarchaea (Uemori et al., 1997, Ishino et al., 1998 ; Cann et al., 1998). D'un point de vue structural, Pol D est un hétérodimère composé d'une petite sous-unité nommée DP1 qui a un rôle de correction des erreurs (domaine exonucléase) et d'une grande sous-unité DP2 portant l'activité ADN polymérase. Alors que la séquence primaire de DP2 présente des similarités avec celles des autres ADN polymérases archéennes, celle de DP1, en revanche, montre des homologies de séquence avec les petites sous-unités des pols répliquatives eucaryotes telles que la sous-unité p70 (Pol  $\alpha$ ), la sous-unité Cdc27p (Pol  $\delta$ ) et la sous-unité p55 (Pol  $\epsilon$ ) (Makiniemi et al., 1999) (Yamasaki et al., 2010). De ce fait, Pol D pourrait être l'homologue fonctionnel archéen de pol  $\epsilon$  (Tahirov et al., 2009). DP2, isolée de DP1, a une faible activité catalytique polymérase alors que l'association en hétérodimère stimule l'activité polymérase et l'activité 3'  $\rightarrow$  5' exonucléase (Uemori et al., 1997 ; Ishino et al., 1998 ; Gueguen et al., 2001 ; Jokela et al., 2004). Ces observations suggèrent que Pol D soit une pol répliquative chez les Euryarchaea. De plus, la stimulation de cette pol par le PCNA et le RFC (Cann et al., 2001 ; Henneke et al., 2002) conforte cette hypothèse. Plus particulièrement, des études biochimiques ont montré que la Pol D de *P. abyssi* se liait au PCNA par l'intermédiaire de deux motifs PIP (Castrec et al., 2009). Henneke ses collaborateurs ont montrés que Pol B et Pol D étaient nécessaires à la réplication de l'ADN de *Pyrococcus abyssi* (Henneke et al., 2005).

Aucune structure de l'ensemble de l'hétérodimère n'a été obtenue à ce jour, néanmoins, la structure du domaine N-terminal de la grande sous-unité DP2 et la structure d'une partie de la sous-unité DP1 ont été résolues (Matsui et al., 2011). Une partie de DP1 semble disposer d'un repliement similaire à celui des AAA+ ATPases et pourrait lier le PCNA (Tang et al., 2004). D'autre part, des études suggèrent que Pol D participe aux mécanismes de la recombinaison et de la réparation de l'ADN en raison de son interaction démontrée avec un homologue archéen de Rad51 (Hayashi et al., 1999).

Des études génétiques d'extinction de gène chez *Halobacterium* indiquent que les deux ADN polymérase Pol B et Pol D seraient essentielles à la viabilité de l'organisme (Berquist et al., 2007). Tandis que chez les bactéries une seule pol, Pol III, se dimérise à la fourche de réplication pour mener à bien la synthèse des brins avancé et retardé (Kelman and O'Donnell, 1995), chez les Archaea, deux pols de famille distinctes, Pol B et Pol D agissent de concert à la fourche de réplication (Rouillon et al., 2007), à l'image des Eukarya, où Pol  $\epsilon$  est en charge de la réplication du brin avancé et pol  $\delta$  du brin retardé (Burgers, 2009). En revanche, des études récentes utilisant des outils de génétique chez *Thermococcus kodakarensis* ont montré que seule Pol D, et non Pol B, serait essentielle à la survie de l'organisme (Cubonovà 2013).

### II.3. Les ADN polymérase archéennes face aux dommages à l'ADN

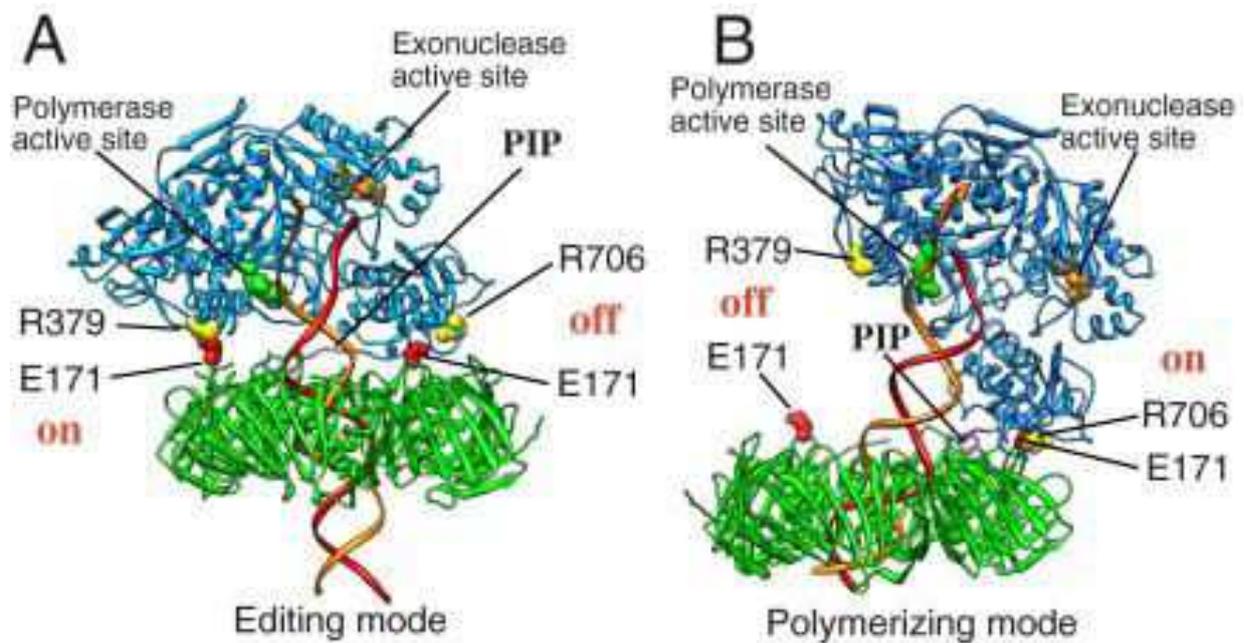
Dans la deuxième partie de cette étude bibliographique nous avons vu que Pol B et Pol D sont impliquées dans la reconnaissance des bases désaminées, lésion de l'ADN supposée accélérée à haute température (Lindahl and Karlström, 1973 ; Lindahl and Nyberg, 1974). *In vitro*, en condition d'excès molaire, la synthèse translésionnelle est possible lorsque Pol B rencontre une base uracile (Gouge et al., 2012) et également lorsqu'elle est confrontée à un site apurinique, avec une incorporation préférentielle d'un dAMP en face de cette lésion (Palud et al., 2008) ; respectant la règle du A (Matray and Kool, 1999 ; Zahn et al., 2007 ; Obeid et al., 2010). Ainsi, ces deux ADN polymérase présentent une activité translésionnelle similaire à celle des pols de la famille Y présentes uniquement chez les Crenarchaea.

Outre la propriété de reconnaissance très spécialisée des bases endommagées désaminées par les deux familles de pols présentes chez les Euryarchaea, à ce jour, *in vivo*, aucune pol de la réparation de l'ADN n'a été décrite chez les Archaea. Cependant, des similarités de séquences et de structures entre la pol de la réparation humaine (Pol  $\beta$ ) et la primase de *P. furiosus* (Augustin et al., 2001) suggèrent que la primase pourrait avoir un rôle dans la réparation de l'ADN (Le Breton et al., 2007). De façon intéressante, une nouvelle famille d'ADN polymérase a été identifiée sur un plasmide *Pyrococcus sp* Pp12-1 ; six gènes, dont des homologues codés par le virus PAV1 de *P. abyssi* (Geslin et al., 2003 ; Geslin et al., 2007), ont été décrits, cette étude révèle que les plasmides et les virus sont de véritables réservoirs de nouvelles protéines impliquées dans la réplication de l'ADN (Soler et al., 2010). Il serait intéressant d'étudier le comportement de ces pols face aux dommages à l'ADN.

L'ensemble de ces données suggère que soit d'autres ADN polymérase sont impliquées dans la réplication/réparation de l'ADN, soit que d'autres mécanismes spécialisés doivent être impliqués mais sont, à ce jour méconnus, ou encore que les mécanismes de réparation de l'ADN soient très actifs (constitutifs) chez les Archaea hyperthermophiles.

### **III. Les ADN polymérases en interaction avec les acteurs protéiques du réplisome**

Le « proliferating cell nuclear antigen » (PCNA) également appelé « clamp » est une protéine qui accompagne le déplacement de l'ADN polymérase le long du brin d'ADN en cours de réplication et peut favoriser sa processivité qui, par définition, est le nombre de nucléotides incorporés par unité de temps au cours d'un cycle d'attachement de la pol à l'ADN. Le facteur de processivité PCNA *Pyrococcus abyssi* se lie à Pol B par le peptide d'interaction au PCNA (motif PIP) et à Pol D par deux motifs PIP et augmente l'efficacité d'extension respective des deux polys (Castrec et al., 2009). Cependant, des études ont souligné que le PCNA n'a pas toujours une fonction d'adjuvant de la processivité. En effet, tout récemment, un nouveau motif PIP a été identifié dans le domaine C-terminal de la Pol B1 de *Sulfolobus solfataricus*. Il serait essentiel à la formation du complexe Pol-PCNA mais aussi, et principalement, à son attachement continu à la pol tout au long de la synthèse d'ADN (Bauer et al., 2013). De plus, l'association éventuelle des polys en trimère, comme c'est le cas pour la Pol B1 de *S. solfataricus*, contribue fortement à la processivité intrinsèque de l'enzyme (Mikheikin et al., 2009). La structure du complexe pol-PCNA-ADN de *Pyrococcus furiosus* a été analysée par microscopie électronique en conformation exonucléase ou polymérase. Les résultats structuraux des interactions entre les protéines suggèrent que le PCNA aurait un rôle dans la transition structurale du mode polymérase au mode exonucléase (Nishida et al., 2009 ; Mayanagi et al., 2011) (Figure 14). De plus, au niveau fonctionnel, lorsque le complexe Pol B-PCNA rencontre une base désaminée (uracile ou hypoxanthine), la présence du PCNA inhiberait l'action des enzymes de la réparation UDG et endonucléase V (Emptage et al., 2008). De la même façon, chez l'humain, PCNA aurait un rôle dans la régulation de l'activité mutagène de l'ADN polymérase translésionnelle (Nelson et al., 1996 ; Pavlov et al., 2001 ; Guo et al., 2001 ; Northam et al., 2006).



**Figure 14. Le mécanisme de transition structurale du mode exonucléase (« editing ») vers le mode polymérase du complexe Pol B-PCNA-ADN chez *Pyrococcus furiosus* et les interactions protéiques mises en jeu. (A) Le complexe en mode exonucléase. Pol B et PCNA sont respectivement en bleu et vert. Les brins d'ADN matrice et amorce sont représentés respectivement en orange et rouge. Les deux acides aminés E171 de la zone d'interaction du PCNA sont représentés par des sphères rouges. Les acides aminés de la pol, R706, dans le domaine « pouce » et R379, dans le domaine « paume » sont représentés par des sphères jaunes. Le motif DIE, dans le site actif d'exonucléase, et le motif DTDG, dans le site actif de l'ADN polymérase sont représentés respectivement par des sphères orange et verte. Alors que E171 est détaché de R706, la zone d'interaction proche du motif PIP-box impliquant R171, est en contact avec le résidu R379 de la Pol B et permet de stabiliser la configuration fermée du complexe. (B) Le complexe en mode polymérisation. Les zones d'interactions entre le PCNA et Pol B sont inversées par rapport à celles du mode d'édition, d'après Mayanagi et al., 2011).**

Une étude concernant la caractérisation de l'interaction *in vitro* et *in vivo* entre Pol Y et les protéines de la reconnaissance de l'origine de replication Orc1/cdc6 chez *Sulfolobus solfataricus*, a mis en évidence l'existence d'un mécanisme d'interaction à la fois physique et fonctionnel. En effet, lorsque l'ADN est sain, Orc1/cdc6 forme un complexe avec Pol Y stimule la liaison à l'ADN et coopère avec PCNA dans l'activité de polymérisation. En revanche, lorsque l'ADN comporte un mésappariement, Orc1/cdc6 inhibe la liaison à l'ADN et l'activité de polymérisation de Pol Y (Yang et al 2010).

#### **IV. Thermostabilité des ADN polymérases d'archées et intérêt biotechnologique**

En raison des conditions de vie des Archaea hyperthermophiles, la question de l'effet de la température sur les protéines se pose. De façon générale, les protéines issues d'organismes vivants à haute température sont dites thermostables en raison de l'adaptation même de leur structure primaire, secondaire, tertiaire et de leur repliement dans l'espace. Par rapport aux protéines dites mésophiles, une diminution du nombre d'acides aminés thermolabiles (alanine, histidine, glutamine, thréonine) a été rapportée chez les organismes thermophiles (Jaenicke et., al 2001). Par exemple, la Pol B de *P.*

*furiosus* comporte quatre cystéines qui établissent des ponts disulfures et ont un rôle dans la propriété de thermostabilité (Killelea and Connolly, 2011). D'un point de vue structure secondaire, les thermophiles ont une plus grande fraction de leur acides aminés organisés en hélices  $\alpha$  (Kumar et al., 2000). Les liaisons ioniques entre les acides aminés sont également plus nombreuses et participent au compactage global de la protéine et donc à la réduction de l'exposition des surfaces (Jaenicke et al., 1991 ; Jaenicke et al., 1996). De plus, l'augmentation de la densité de charge induit une réduction de la taille des cavités lors du repliement de la protéine. De même que la limitation de l'accès aux extrémités N-terminale et C-terminale augmenterait aussi la stabilité chez les Archaea (Hennig et al 1995). Une étude de thermostabilité des pols de *P. abyssi* a montré que la Pol B conservait 100 % de son activité enzymatique après un chauffage de 2h à 90°C ; à 100°C, son activité est diminuée de 30% (Gueguen et al., 2011).

Cette propriété de thermostabilité rend les Pols B issues d'archées hyperthermophiles particulièrement intéressantes pour leur utilisation en technique de réaction de polymérisation en chaîne (PCR). La PCR a été mise au point en 1986 par Kary Mullis (Mullis et al., 1986) et est, depuis lors, l'outil premier de la biologie moléculaire dans de nombreux domaines scientifiques. C'est une technique *in vitro* qui permet d'amplifier de façon exponentielle des fragments d'ADN. La première polymérase thermostable, utilisée dans cette technique, fut Pol I (*Taq* polymérase) de la bactérie *Thermus aquaticus* (Saiki et al., 1988) et c'est la pol la plus communément utilisée encore aujourd'hui. Cependant, cette pol répliquative est dépourvue de fonction exonucléase et peut donc effectuer des erreurs de réplication et les cumuler. C'est la raison pour laquelle, au cours des 25 dernières années, l'utilisation des Pols B archéennes en PCR et leur commercialisation s'est développée. On peut citer l'exemple des Pols B de *Pyrococcus furiosus* (Lundberg et al., 1991), *Thermococcus fumicolans* (Cambon-Bonavita et al., 2000) et *Pyrococcus abyssi* (Dietrich et al., 2002). Ces pols sont hautement fidèles mais moins rapides que la *Taq* pol. Actuellement, grâce aux connaissances approfondies des rôles des acides aminés au sein des pols il est possible de moduler leur activité spécifique et de générer des nouvelles méthodes variantes de la PCR classique. Par exemple, des mutations sur des régions précises de *Pfu* ont permis de rendre la pol extrêmement peu fidèle et ainsi d'utiliser cette propriété pour effectuer des PCR « error-prone » pour de la mutagenèse aléatoire de protéines (Biles and Connolly, 2004). La méthode de mutagenèse aléatoire est une méthode utilisée dans le développement de protéines aux propriétés modifiées et dans l'étude structure-dynamisme-fonction des protéines. Car en effet, la relation structure-fonction est un concept très important dans l'étude des systèmes et notamment en Biologie. Il met en relation la forme d'une entité et l'activité qui en découle car la structure d'une protéine détermine sa fonction.



## **CHAPITRE II**

Etude biochimique comparée de l'ion  $Mg^{2+}$  et de l'ion  $Ca^{2+}$ , en tant que cofacteurs enzymatiques, sur les propriétés fonctionnelles de l'ADN polymérase B de *Pyrococcus abyssi*.

---

Présentation de l'étude

Article

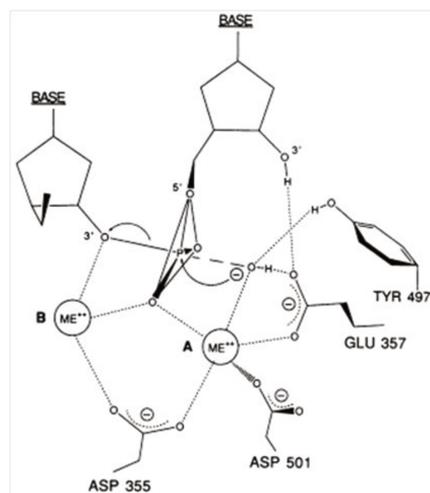
Résultats supplémentaires



# Première partie : Présentation de l'étude

Au cours de la réplication, les ADN polymérases utilisent un mécanisme catalytique identique qui consiste en l'élongation de l'ADN dans le sens 5' → 3'. Le site actif du domaine polymérase de ces enzymes comporte deux ions métalliques divalents, dont le plus couramment utilisé est le  $Mg^{2+}$ . Ce sont des cofacteurs essentiels à la synthèse du brin d'ADN complémentaire car ils interviennent dans la catalyse de la liaison phosphodiester entre le groupement 5'-phosphate d'un désoxyribonucléotide triphosphate (dNTP) et le groupement 3'-hydroxyle (OH) du brin d'ADN en cours de synthèse. Tandis que le métal A active le 3'-OH de l'amorce ADN pour l'attaque du phosphate  $\alpha$  dNTP, le métal B stabilise l'état de transition et chélate les phosphates  $\beta$  et  $\gamma$  ; un pyrophosphate ( $PP_i$ ) est alors libéré (Figure 7, page 41).

Le domaine exonucléase des ADN polymérases, responsable de la fonction de correction des erreurs d'insertion de base, possède également deux ions catalytiques. On parle d'activité exonucléasique 3' → 5'. Lors d'un processus de correction, dans le site exonucléase de la polymérase, une molécule d'eau coordonnée par l'ion métal A, forme un ion métal-hydroxide orienté de façon à attaquer le phosphate de la liaison à cliver. L'intermédiaire est alors stabilisé par la présence des deux ions catalytiques maintenus dans le site actif par trois acides aminés acides (D355, E357 et D501) dans le cas de l'ADN polymérase I d'*Escherichia coli* (Beese and Steitz, 1991).



**Figure 15. Mécanisme réactionnel de l'activité catalytique exonucléase.**

Le métal A tétrahédrique lie un des atomes d'oxygène de la liaison à cliver. Un hydroxyde (ou deux molécules d'eau) attaque(ent) le phosphate de la liaison. L'état de transition est stabilisé par les deux ions métalliques qui sont maintenus par les acides aminés acides (D355, E357, D501).

Comme nous l'avons vu dans le chapitre I de ce manuscrit, les ions catalytiques divalents sont au cœur des réactions enzymatique des pols. La nature chimique des ions métalliques utilisés par les polymérase *in vivo* n'est pas simple à déterminer. En effet, les structures cristallographiques sont obtenues dans des solutions tampons synthétiques, elles contiennent du  $Mg^{2+}$ , du  $Zn^{2+}$  ou encore du  $Mn^{2+}$  dans les sites actifs des pol. La structure du complexe ternaire cristallographique (Pol/ADN/dTTP) de l'ADN polymérase du bactériophage RB69 a été obtenue avec deux ions  $Ca^{2+}$  au sein du site actif polymérase et du site actif exonucléase (Franklin et al., 2001).

Dans les études *in vitro* de mesures d'activité polymérasique des pols, le cofacteur exclusivement utilisé est le  $Mg^{2+}$  sous forme de chlorure de magnésium, de sulfate de magnésium ou encore d'acétate de magnésium. En ce qui concerne les études de l'activité exonucléase, l'ion métallique divalent manganèse ( $Mn^{2+}$ ) est aussi utilisé (Pisani and Rossi, 1993). Ainsi, de façon surprenante, il s'est avéré que PabPolB était capable d'utiliser l'ion calcium en tant que cofacteur pour mener à bien les réactions d'extension d'amorce. Ce mécanisme est rare, et à notre connaissance, n'a été décrit jusqu'à présent que chez une Pol translésionnelle de la famille Y de l'Archaea *Sulfolobus solfataricus* (Irimia et al., 2006).

Afin de mieux comprendre le mécanisme réactionnel mis en jeu lorsque l'activité catalytique est menée par le cofacteur  $Ca^{2+}$ , nous avons comparé les propriétés intrinsèques de PabPolB, *in vitro* en présence de  $Mg^{2+}$  ou de  $Ca^{2+}$  comme cofacteur. Ainsi, des expériences de caractérisation enzymatique de l'ADN polymérase B de *P. abyssi* ont été menées dans le but de :

- Définir la concentration optimale en ion métallique divalent en tant que cofacteur lors de la réaction d'extension d'amorce (activité polymérase) et de dégradation d'amorce (activité exonucléase).
- Comparer la rapidité de polymérisation en présence de  $Mg^{2+}$  ou de  $Ca^{2+}$  par des expériences de cinétique d'extension d'amorce.
- Définir les paramètres de la cinétique enzymatique à savoir, l'affinité pour chaque dNTP et la sélectivité pour les dNTPs, l'efficacité catalytique d'incorporation de chaque dNTP.
- Etudier la flexibilité des sites catalytiques pol et exo par des expériences d'échange d'ions.
- Comparer l'utilisation du  $Ca^{2+}$  comme cofacteur par d'autres familles d'ADN polymérase des trois domaines du Vivant.

# Deuxième partie: Article en préparation

---



## Unexpected DNA synthesis with $\text{Ca}^{2+}$ as cofactor by archaeal Family B DNA polymerase.

Céline Ralec<sup>1,2,3</sup>, Sébastien Laurent<sup>1,2,3</sup> and Ghislaine Henneke<sup>1,2,3\*</sup>

In preparation

<sup>1</sup>Ifremer, UMR 6197, Laboratoire de Microbiologie des Environnements Extrêmes, BP 70, 29280 Plouzané, France.

<sup>2</sup>Université de Bretagne Occidentale, UMR 6197, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France.

<sup>3</sup>Centre National de la Recherche Scientifique, UMR 6197, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France.

\*To whom correspondance should be addressed. Tel : +33 298224609 ; Fax: +33 298224757; Email: Ghislaine.Henneke@ifremer.fr

Keywords : DNA replication, DNA polymerase, Archaea, divalent metal ion, cofactor.

### ABSTRACT

Family B DNA polymerase, (PabPolB) is one of the two major DNA polymerases in the hyperthermophilic Euryarchaeon *Pyrococcus abyssi*. High fidelity DNA replication is ensured by the coordinated polymerization and exonuclease proofreading. Like all biochemical processes, both polymerase and exonuclease activities require divalent metal ion cofactors to be effective,  $\text{Mg}^{2+}$  being the most common. The consequence of replacing  $\text{Mg}^{2+}$  by  $\text{Ca}^{2+}$  as a cofactor for family B DNA polymerase have not been investigated to date. Here we show that not only  $\text{Mg}^{2+}$  but also  $\text{Ca}^{2+}$  can be used by PabPolB to carry out polymerization on primed DNA. However, no exonuclease activity occurs in the presence of  $\text{Ca}^{2+}$ . To explore the ionic requirements for PabPolB,  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  are employed as catalytic metal ions in primer extension assays in order to compare the polymerization efficiency and the selective dNTPs incorporation. Then, the ionic accessibility of polymerase and exonuclease catalytic sites are examined in ion displacement assays. The results clearly show that the velocity is reduced with  $\text{Ca}^{2+}$  compared to  $\text{Mg}^{2+}$ , highlighting a novel cofactor feature. Moreover, we observe that higher amounts of heparin are required to trap PabPolB in a  $\text{Ca}^{2+}$ -dependent polymerizing mode, likely indicating that PabPolB affinity for the DNA template is more pronounced in the presence of  $\text{Ca}^{2+}$  compare to  $\text{Mg}^{2+}$ . Finally, dNTPs incorporation does not vary with  $\text{Ca}^{2+}$ , suggesting that PabPolB selectivity is likely comparable to  $\text{Mg}^{2+}$ . All these observations suggest that the catalytic mechanism with  $\text{Ca}^{2+}$  is totally different than  $\text{Mg}^{2+}$ . Moreover, we show that  $\text{Ca}^{2+}$  utilization as a catalytic ion by other DNA polymerase families seems to be exclusive to the family B DNA polymerase.

## INTRODUCTION

PabPolB is a family-B DNA polymerase from an Archaeon organism named *Pyrococcus abyssi* (*P. abyssi*). This characterized hyperthermophilic and anaerobic organism (Erauso et al., 1993) was isolated from marine hydrothermal fluids at 2 000 m depth, Fiji bassin (Auzende et al., 1991). The ability of *P. abyssi* to strictly colonize hot marine environment confers enzyme hyperthermostability (Dietrich et al., 2002). Particularly, PabPolB remains enzymatically active at 90°C (Gueguen et al., 2001) and sustains temperature variations, makes it useful for PCR reactions (ISIS commercial version, MP Biomedicals), in addition to the high level efficiency and accuracy of DNA copying generally held by the family B DNA polymerase.

Alone, PabPolB is able to synthesize long length DNA fragments (7-8 kb). However it cannot extend and displaces RNA primers. Strand displacement of DNA fragments is fully enhanced with the proliferating cell nuclear antigen (PCNA) (Henneke et al., 2005). In 2005, Henneke *et al* proposed that at the replication fork of *P. abyssi*, PabPolB deals with PabPolD to carry out accurate DNA replication. It has been proposed that PabPolB, which restrictively extend DNA primer, synthesizes the leading strand while PabPolD acts on the lagging strand. PabPolB interacts with PCNA via the universally conserved consensus sequence PCNA-interacting protein (PIP) box and is able to stimulate the loading of PCNA (Rouillon et al., 2007). That ability of *P. abyssi* to survive in very high temperature environments can be attributed, in part, to the ability of PabPolB to tolerate one of the most common DNA lesion, the endogenous abasic apurinic/aprimidinic (AP) site (Palud et al., 2008). In addition, the so-called read-head function to detect promutagenic uracil bases, firstly identified in the family-B DNA polymerase in the euryarchaeon *Pyrococcus furiosus* (*Pfu*) (Greagg et al., 1999) through its uracil recognition pocket, (Fogg et al., 2002 ; Shuttleworth et al., 2004 ; Firbank et al., 2008) has been recently found in PabPolB. Due to the fact that in PabPolB, uracil as well as canonical and non canonical bases can enter the uracil recognition pocket, we redefined it as the base checking pocket (Gouge et al., 2012).

PabPolB, as all DNA replicative polymerases, catalyzes the addition of mononucleotide units derived from deoxynucleoside 5'-triphosphates (dNTP) to the 3'-hydroxyl terminus of a primer chain in a reaction that requires a DNA template which directs the enzyme in its selection of the specific incoming nucleotide (Kornberg and Baker, 1991). Replicative family B DNA polymerases during DNA replication have dual activities. One is the polymerase phosphodiester bond creation which is responsible for DNA elongation. It occurs in the polymerase active site located in the « palm » domain. The other one is the reverse reaction consisting of DNA degradation. It takes place in the exonuclease active site located in the exonuclease domain that is conserved among replicative pol. It is known to confer high fidelity to DNA polymerase (McCulloch and Kunkel, 2008).

Divalent metal ions are essential for both polymerase and exonuclease activities (Lehman and Richardson, 1964 ; Kornberg, 1980). Their chelation by EDTA or their elimination by mutagenesis result in the loss of dNMP into the solution structure of the exonuclease domain (Derbyshire et al, 1988) therefore abolishing the hydrolytic phosphoryl transfer reaction. The proposed two metal ion

mechanism for the exonuclease activity of the Klenow fragment consists of metal ion A attacking the hydroxide ion with one electron pair oriented towards the phosphorus for the cleavage of the phosphodiester bond. Metal B is hypothesized to facilitate the leaving of the 3'-hydroxyl group and to stabilize the transition state. Similarly, in the polymerase active site, two divalent metal ions are also required. Whereas the metal A reduces the pKa of the 3'-hydroxyl group of the growing primer and attacks the phosphorous of the dNTP, metal B stabilizes the transition state (Beese and Steitz, 1991 ; Brautigam and Steitz, 1998). The common strategy for DNA synthesis is evidenced by the similar geometry of the bound DNA, nucleotide, and metal ions. Several triplex crystal structures of DNA polymerases with DNA and nucleotide substrate were solved. All these structures feature a conserved active-site geometry that includes a kink in the 5'-end of the template strand and, most strikingly, a large conformational change in the fingers for substrate binding.

To date, only one study reports the use of  $\text{Ca}^{2+}$  by archaeal family Y DNA polymerase (Dpo4) to perform DNA synthesis with very low efficiency (Irimia et al., 2006). Here we show that PabPolB is able to use  $\text{Ca}^{2+}$  as cofactor to perform DNA polymerization but not DNA degradation. The detailed examination of the functional intrinsic properties of PabPolB with  $\text{Ca}^{2+}$  suggests that the catalytic mechanism is totally different than  $\text{Mg}^{2+}$ . Interestingly,  $\text{Ca}^{2+}$  usage for polymerase activity seems to be restricted to family B DNA polymerases.

## MATERIAL AND METHODS

### Enzymes and DNA substrates

Wild-type (PabPolB exo+) and exonuclease-deficient (PabPolB exo -) (D215A) (Evans et al., 2000) were produced and purified as previously described (Gouge et al., 2012). PabPol D was produced and purified as previously described (Palud et al., 2008). Bacterial recombinant clone for production of family B DNA polymerase from the mesophilic crenarchaeon *Methanosarcina acetivorans* (MacPolBI) was kindly gifted by Isaac K. O. Cann, and purification of the MacPolBI was done as previously described (Robbins et al., 2004). Family X mesophilic DNA polymerase  $\beta$  from human was kindly provided by Ulrich Hübscher. Family Y DNA polymerase Dpo4 from the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* was purchased from Enzymax, family B DNA polymerase *Pfu* from the hyperthermophilic euryarchaeon *Pyrococcus furiosus* was purchased from Promega, family B DNA *KOD HiFi* polymerase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* was purchased from Novagen and family B DNA polymerase from the mesophilic bacteriophage T4 was purchased from New England Biolabs.

All oligonucleotides were purchased from Eurogentec (Seraing, Belgium). The unlabelled and 5'-Cy5 indocarbocyanin<sup>®</sup> labelled oligonucleotides were respectively purified by RP-cartridge and RP-HPLC. The sequences used are dependent on the experiment and are indicated in Figures legends. The oligonucleotide templates were annealed to 5'-end labelled oligonucleotide primers in the presence of 10 mM Tris-HCl, pH 8.0, and 50 mM NaCl at a 1:1 molar ratio by heating to 95 °C for 3 min and cooling to room temperature. M13ssc (single-stranded circular M13mp18 DNA) purchased from New

England Biolabs was annealed to the 32-nt oligonucleotide at a 1:3 molar ratio in 10 mM Tris–HCl, pH 8.0, and 50 mM NaCl, by heating at 75 °C for 15 min followed by cooling to room temperature.

### Primer extension assays

Primer extension reaction of the fluorescent-labeled 17-mer primer 5'-Cy5-TGCCAAGCTTGCATGCC-3' annealed to the 87-mer template 5'-CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGG TACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGC-3' (25 nM) were carried out in final volume (10 µl) containing 50 mM Tris (pH 8.8), 1 mM DTT, 10 mM KCl, indicated MgCl<sub>2</sub>, CaCl<sub>2</sub>, SrCl<sub>2</sub> or BaCl<sub>2</sub> concentrations, as indicated and 200 µM each of the dNTPs. Polymerizations were initiated by adding DNA polymerases at various amounts. 75 nM for PabPolB exo+ /exo -, MacPolBI, Dpo4 and Pol β ; 148 nM for PabPolD ; 0.75U for T4 and Pfu and 0.06U for KOD HiFi. Reactions were carried out at 37°C for mesophilic DNA polymerases and at 55°C for thermophilic DNA polymerases at the indicated times. Reactions were stopped by adding 15 µl of stop buffer [84% formamide, 10 mM EDTA, 0.01 N NaOH, and 1 µM of competitor oligonucleotide (an exact complement of the template)] followed by heating at 95°C for 8 min. Products were detected using 8 M urea denaturing polyacrylamide (17%) gel electrophoresis followed by visualization with a Mode Imager Typhoon 9400.

Primer extension reaction of the fluorescent-labeled 32-mer primer 5'-Cy5-TGCCAAGCTTGCATGCC TGCAGGTCGACTCTA-3' annealed to the M13mp18 template (7 nM) were performed in final volume (10 µl) containing 50 mM Tris (pH 8.8), 1 mM DTT, 10 mM KCl, indicated MgCl<sub>2</sub> or CaCl<sub>2</sub> concentrations. When required P/T was preincubated with PabPCNA for 8 min at 60°C. Then, reactions were initiated with 200 nM or 16 nM of PabPolB exo+ /exo - and 200 µM each of the dNTPs at 60°C at the indicated times. Reactions were quenched by addition of a 2-fold excess of EDTA (over divalent ion cofactor) and 40% formamide. Products were denatured for 5 min at 100°C, resolved on 1% alkaline agarose gel and visualized with a Mode Imager Typhoon 9400.

### PabPolB exonuclease assays

Primer degradation reaction of the fluorescent-labeled 17-mer primer 5'-Cy5-TGCCAAGCTTGCATGCC-3' annealed to the 87-mer template 5'-CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTC GGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGC-3' were carried out as described above in the absence of dNTPs.

### Steady-state kinetic analysis

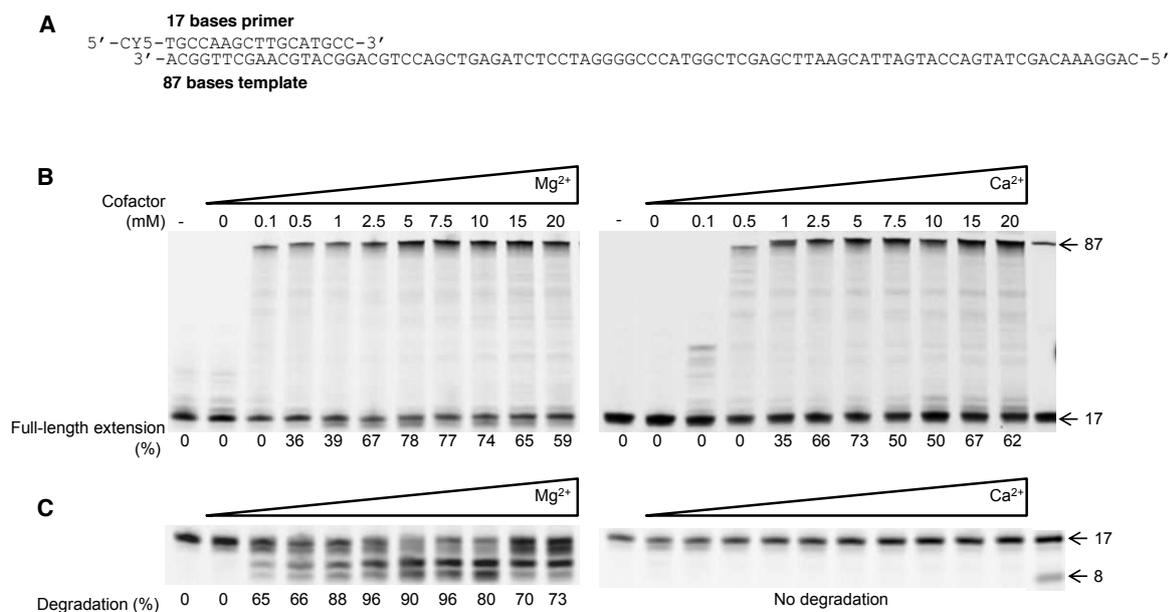
Reactions were performed as described for primer extension assays except that 5 nM polymerase and fluorescent-labeled 26-mer primer 5'-Cy5-TGCCAAGCTATGCCTGCAGGTCG-3' annealed to the 34-mer template 5'-GGATCCTGCGACCTGCAGTGCAAGCTTGGCA-3' were used. Single dNTP incorporation opposite G base is measured with either dCTP or dTTP at increasing concentrations as shown. Time points and pol concentrations were chosen in which primer extension was < 20% and the ratio of extended primers and unextended primers was plotted against time to give a linear plot of product formation. The band intensities of extended primers (+1 nt) and unextended primers were

quantified using the ImageQuantTL 8.0 software. The initial velocities of reactions were calculated as the slopes of the lines for each concentration of dNTP, plotted against dNTP concentration and fitted to a square hyperbola ( $V = V_{\max}[\text{dNTP}]/(K_m + [\text{dNTP}])$ ) with GraphPadPrim 5.2 software. Apparent  $K_m$  and  $V_{\max}$  obtained from the fit were used to calculate the efficiency of deoxyribonucleotide incorporation.

## RESULTS

### Effect of calcium concentration on catalytic activities of PabPoIB

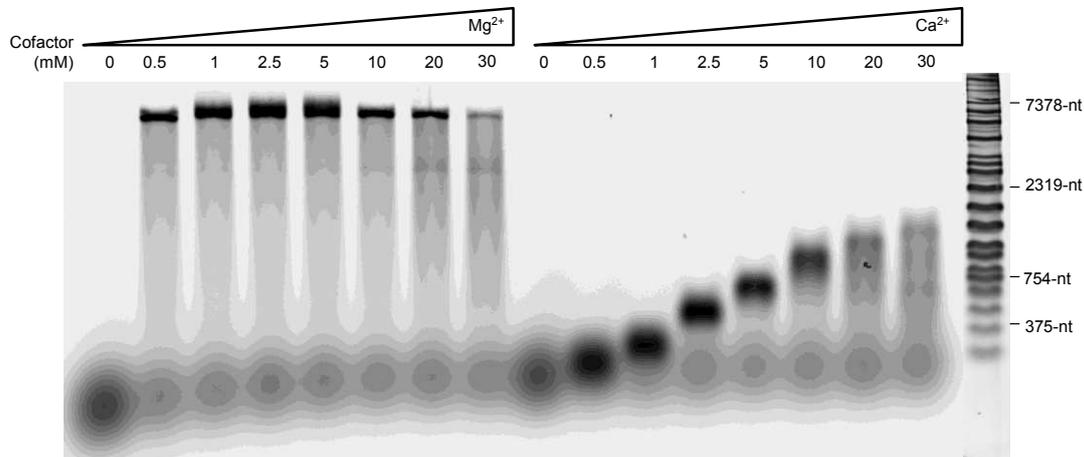
$\text{Ca}^{2+}$  is known to be an inhibitor of DNA polymerase. Figure 1 shows that without the addition of a divalent metal cation, no extension occurs, indicating that cofactors are essential for DNA synthesis.  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions were included in primer extension reactions with short primer-template (P/T) oligonucleotides as shown Figure 1A. PabPoIB has an optimum polymerization activity at 5 mM for both divalent metal ions (Figure 1B). However, exonuclease activity is deficient in the presence of  $\text{Ca}^{2+}$  (Figure 1C). Based on these observations, we decided to explore overall PabPoIB intrinsic properties with  $\text{Ca}^{2+}$  and to systematically compare to  $\text{Mg}^{2+}$ .



**Figure 1.** Catalytic activities of PabPoIB in the presence of increasing  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  concentrations.

The DNA polymerase (75 nM) was added to the primer-template (A) with dNTPs (200  $\mu\text{M}$ ) (B) or without dNTPs (C) and the extension (B) or the degradation (C) of the Cy5-labelled DNA P/T is obtained at increasing  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  concentrations at 55°C for 30 min. First lane of each gel is the reaction without enzyme. Total percentage under each gel represents either extended (B) or degraded (C) products. Quantification of the extension is defined as the ratio of extended full length products to total products. Reference oligonucleotides of 8, 17 and 87 bases are indicated by arrows.

When the influence of  $Mg^{2+}$  concentrations is examined in primer extension assays with long-length DNA P/T (32-nt/M13mp18), DNA polymerization works best at about 5 mM. In contrast, optimal concentration is never obtained with  $Ca^{2+}$  (Figure 2).

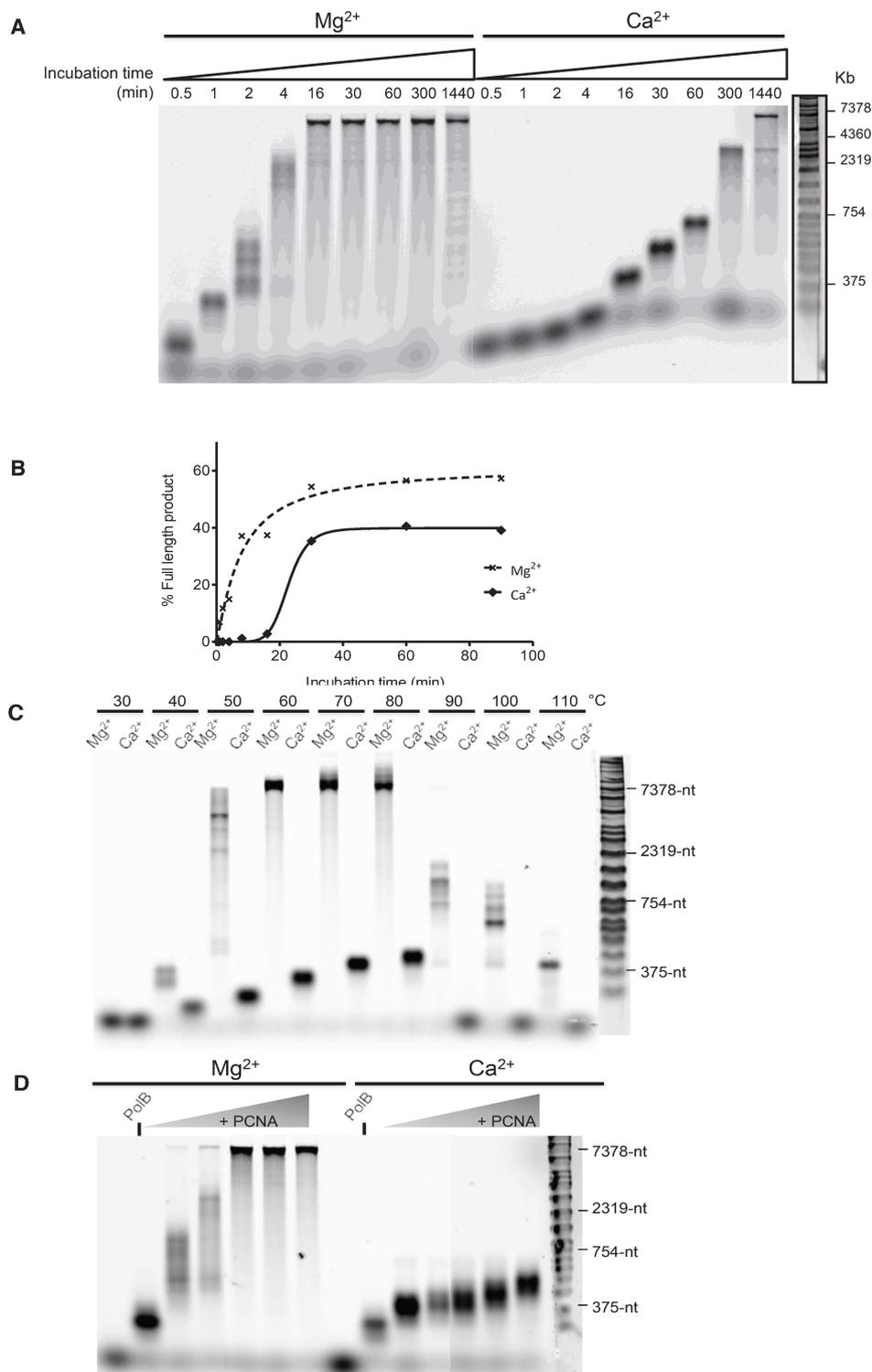


**Figure 2.** DNA synthesis on primed M13mp18 DNA template by PabPolB in the presence of increasing  $Mg^{2+}$  or  $Ca^{2+}$  concentrations.

The DNA polymerase (200 nM) was added to the primer-template in the presence of dNTPs (200  $\mu$ M). Effect of increasing  $Mg^{2+}$  or  $Ca^{2+}$  concentrations on primer extension by PabpolB was measured as described in Materials and methods. First lane of each gel is the reaction without enzyme. Molecular weight markers are labeled on the right of the gel.

### Effect of $Ca^{2+}$ on full-length extension ability by PabPolB

We next assess the ability of PabPolB to perform full-length DNA synthesis in the presence of  $Ca^{2+}$  on either on a long or short P/T (Figure 3A and 3B). As shown in Figure 3A, full-length DNA synthesis is completed at about 16 minutes with  $Mg^{2+}$ , whereas 24 hours are required with  $Ca^{2+}$  to obtain full-length replicated M13mp18. This result demonstrates that the rate of DNA extension by PabPolB is clearly reduced with  $Ca^{2+}$  compared to  $Mg^{2+}$ . This observation holds true when primer-extension assay is carried out on short P/T as highlighted by the fitted values (Figure 3B).  $Ca^{2+}$  is not as suitable as  $Mg^{2+}$  but clearly is not an inhibitor of PabPolB polymerase activity.  $Ca^{2+}$  seems to induce a different DNA polymerization mechanism for PabPolB. We then searched for DNA extension reaction conditions which could improve  $Ca^{2+}$ -mediated rate extension. Interestingly, as shown in Figures 3C and 3D, the increase of reaction temperature in one hand, and the addition of PabPCNA on the other hand, respectively enhance primer extension activity by PabPolB with either  $Ca^{2+}$  or  $Mg^{2+}$ . As a result, replicated M13mp18 products were longer corresponding to the attachment of 200 extra bases on DNA fragments.



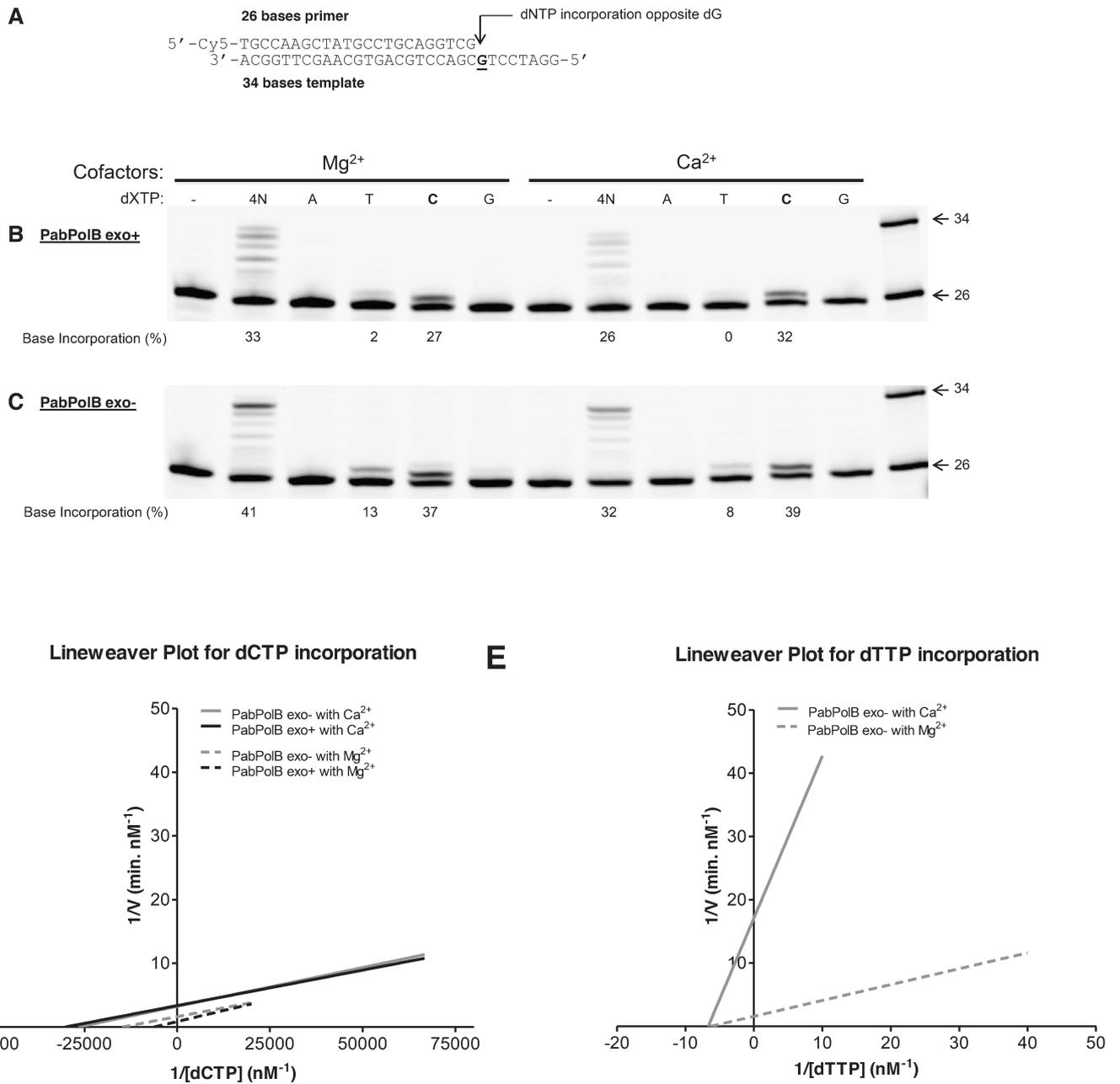
**Figure 3.** Time course experiments of polymerase activity by PabPolB on short and long primer/template (P/T).

**(A)** Primer extension assay by PabPolB (200 nM) with the 5'-end labeled (P/T) 32/M13mp18 (7 nM) in the presence of dNTPs (200  $\mu$ M) and either 5 mM  $Mg^{2+}$  or  $Ca^{2+}$ . Reactions were carried out at 60°C at the indicated incubation times. Molecular weight markers are indicated. **(B)** Primer extension assay by PabPolB (75 nM) with the 5'-end labeled (P/T) 17/87 (25 nM) in the presence of dNTPs (200  $\mu$ M) and either 5 mM  $Mg^{2+}$  or  $Ca^{2+}$ . Reactions were carried out at 55°C at the indicated incubation times. Percentages of full-length extension products were obtained using ImageQuantTL 8.0 software. Data were fitted using One site-specific binding will Hill Slope plot with GraphPadPrism software. **(C)** Primer extension assay by PabPolB (200 nM) is performed as in **(A)** for 30 min at the indicated temperatures. **(D)** Primer extension assay by PabPolB (16 nM) with the 5'-end labeled (P/T) 32/M13mp18 (7 nM). Preincubation was carried out with increasing PabPCNA concentrations (from 4 to 128 nM). Reactions were performed in the presence of dNTPs (200  $\mu$ M) and either 5 mM  $Mg^{2+}$  or  $Ca^{2+}$  for 30 min at 60°C.

### A higher dNTP affinity for PabPolB with $\text{Ca}^{2+}$

Initial primer extension assays using the labeled P/T 32/M13mp18 displayed reduced polymerase activity of PabPolB with  $\text{Ca}^{2+}$  compared to  $\text{Mg}^{2+}$ . This observation prompted us to investigate the single nucleotide incorporation of PabPolB *exo+* and PabPolB *exo-* in the presence of either  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$ . Figure 4A and B shows that dNTP incorporation is similar with the two divalent ions; as expected, for PabPolB *exo+*, only dCTP is incorporated opposite G. In contrast, the lacking proof-reading PabPolB (PabPolB *exo-*) incorporates dCTP opposite G but also dTTP with a lower efficiency. Besides, no significant difference were observed when  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  concentrations varied (2 mM, 5 mM and 10 mM) (Supplementary data, Figure S1 and Table S1). According to these results, the determination of the kinetic constants for single incorporation of dNTP was performed. Analysis of these data shows that the enzymatic reactions obey Michaelis-Menten kinetics (Figure 4D and E) (Supplementary data, Figure S2 and S3). The calculated values are summarized in Table 1. Interestingly,  $k_{\text{cat}}$  are lower with  $\text{Ca}^{2+}$  than with  $\text{Mg}^{2+}$  ( $0.07 \text{ min}^{-1}$  versus  $0.26 \text{ min}^{-1}$  for PabPolB *exo+* and  $0.066 \text{ min}^{-1}$  versus  $0.147 \text{ min}^{-1}$  for PabPolB *exo-*), suggesting that dCTP is slowly incorporated per protein with  $\text{Ca}^{2+}$ . This observation is consistent with the results obtained in Figure 3B, demonstrating a rate of DNA synthesis reduced in the presence of  $\text{Ca}^{2+}$ . The calculated  $K_m$  values for dCTP with  $\text{Mg}^{2+}$  ( $0.196 \mu\text{M}$  for PabPolB *exo+* and  $0.076 \mu\text{M}$  for PabPolB *exo-*) and with  $\text{Ca}^{2+}$  ( $0.042 \mu\text{M}$  for PabPolB *exo+* and  $0.033 \mu\text{M}$  for PabPolB *exo-*) show a higher affinity of PabPolBs for the correct nucleotide in the presence of  $\text{Ca}^{2+}$ .

The catalytic efficiency ( $k_{\text{cat}}/K_m$ ) for PabPolB *exo+* and PabPolB *exo-* did not vary significantly in the presence of  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$ ; the values were 1.2-fold higher for PabPolB *exo+* and 1-fold higher for PabPolB *exo-* when  $\text{Ca}^{2+}$  is used instead of  $\text{Mg}^{2+}$ . Under these steady-state conditions, the turnover of PabPolB *exo+* and PabPolB *exo-* is respectively 3.5-fold and 2-fold lower with  $\text{Ca}^{2+}$  compared to  $\text{Mg}^{2+}$ . Interestingly, the turnover of PabPolB *exo+* with  $\text{Ca}^{2+}$  is comparable to PabPolB *exo-*, probably because  $\text{Ca}^{2+}$  inhibits the proof-reading activity of the wild-type PabPolB as shown in Figure 1.



**Figure 4.** Single nucleotide incorporation by PabPolB<sub>exo+</sub> and PabPolB<sub>exo-</sub> in the presence of  $Mg^{2+}$  or  $Ca^{2+}$ .

The primer-template oligonucleotide sequences are shown and imply single dNTP incorporation opposite G (**A**). Extension reactions of the P/T (26/34) (25 nM) in the presence of a single or all four dNTPs (4N) (200  $\mu$ M) with either  $Mg^{2+}$  or  $Ca^{2+}$  (5 mM) were performed at 55°C for 30 min by PabPolB exo+ (5 nM) (**B**) and by PabPolB exo- (5 nM) (**C**). Percentages of dNTP incorporation are indicated under each gel. Reference oligodeoxynucleotides of 26 and 34 bases are indicated by the arrows.

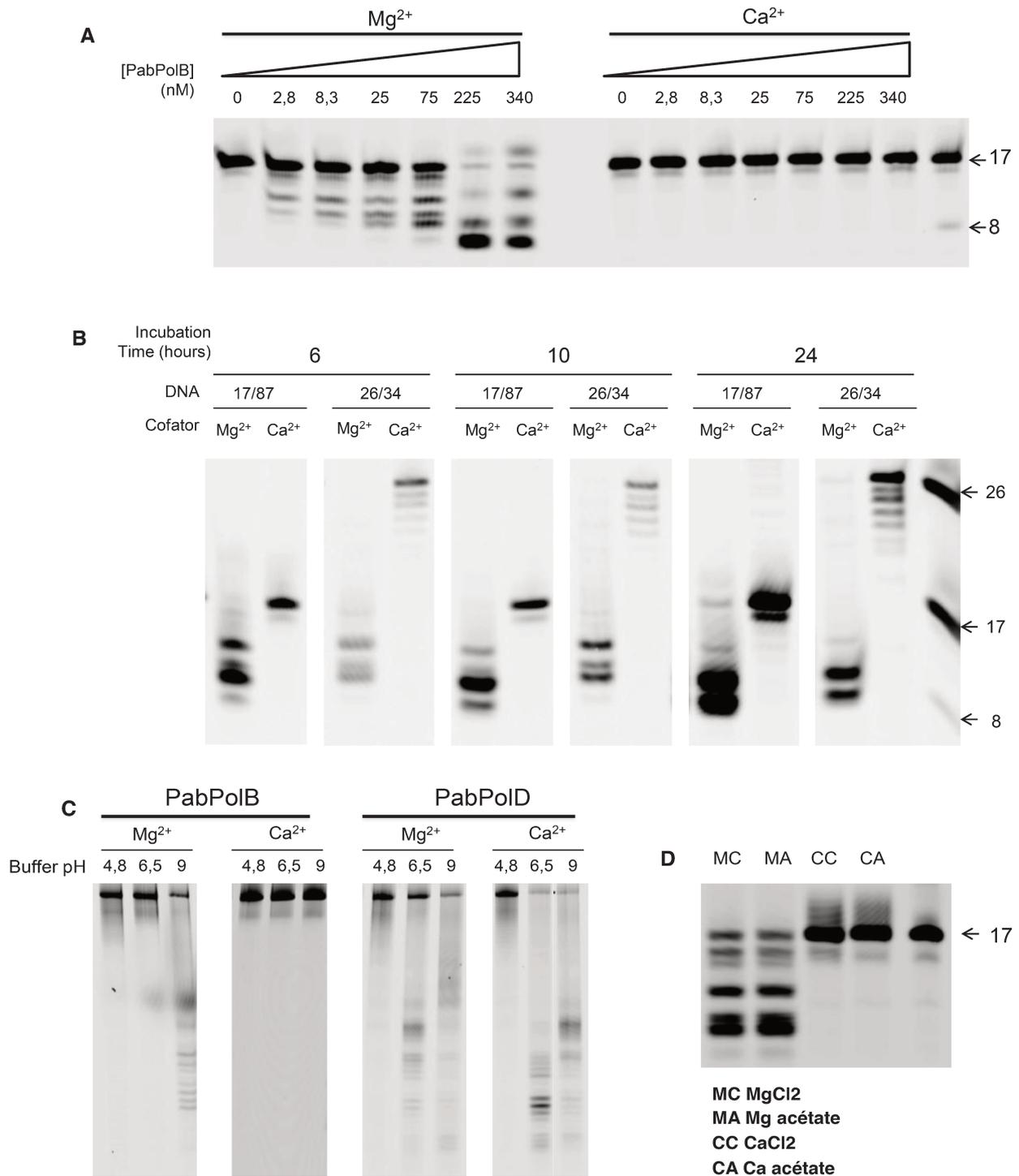
n	PabPoIB	Cofactor	Template : incoming dNTP	$V_{max}$ (nM/min)	$K_m$ ( $\mu$ M)	$k_{cat}$ ( $\text{min}^{-1}$ ) $\times 10^{-3}$	Catalytic efficiency $k_{cat}/K_m$  ( $\mu\text{M}^{-1} \text{min}^{-1}$ ) $\times 10^{-3}$	Efficiency relative to $\text{Mg}^{2+}$	f(misinsertion frequency)
7	Wild-type	$\text{Mg}^{2+}$	G : dC	1.305 $\pm$ 0.055	0.1963 $\pm$ 0.038	261.00	1329.6		
8	Wild-type	$\text{Ca}^{2+}$	G : dC	0.348 $\pm$ 0.019	0.0424 $\pm$ 0.007	69.58	1640.7	0.81 ( $\uparrow$ 1.2 x)	
17	Exo-	$\text{Mg}^{2+}$	G : dC	0.736 $\pm$ 0.025	0.0756 $\pm$ 0.015	147.12	1946.5		1
9	Exo-	$\text{Ca}^{2+}$	G : dC	0.331 $\pm$ 0.017	0.0327 $\pm$ 0.006	66.12	2024.5	0.96 ( $\uparrow$ 1 x)	1
5	Exo-	$\text{Mg}^{2+}$	G : dT	0.743 $\pm$ 0.082	180.9 $\pm$ 40.43	148.56	0.8212		0.0004
5	Exo-	$\text{Ca}^{2+}$	G : dT	0.111 $\pm$ 0.015	396.6 $\pm$ 133.33	22.26	0.0561	0.07 ( $\downarrow$ 15x)	0.00003

**Table 1.** Kinetic parameters of nucleotide incorporation opposite template G with  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  by PabPoIB exo+ and PabPoIB exo-.

For kinetic analysis, experimental data of the rate of incorporation as a function of dNTP concentration were fitted by nonlinear regression, using the program GraphPad Prism 4. The data were fitted to a Michaelis-Menten equation: velocity =  $V_{max} \cdot [\text{dNTP}] / (K_m + [\text{dNTP}])$ . Apparent  $K_m$  and  $V_{max}$  kinetic parameters obtained from the fit were used to calculate the efficiency of dNTP incorporation ( $k_{cat}/K_m$ ). The kinetic values are the average of 5 to 17 determinations and are showed with SD. The f(misinsertion frequency) for PabPoIB exo- is the ratio of efficiency for the incorrect nucleotide to the efficiency for the correct nucleotide.

### Ionic accessibility of polymerase and exonuclease active sites.

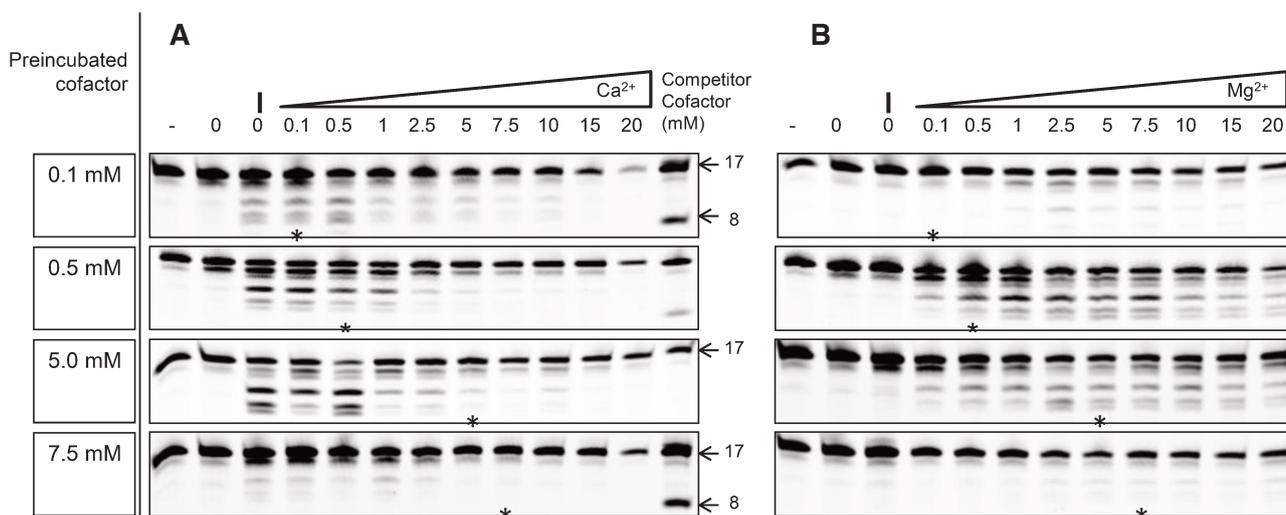
The ability of PabPoIB to degrade an oligonucleotide in the 3'→5' direction through its exonuclease domain in the presence of  $\text{Ca}^{2+}$  was investigated. No exonuclease activity was detected upon incubating with either P/T DNA (Figure 1C, right) or with increasing amounts of PabPoIB (Figure 5A). Moreover, longer incubation times (24 hours) and the use of various P/T oligonucleotide sequences (Figure 5B) did not trigger exonuclease activity. Modification of pH buffer (Figure 5C) or counterion of calcium (Figure 5D) had no effect either. Positive controls show that exonuclease activity of PabPoIB was proficient in all conditions tested with 5 mM  $\text{Mg}^{2+}$  optimal concentration (Figure 1C, left and Figure 5A-C). These observations support the conclusion that  $\text{Ca}^{2+}$  cannot be used as a catalytic metal ion for the exonuclease activity of PabPoIB.



**Figure 5.** Exonuclease activity of PabPolB *exo*<sup>+</sup> in various experimental conditions with Mg<sup>2+</sup> and Ca<sup>2+</sup>.

(A) Primer degradation assay by increasing concentrations of PabPolB *exo*<sup>+</sup> with the 5'-end labeled (P/T) 17/87 (25 nM) in the presence of either 5 mM Mg<sup>2+</sup> or Ca<sup>2+</sup>. Reactions were carried out at 55°C for 30 min. Molecular weight markers are indicated. (B) Primer degradation assay by PabPolB *exo*<sup>+</sup> (75 nM) with the 5'-end labeled (P/T) 17/87 or the (P/T) 26/34 in the presence of either 5 mM Mg<sup>2+</sup> or Ca<sup>2+</sup>. Reactions were carried out at 55°C at the indicated incubation times of 6, 10 and 24 hours. (C) Primer degradation assay by PabPolB *exo*<sup>+</sup> (75 nM) (left) and PabPolD *exo*<sup>+</sup> (148 nM) (right) with the 5'-end labeled single-stranded DNA (87-nt) (25 nM) in the presence of either 5 mM Mg<sup>2+</sup> or Ca<sup>2+</sup> in various pH buffers. Reactions were carried out at 55°C for 30 min. (D) Primer degradation assay by PabPolB (75 nM) with the 5'-end labeled (P/T) 17/87 (25 nM) in the presence of various 5 mM Mg<sup>2+</sup> or Ca<sup>2+</sup> counterions. Reactions were carried out at 55°C for 30 min.

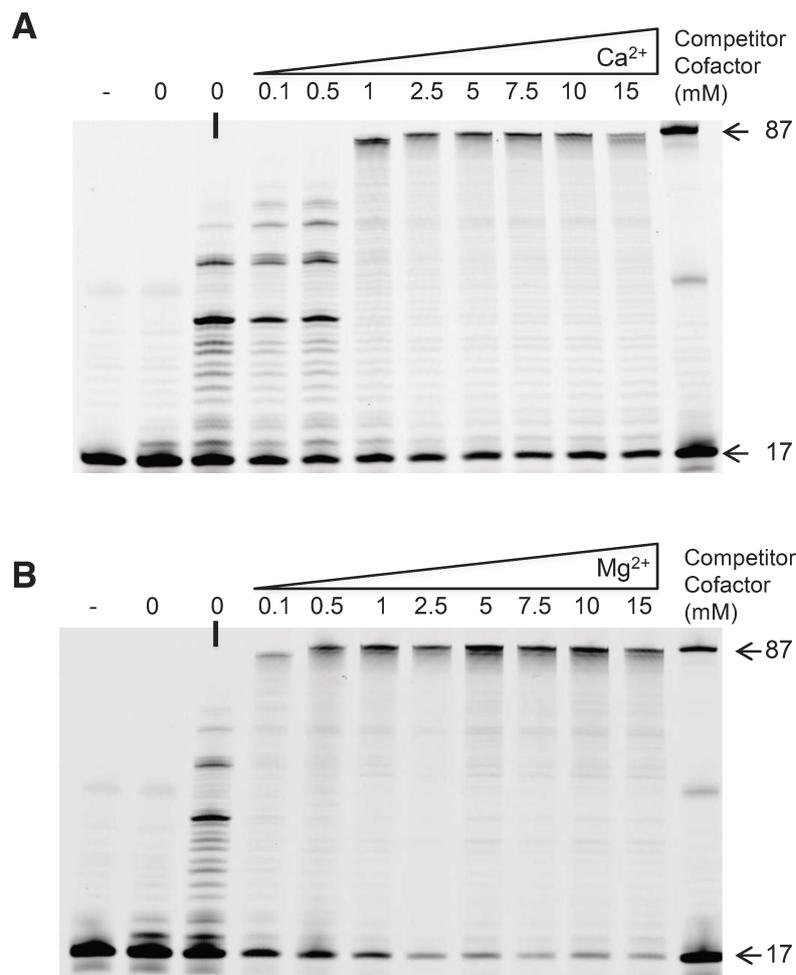
Since  $\text{Ca}^{2+}$  impairs PabPolB exonuclease activity, we asked questions about the mechanism involved in the inhibitory effect. For instance, is  $\text{Ca}^{2+}$  entering the exonuclease catalytic pocket ? To address this point, ion displacement experiments were performed. Reactions consisted of a preincubation of PabPolB  $\text{exo}^+$  with one metal ion followed by addition of increasing concentrations of the competitive metal ion. The replacement of  $\text{Mg}^{2+}$  metal ion by  $\text{Ca}^{2+}$  is followed by the suppression of exonuclease activity. Figure 6A shows that the increase of  $\text{Ca}^{2+}$  concentrations progressively inhibits  $\text{Mg}^{2+}$ -dependent primer degradation. This suggests that  $\text{Ca}^{2+}$  can easily enter the exonuclease pocket and displace preincubated  $\text{Mg}^{2+}$  ions. Different concentrations of preincubated  $\text{Mg}^{2+}$  ions were tested and showed that 0.1 to 20 mM  $\text{Ca}^{2+}$  can gradually dislodge preincubated  $\text{Mg}^{2+}$  ions with concentrations ranging from 0.1 to 7.5 mM (Figure 6A). Interestingly, the inhibitory effect happened exactly when  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  are used at equimolar concentrations (Figure 6A, see stars (\*) on the gels). These results seem to indicate that  $\text{Ca}^{2+}$  can enter the catalytic exonuclease pocket and becomes catalytically inactive. To give consistency, reverse reactions were performed (Figure 6B). When PabPolB $\text{exo}^+$  is preincubated with constant  $\text{Ca}^{2+}$  concentrations (0.1, 0.5, 5 and 7.5 mM) and  $\text{Mg}^{2+}$  is used as a metal ion competitor, PabPolB exonuclease activity is restored but limited to 5 mM  $\text{Ca}^{2+}$ . Indeed,  $\text{Ca}^{2+}$  displacement is not triggered when preincubated at 7.5 mM. Interestingly,  $\text{Ca}^{2+}$  dependent inhibition of exonuclease activity happens exactly when  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  are used at equimolar concentrations (Figure 6B, see stars (\*) on the gels). This result shows that  $\text{Ca}^{2+}$  displacement is not as efficient as  $\text{Mg}^{2+}$  and suggests that  $\text{Mg}^{2+}$  divalent metal ion is more labile than  $\text{Ca}^{2+}$  in the exonuclease active site.



**Figure 6.** Ion displacement experiment in exonuclease mode.

Since  $\text{Ca}^{2+}$  is shown to be an inhibitor for exonuclease activity, the replacement of a metal ion in the exonuclease active site can be followed by the decrease of exonuclease activity. Ion displacement experiment was carried out with constant metal ion concentrations and by increasing metal ion competitor concentrations in exonuclease assays as shown respectively on the left and the top of each gel. PabPolB $\text{exo}^+$  was preincubated at the indicated ion concentrations in PabPolB reaction buffer. Reaction displacement was initiated by coaddition of P/T (17/87) (25 nM) together with increased ion competitor concentrations and performed for 30 min at 55°C. (A) Displacement of  $\text{Mg}^{2+}$  by  $\text{Ca}^{2+}$  from the exonuclease active site (B) Displacement of  $\text{Ca}^{2+}$  by  $\text{Mg}^{2+}$  from the exonuclease active site. Second lane of each gel is the control with PabPolB without metal ion. Third lane corresponds to the basal activity of the enzyme at 0.1, 0.5, 5 and 7.5 mM  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$ . Reference oligodeoxynucleotides of 8 and 17 bases are indicated by the arrows.

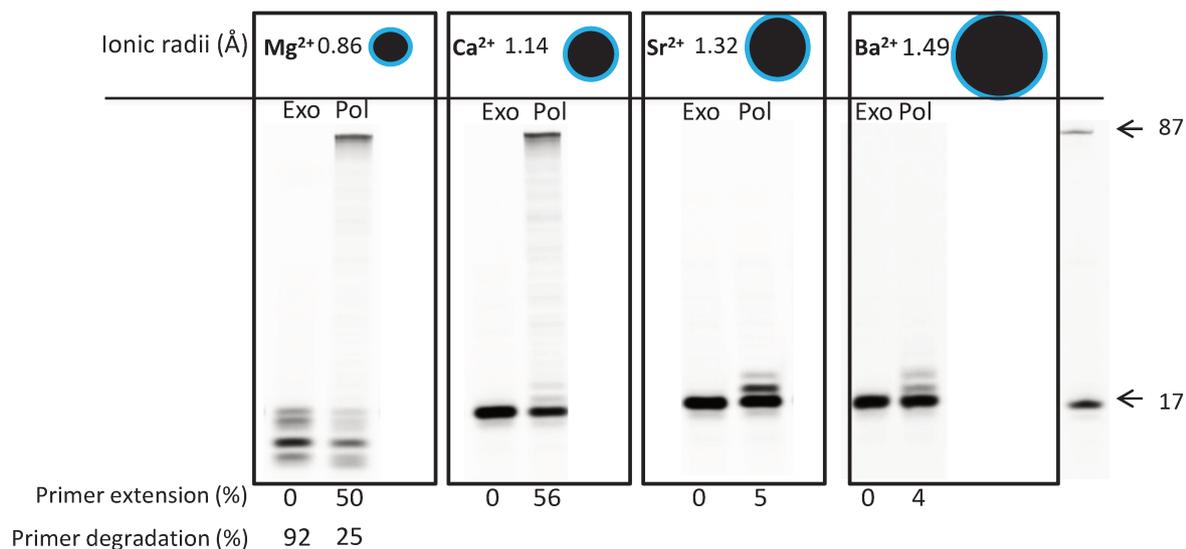
Since we gave evidence that  $\text{Ca}^{2+}$  can be the polymerase catalytic metal ion, we decided to explore the replacement of a metal ion in the polymerase active site. This phenomenon can be followed by an increase in polymerase activity. Ion displacement experiments were performed in primer extension assays with P/T DNA (17/87) by increasing competitor ion concentrations. PabPolBexo- is preincubated with either  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  constant metal ion concentrations (Figure 7A and B, respectively). DNA polymerase activity is progressively stimulated upon addition of  $\text{Ca}^{2+}$  to the  $\text{Mg}^{2+}$  preincubated enzyme and full-length synthesis is observed for 5 mM  $\text{Ca}^{2+}$  (Figure 7A). The reverse reaction shows that addition of  $\text{Mg}^{2+}$  to the  $\text{Ca}^{2+}$  preincubated enzyme enhances DNA synthesis and full-length is obtained at 0.5 mM  $\text{Mg}^{2+}$  (Figure 7B).



**Figure 7.** Ion displacement experiment in polymerase mode.

Since  $\text{Ca}^{2+}$  can be used as a cofactor for polymerase activity, the replacement of a metal ion in the polymerase active site can be followed by the increase in polymerase activity. Ion displacement experiment was carried out with constant metal ion concentrations and by increasing metal ion competitor concentrations in primer-extension assays. PabPolBexo- was preincubated with the indicated ion concentrations in PabPolB reaction buffer. Reaction displacement was initiated by coaddition of P/T (17/87) (25 nM) together with increased ion competitor concentrations and all four dNTP (200  $\mu\text{M}$ ) and performed for 30 min at 55°C. **(A)** Displacement of  $\text{Mg}^{2+}$  by  $\text{Ca}^{2+}$  from the polymerase active site **(B)** Displacement of  $\text{Ca}^{2+}$  by  $\text{Mg}^{2+}$  from the polymerase active site. Second lane of each gel is the control with PabPolB without metal ion. Third lane corresponds to the basal activity of the enzyme at 0.1 mM  $\text{Mg}^{2+}$  or 0.5 mM  $\text{Ca}^{2+}$ . Reference oligodeoxynucleotides of 17 and 87 bases are indicated by the arrows.

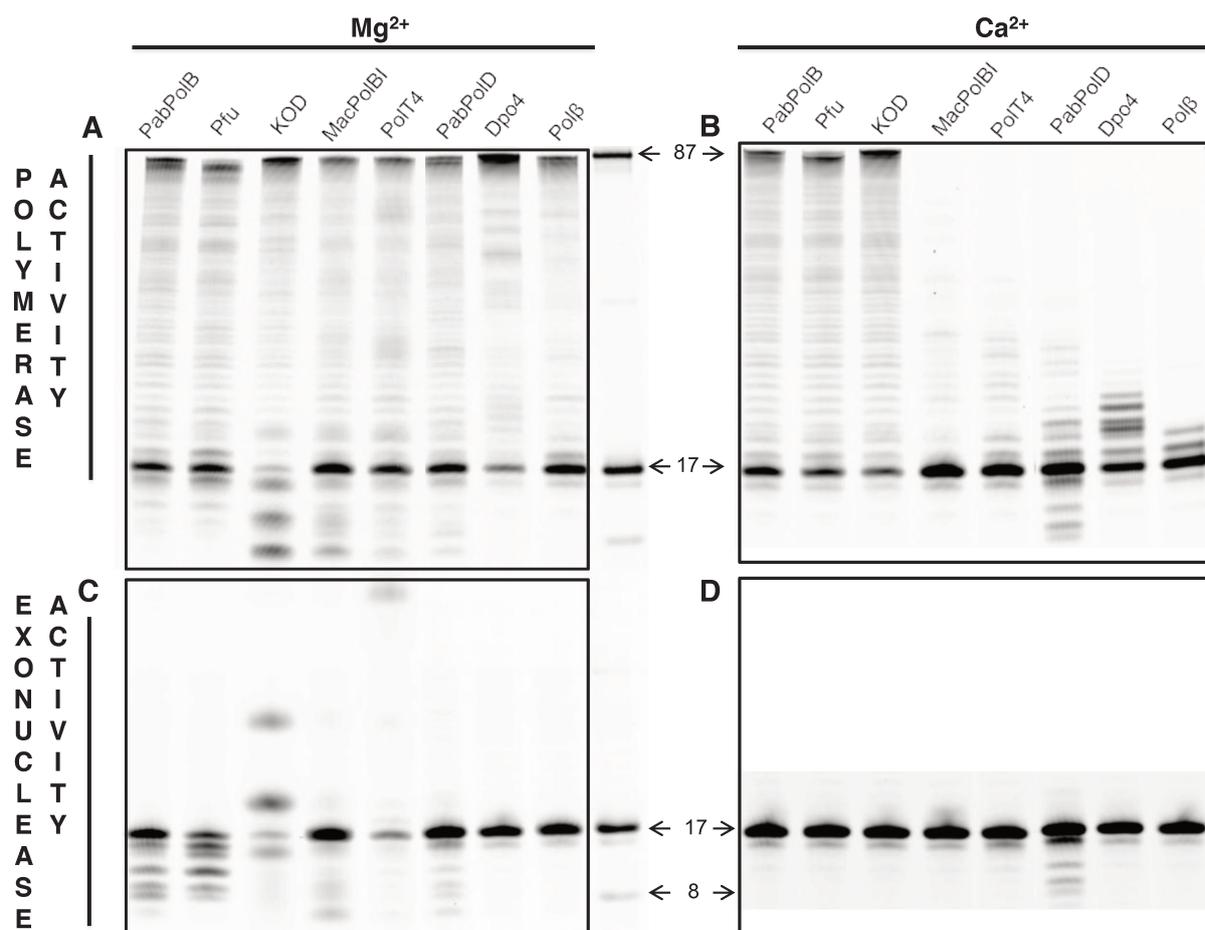
The results show that ion exchange can take place within the active sites of the polymerase and exonuclease domain but this was dependent on the preincubated metal ion concentrations. Thus we ask whether there was another parameter which could influence these ion exchanges and look at the size of the polymerase and exonuclease active sites. To address this question, divalent metal ions with increased ionic radii were assayed. All belonged to the alkaline earth group according to the periodic table of Mendeleev. The results indicate that above 1.14 Å ionic radius of metal ion DNA synthesis and exonuclease activity are inhibited (Figure 8).



**Figure 8.** Influence of the size of metal divalent ions on primer extension and degradation by PabPoIB *exo*<sup>+</sup>.

Reactions were performed at 55°C for 30 min with labeled P/T (17/87) (25 nM), PabPoIB *exo*<sup>+</sup> (75 nM) and 5 mM Mg<sup>2+</sup>, Ca<sup>2+</sup>, Sr<sup>2+</sup> or Br<sup>2+</sup>. Percentage of products formation is indicated under the gel. Reference oligodeoxynucleotides of 17 and 87 bases are indicated by the arrows.

Then we asked if Ca<sup>2+</sup> utilization was restricted to archaeal and/or to family B DNA polymerases. To answer that question various families of DNA polymerases from archaeal and non archaeal organisms were assayed in either polymerase or exonuclease reactions. As shown in Figure 9, only archaeal members of the family B DNA polymerases were able to perform full-length DNA synthesis with Ca<sup>2+</sup>. In contrast, PolT4, Dpo4 and PabPoID showed limited primer extension by adding few nucleotides and, in these conditions only PabPoID was capable to degrade primer DNA with Ca<sup>2+</sup>. Interestingly, this replicative DNA polymerase, restricted to euryarchaeal organisms, is composed of two distinct catalytic domains to achieve both polymerase and exonuclease activities. Thus, this unique structural organization could possibly facilitates Ca<sup>2+</sup> utilization in exonuclease activity.



**Figure 9.** Primer extension and degradation assays with  $Mg^{2+}$  or  $Ca^{2+}$  as a cofactor by different DNA polymerase families from the three domains of life.

Primer extension assays with P/T 17/87 (25 nM) in the presence of 5 mM  $Mg^{2+}$  (A) or 5 mM  $Ca^{2+}$  and dNTP (200  $\mu$ M)(B). Primer degradation assays with P/T 17/87 (25 nM) in the presence of 5 mM  $Mg^{2+}$  (C) or 5 mM  $Ca^{2+}$  (D) without dNTP. Reactions were carried out for 30 min at the indicated concentrations of various DNA polymerases (as indicated) at 37°C for PolBIMa (75 nM), PolT4 (0,75 U), Dpo4 (75 nM), Polβ (75 nM), and at 55°C for PabPolB (75 nM), Pfu (0,75 U), KOD (0,06 U) and PabPolD (148 nM). Reference oligodeoxynucleotides of 8 and 17 and 87 bases are indicated by the arrows.

### Acknowledgments

We thank Ulrich Hübscher for the gift of aliquots of human DNA polymerase β. We also thank Isaac K. O. Cann for the gift of bacterial expression plasmid of DNA polymerase BI from *Methanosarcina acetivorans*. This work was supported by Ifremer and the Région Bretagne.

## REFERENCES

- Auzende, J.-M., Urabe, T., Bendel, V., Deplus, C., Eissen, J.-P., Grimaud, D., Huchon, P., Ishibashi, J., Joshima, M., Lagabrielle, Y., et al. (1991). In situ geological and geochemical study of an active hydrothermal site on the North Fiji Basin ridge. *Marine Geology* *98*, 259–269.
- Beese, L.S., and Steitz, T.A. (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* *10*, 25–33.
- Brautigam, C.A., and Steitz, T.A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* *8*, 54–63.
- Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L., Friedman, J.M., Joyce, C.M., and Steitz, T.A. (1988). Genetic and crystallographic studies of the 3',5'-exonucleolytic site of DNA polymerase I. *Science* *240*, 199–201.
- Dietrich, J., Schmitt, P., Zieger, M., Preve, B., Rolland, J.-L., Chaabihi, H., and Gueguen, Y. (2002). PCR performance of the highly thermostable proof-reading B-type DNA polymerase from *Pyrococcus abyssi*. *FEMS Microbiol. Lett.* *217*, 89–94.
- Erauso, G., Reysenbach, A.-L., Godfroy, A., Meunier, J.-R., Crump, B., Partensky, F., Baross, J.A., Marteinson, V., Barbier, G., Pace, N.R., et al. (1993). *Pyrococcus abyssi* sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* *160*, 338–349.
- Evans, S.J., Fogg, M.J., Mamone, A., Davis, M., Pearl, L.H., and Connolly, B.A. (2000). Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* *28*, 1059–1066.
- Firbank, S.J., Wardle, J., Heslop, P., Lewis, R.J., and Connolly, B.A. (2008). Uracil recognition in archaeal DNA polymerases captured by X-ray crystallography. *J. Mol. Biol.* *381*, 529–539.
- Fogg, M.J., Pearl, L.H., and Connolly, B.A. (2002). Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat. Struct. Biol.* *9*, 922–927.
- Gouge, J., Ralec, C., Henneke, G., and Delarue, M. (2012). Molecular recognition of canonical and deaminated bases by *P. abyssi* family B DNA polymerase. *J. Mol. Biol.* *423*, 315–336.
- Greagg, M.A., Fogg, M.J., Panayotou, G., Evans, S.J., Connolly, B.A., and Pearl, L.H. (1999). A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc. Natl. Acad. Sci. U.S.A.* *96*, 9045–9050.
- Gueguen, Y., Rolland, J.L., Lecompte, O., Azam, P., Le Romancer, G., Flament, D., Raffin, J.P., and Dietrich, J. (2001). Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi*. *Eur. J. Biochem.* *268*, 5961–5969.
- Henneke, G., Flament, D., Hübscher, U., Querellou, J., and Raffin, J.-P. (2005). The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. *J. Mol. Biol.* *350*, 53–64.
- Irimia, A., Zang, H., Loukachevitch, L.V., Eoff, R.L., Guengerich, F.P., and Egli, M. (2006). Calcium is a cofactor of polymerization but inhibits pyrophosphorolysis by the *Sulfolobus solfataricus* DNA polymerase Dpo4. *Biochemistry* *45*, 5949–5956.
- Lehman, I.R., and Richardson, C.C. (1964). The deoxyribonucleases of *Escherichia coli* IV. An exonuclease activity present in purified preparations of deoxyribonucleic acid polymerase. *J. Biol. Chem.* *239*, 233–241.
- McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and

translesion synthesis polymerases. *Cell Res.* *18*, 148–161.

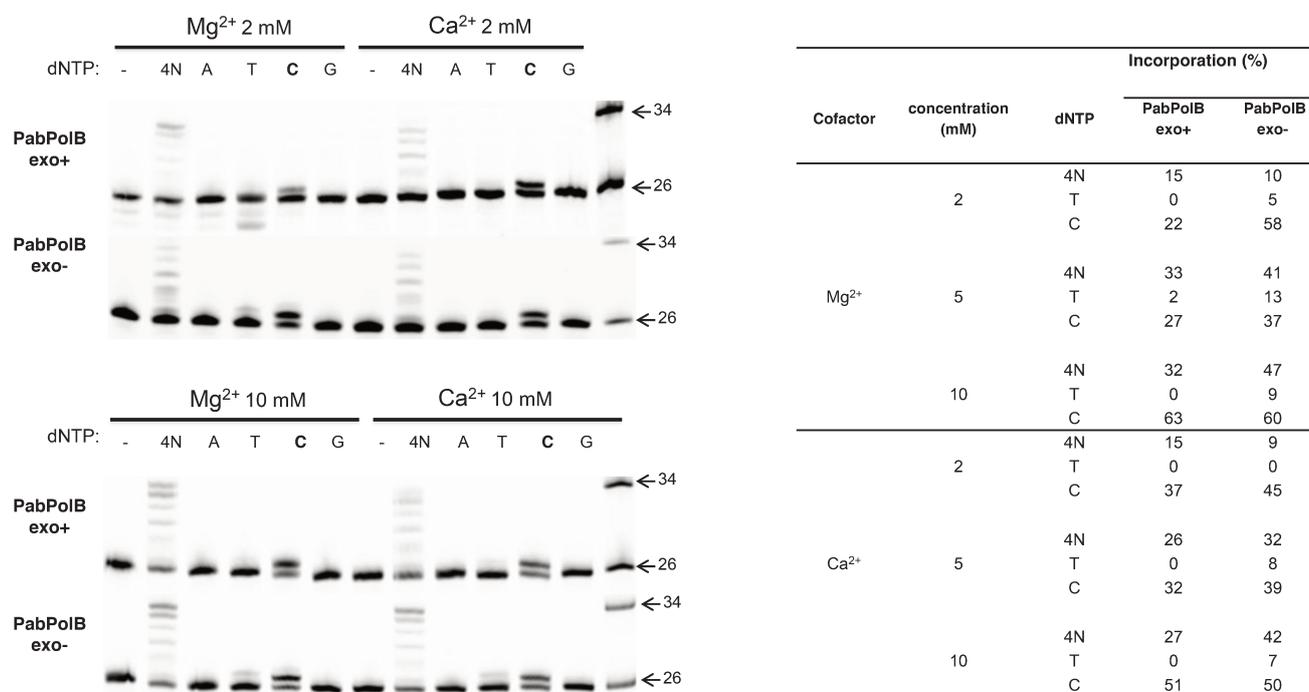
Palud, A., Villani, G., L'Haridon, S., Querellou, J., Raffin, J.-P., and Henneke, G. (2008). Intrinsic properties of the two replicative DNA polymerases of *Pyrococcus abyssi* in replicating abasic sites: possible role in DNA damage tolerance? *Mol. Microbiol.* *70*, 746–761.

Robbins, J.B., Murphy, M.C., White, B.A., Mackie, R.I., Ha, T., and Cann, I.K.O. (2004). Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase B1. *J. Biol. Chem.* *279*, 6315–6326.

Rouillon, C., Henneke, G., Flament, D., Querellou, J., and Raffin, J.-P. (2007). DNA polymerase switching on homotrimeric PCNA at the replication fork of the euryarchaea *Pyrococcus abyssi*. *J. Mol. Biol.* *369*, 343–355.

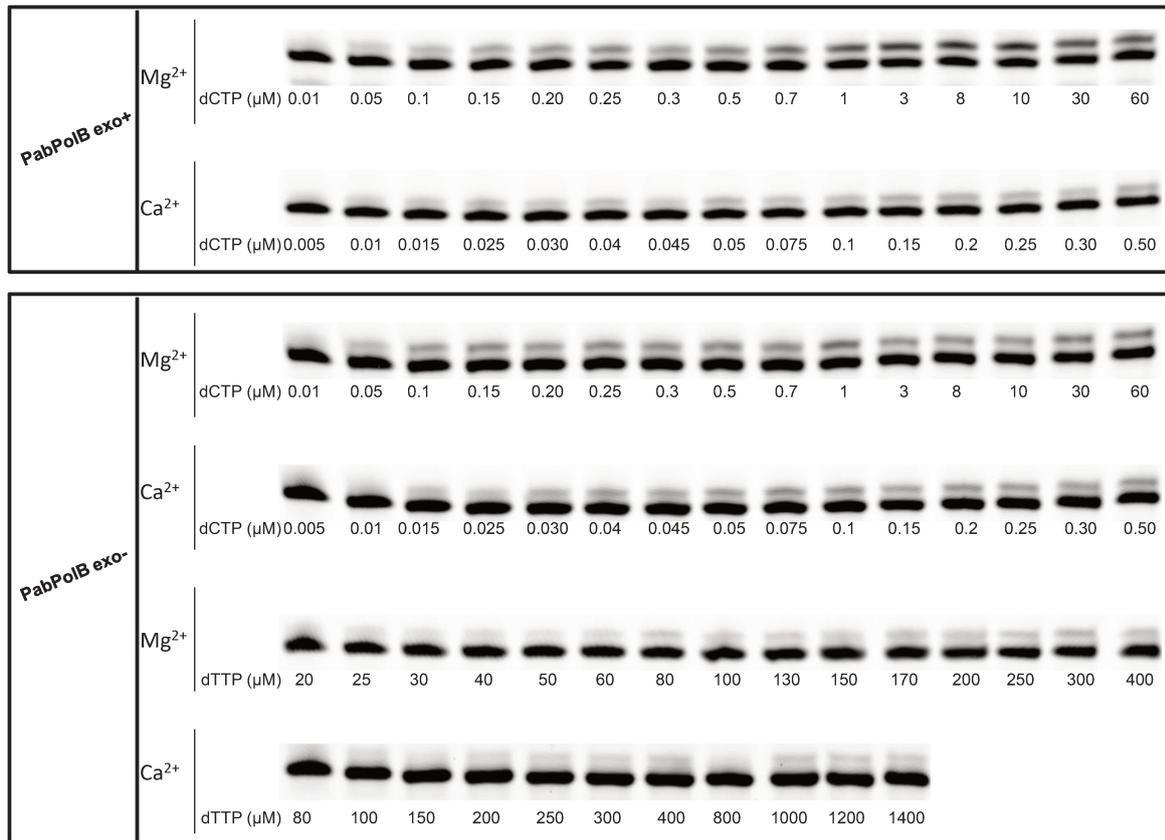
Shuttleworth, G., Fogg, M.J., Kurpiewski, M.R., Jen-Jacobson, L., and Connolly, B.A. (2004). Recognition of the pro-mutagenic base uracil by family B DNA polymerases from archaea. *J. Mol. Biol.* *337*, 621–634.

## SUPPLEMENTARY DATA

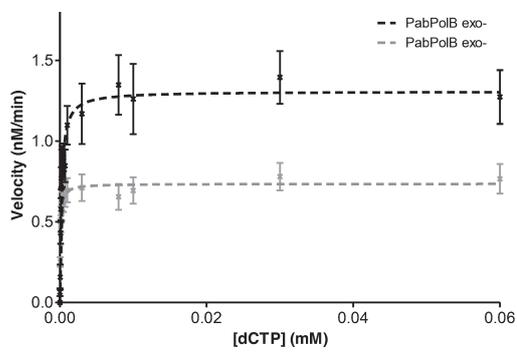


**Figure S1 and Table S1.** Influence of the concentration of metal ion on single nucleotide incorporation by PabPolB.

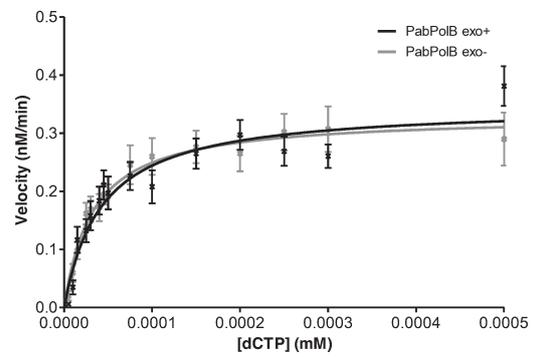
Primer extension assays were carried out at 55°C for 30 min in the presence of P/T (26/34) (25 nM), PabPolB exo+ (5 nM), and either all four dNTPs (4N) or a single dNTP (as indicated) (200 μM). Reactions performed with 2 mM Mg<sup>2+</sup> (A), 2 mM Ca<sup>2+</sup> (B), 10 mM Mg<sup>2+</sup> (C), 10 mM Ca<sup>2+</sup> (D). Reference oligodeoxynucleotides of 26 and 34 bases are indicated by the arrows. Percentages of nucleotides incorporation are indicated in Table S1.



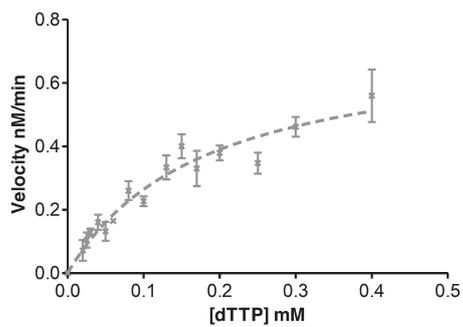
**A** Michaelis-Menten Plot for dCTP incorporation with Mg<sup>2+</sup>



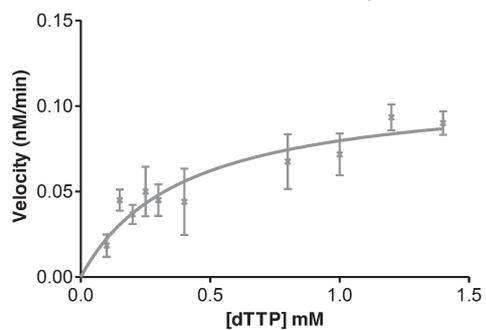
**B** Michaelis-Menten Plot for dCTP incorporation with Ca<sup>2+</sup>



**C** Michaelis-Menten Plot for dTTP incorporation with Mg<sup>2+</sup>



**D** Michaelis-Menten Plot for dTTP incorporation with Ca<sup>2+</sup>



**Figure S2 and Table S2.** Steady-state kinetics of PabPolB *exo+* and PabPolB *exo-* with either  $Mg^{2+}$  or  $Ca^{2+}$ .

Figure S2 shows representative gels of steady-state kinetic measurements. 5 nM of wild-type or *exo-* PabPolB were incubated with P/T (26/34) (25 nM) in the presence of dCTP or dTTP at various concentration ranges as indicated with either  $Mg^{2+}$  or  $Ca^{2+}$  (5 mM). Reactions were carried out at 55°C, for 4 min for dCTP incorporation with  $Mg^{2+}$ , 8 min for dCTP incorporation with  $Ca^{2+}$ , and 8 min for dTTP incorporations with  $Mg^{2+}$  or  $Ca^{2+}$ .

Table S2 shows Michaelis-Menten plots of PabpolB *exo+* and PabpolB *exo-* for dCTP incorporation with  $Mg^{2+}$  (**A**), PabpolB *exo+* and PabpolB *exo-* for dCTP incorporation with  $Ca^{2+}$  (**B**), PabpolB *exo-* for dTTP incorporation with  $Mg^{2+}$  (**C**), PabpolB *exo-* for dTTP incorporation with  $Ca^{2+}$  (**D**).

---

# Troisième partie : Résultats supplémentaires

## I. Matériels et méthodes supplémentaires

### **I.1 Amplification de l'ADN génomique par PCR**

Un fragment d'ADN génomique de *P. abyssi* de 500 kb (20 ng/μL) a été amplifié en présence d'amorce spécifiques reverses et forward (1 μM), de dNTPs (200 μM), dans le tampon d'incubation contenant 50 mM Tris HCl pH 8,8, 1 mM dithiothreitol, 10 mM KCl, 2 mM MgCl<sub>2</sub> et d'une gamme croissante de CaCl<sub>2</sub> dans un volume réactionnel final de 25 μL. Les réactions ont été initiées par l'ajout de 1 Unité d'ADN polymérase Isis™ (version commercialisée par MP Biomedicals de PabPolB exo+) et 30 cycles de réactions de polymérisation en chaînes ont été effectués en thermocycleur (GenAmp® PCR System 9700, Applied Biosystems) selon le programme suivant :

- Dénaturation initiale	(3 min – 94°C)	} X 30 cycles
- Dénaturation	(1 min – 94°C)	
- Hybridation des amorces	(1 min – 55°C)	
- Elongation	(2 min – 72°C)	
- Elongation finale	(5 min – 72°C)	

Les fragments amplifiés ont ensuite été analysés par électrophorèse en gel d'agarose 0,8 % contenant du bromure d'éthidium et visualisés sous UV avec l'Imager BioRad.

## II. Résultats supplémentaires

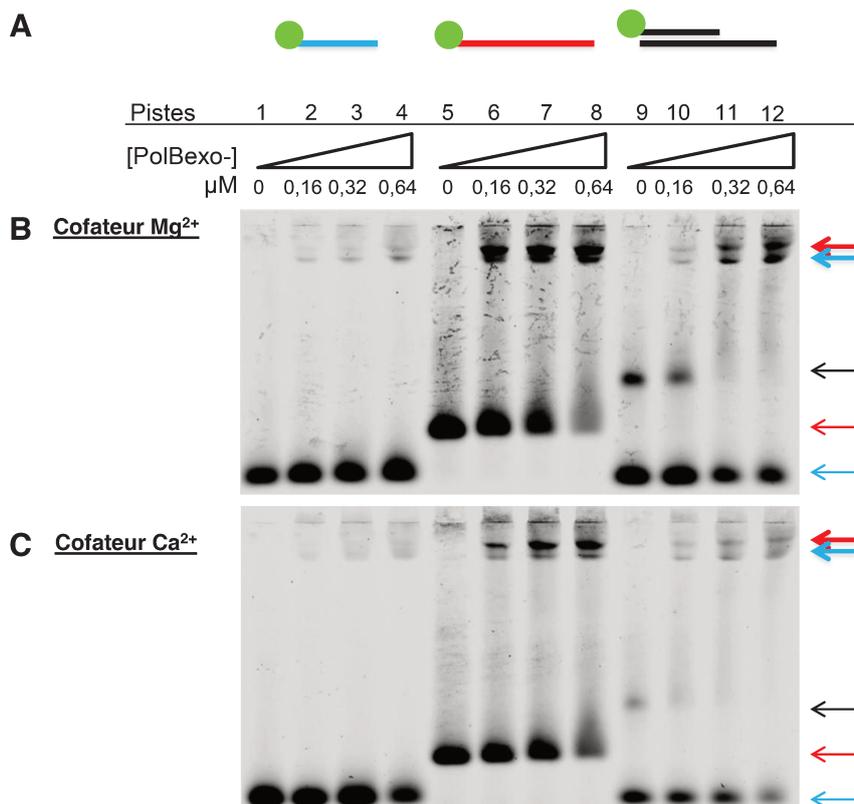
L'ensemble des résultats obtenus souligne que le Ca<sup>2+</sup> peut être utilisé comme cofacteur de PabPolB pour mener à bien la synthèse d'ADN. Toutefois, les expériences montrent que les propriétés intrinsèques de PabPolB sont quelques peu modifiées en présence de Ca<sup>2+</sup>. L'hypothèse d'un mécanisme réactionnel différent de celui induit par le Mg<sup>2+</sup> est alors posée. Les résultats exposés dans cette partie concernent l'impact de l'utilisation du Ca<sup>2+</sup> comme cofacteur sur les propriétés déjà bien identifiées de PabPolB. Notamment, les propriétés de liaison à l'ADN simple brin et double brin, et les propriétés de reconnaissance et de tolérance des dommages au cours de la synthèse d'ADN. Ces résultats sont très préliminaires.

## II.1. Effet du $\text{Ca}^{2+}$ en tant que cofacteur sur les propriétés intrinsèques de PabPolB

### II.1.1. Liaison à différent type de structures ADN

Ces mesures ont été réalisées par la technique de retard sur gel « EMSA » décrite dans l'article publié Gouge et al., 2012. Pour des raisons expérimentales du fait de l'absence de dNTP dans la réaction, et donc pour ne pas activer la potentielle dégradation des brins d'ADN lors de la réaction de liaison, la version *exo-* de PabPolB a été utilisée (mutant dont l'activité exonucléase est nulle). L'absence d'agent dénaturant dans le gel prévient la dénaturation de l'ADN et de la protéine, et conserve, au cours de la migration, les complexes formés lors de la réaction. Deux types de structure d'ADN ont été étudiés : de l'ADN simple brin et de l'ADN double brin amorcé, ainsi que deux longueurs différentes de simple brin (19 et 32 nucléotides). Une gamme de concentration de PabPolB *exo-* a été effectuée en présence de  $\text{Mg}^{2+}$  ou de  $\text{Ca}^{2+}$  afin de détecter une éventuelle modification d'efficacité de liaison à l'ADN. Dans la méthode « EMSA », une liaison de la protéine à l'ADN marqué se traduit par un retard à la migration sur gel nommé « shift » par rapport à l'ADN libre marqué. Sur les figures, les formes libres d'ADN sont indiquées par des flèches, les complexes ADN/PabPolB sont indiqués par des flèches en gras.

La première différence importante à noter est la capacité de liaison de PabPolB à des oligonucléotides simples brins, et que cette liaison est dépendante de la longueur de l'oligonucléotide. En effet, alors qu'un complexe (ADN/PabPolB) est visible à 0,16  $\mu\text{M}$  avec l'oligonucléotide de 32 nt (Figure 16B, piste 6) aucun complexe n'est détectable en présence de l'oligonucléotide de 19 nt aux concentrations en PabPolB testées (Figure 16B, piste 1 à 4). Ce résultat est identique en présence du cofacteur  $\text{Mg}^{2+}$  ou  $\text{Ca}^{2+}$  (comparer Figures 16B et C). La liaison de PabPolB au duplexe amorce/matrice (19/32) est visible entre 0,16 et 0,32  $\mu\text{M}$  de PabPolB (pistes 10 et 11). En présence de  $\text{Ca}^{2+}$ , la liaison de PabPolB au duplexe semble être totale dès 0,16  $\mu\text{M}$  de PabPolB (formation du complexe en piste 10 indiqué par la flèche noire). Cependant, l'interprétation du résultat est incertaine puisque l'intensité de fluorescence globale des bandes du P/T semble être parasitée en présence de  $\text{Ca}^{2+}$ . De façon générale, PabPolB semblerait se lier préférentiellement à l'ADN 32 nt > 19/32 > 19 nt. Lorsque le cofacteur  $\text{Mg}^{2+}$  est remplacé par le  $\text{Ca}^{2+}$ , seule la liaison de PabPolB au duplexe 19/32 semble être augmentée. Les différents complexes observés (indiqués par diverses flèches sur le gel) correspondent très probablement à des monomères (pour la forme la plus pénétrante) et des dimères de PabPolB. Dans ce cas, plusieurs possibilités de fixation existent : une liaison à la jonction amorce/matrice ou une liaison à la partie simple brin de la matrice ADN. Egalement, une fixation coopérative de PabPolB est envisageable. D'autres oligonucléotides ont été étudiés. PabPolB avait plus de capacité à se lier à un ADN simple brin de 87 nt lorsque le cofacteur utilisé était le  $\text{Ca}^{2+}$  ; en revanche, la capacité de liaison à un duplexe 17/87 était similaire en présence de  $\text{Mg}^{2+}$  ou de  $\text{Ca}^{2+}$  (résultats non montrés).



**Figure 16. Liaison de PabPolB *exo-* à différents types d'ADN en présence de Mg<sup>2+</sup> ou de Ca<sup>2+</sup>**

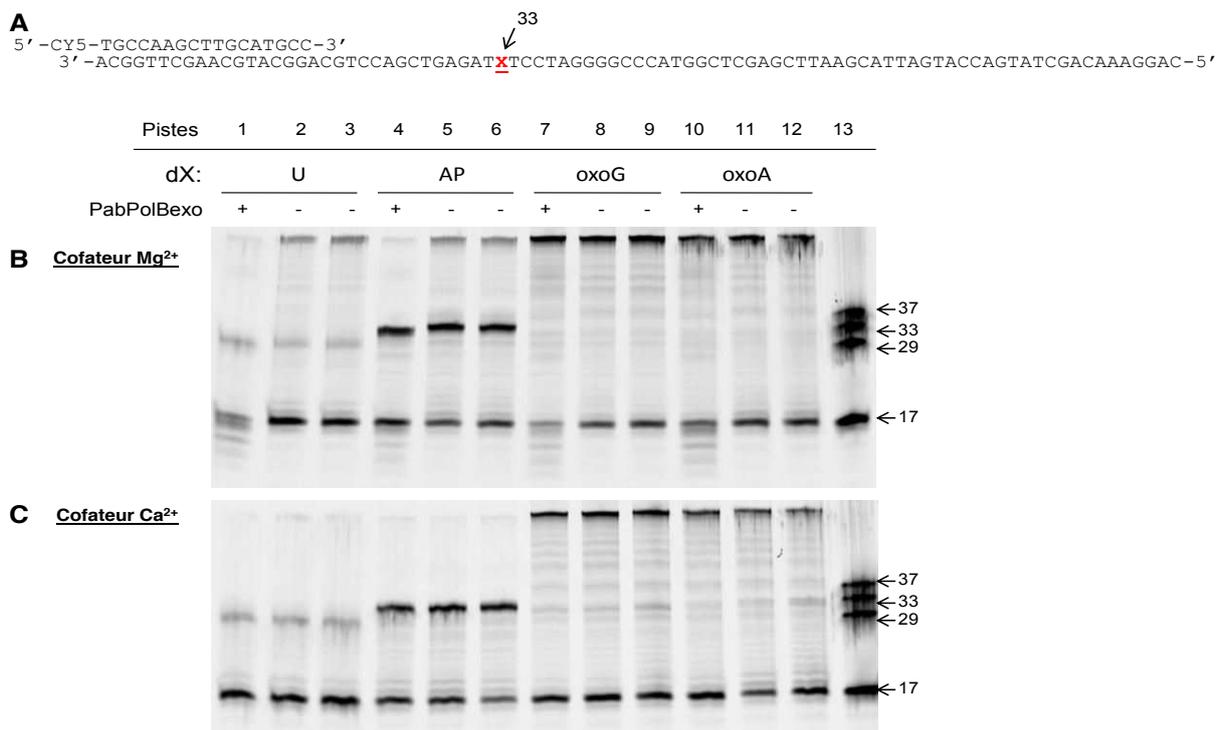
Profils d'électrophorèse en gels non dénaturants de polyacrylamide 8 % des complexes PabPolB/ADN marqués [ADN simple brin de 19 ou 32 nucléotides ou duplexe ADN amorce/matrice (19/32) avec marquage en 5' de l'oligonucléotide ou amorce (**A**)] formés au cours des réactions de liaison d'une gamme de concentrations croissantes de PabPolB *exo-* à l'ADN en présence de 5 mM de cofacteur Mg<sup>2+</sup> (**B**) ou Ca<sup>2+</sup> (**C**) pendant 20 min à 25°C.

## II.2. Comportement de PabPol B face aux dommages à l'ADN en présence de Ca<sup>2+</sup>

### II.2.1. Tolérance de PabPolB à quelques exemples de dommages

PabPolB est une pol répliquative hautement fidèle. Cependant, *in vitro*, une certaine tolérance aux dommages a été mise en évidence, notamment lorsque PabPolB est confrontée à un dommage apurique/apyrimidique (AP) (Palud et al., 2008) et à un dommage uracile (Gouge et al., 2012). La réduction du taux d'extension de PabPolB observée lorsqu'une base uracile est rencontrée dans la matrice ADN s'interprète comme la diminution de quantité de pol libre. Ceci s'explique par la haute capacité de liaison des pols répliquatives archéennes (Shuttleworth et al., 2004) à une matrice contenant un dommage uracile. En effet, la mesure de l'affinité par SPR a montré que la Pol B de *P. furiosus*, *Pfu*, est environ 300 fois plus affine à l'ADN contenant une uracile et est de l'ordre de la dizaine de nM (Gill et al., 2007). Les pols, ainsi fixées à la matrice endommagée, ne sont donc plus aptes à poursuivre l'extension d'amorce au-delà de ce dommage puisque fortement liées.

Dans notre étude, nous montrons, par des expériences d'extension d'amorce hybridée à une matrice contenant différents dommages à l'ADN, en condition « running start », que PabPoIB effectue une pause de la synthèse d'ADN quatre bases en amont de la base uracile (soit en position 29) et que de la synthèse au-delà du dommage est possible (Figure 17B, piste 2 et 3). De façon intéressante, la synthèse au-delà de la lésion uracile est quasiment inexistante lorsque  $\text{Ca}^{2+}$  est utilisé comme cofacteur par rapport à  $\text{Mg}^{2+}$  ; ceci est surtout vrai pour la version *exo-* (comparer Figure 17B et C, pistes 1, 2 et 3). Ainsi, la forte affinité conférée à PabPoIB par le  $\text{Ca}^{2+}$  (montré dans l'article en préparation) diminue la tolérance au dommage uracile par rapport au  $\text{Mg}^{2+}$ . Même dépourvue de son activité de relecture (mutant *exo-*), PabPoIB devient intolérante aux dommages lorsque  $\text{Ca}^{2+}$  est utilisé comme cofacteur à l'image de PabPoIB *exo+* (comparer Figure 17C, pistes 1 et 2). Les mêmes résultats sont observés pour le dommage AP, mais, dans ce cas, l'arrêt de la polymérisation a lieu exactement au niveau de la lésion AP, c'est-à-dire en position 33 (Figure 17B et C, pistes 4 à 6). La tolérance de PabPoIB aux dommages oxydatifs de l'ADN (oxoG et oxoA) se traduit sur les gels par l'absence d'une pause au cours de la polymérisation au moment de la lecture des bases endommagées (Figure 17B, pistes 7 à 13). Cette tolérance n'est quasiment pas affectée en présence de  $\text{Ca}^{2+}$  (Figure 17C, pistes 7 à 13). Ces résultats ne sont que préliminaires, cependant, ils soulignent que la nature des cofacteurs métalliques divalents peut moduler la propriété de tolérance des ADN polymérases à certains dommages.

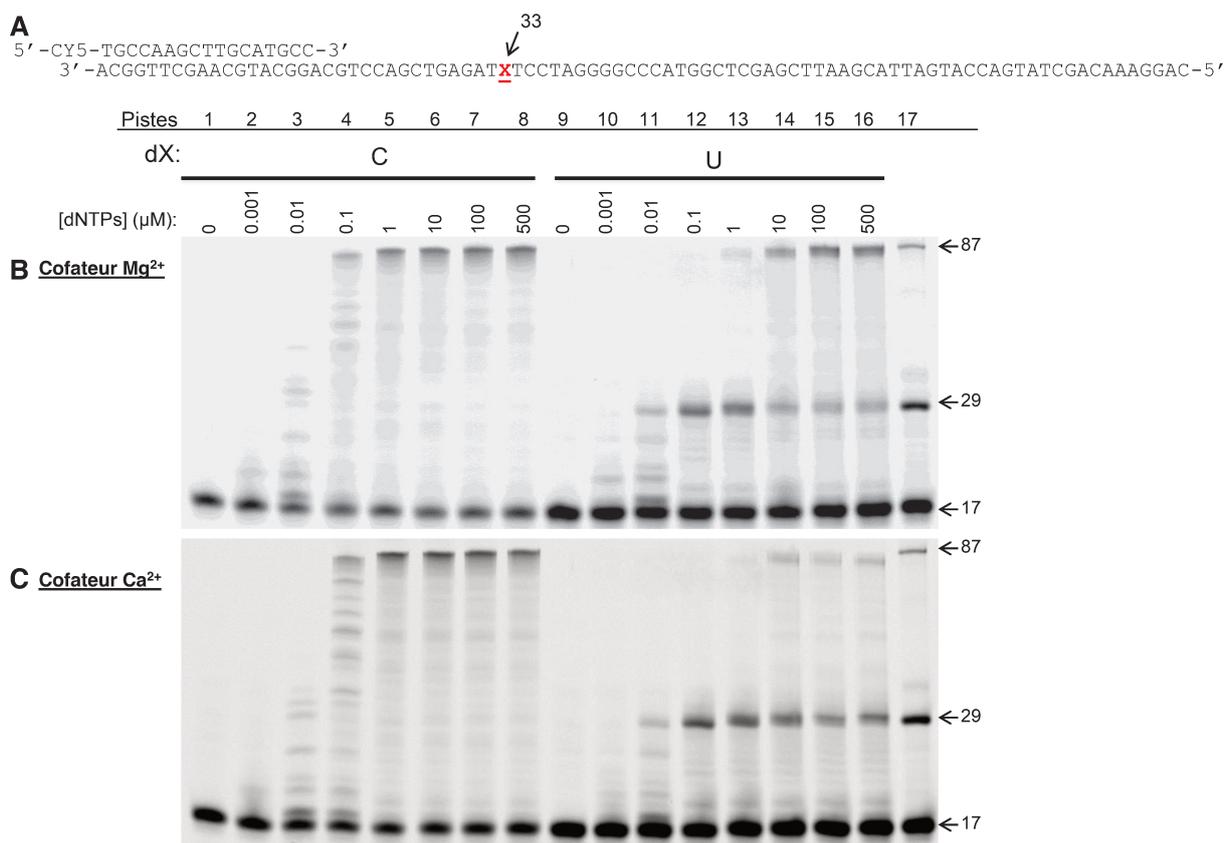


**Figure 17. Réaction d'extension d'amorce en présence d'ADN endommagé contenant une lésion ponctuelle par PabPoIB *exo-* avec du  $\text{Mg}^{2+}$  ou du  $\text{Ca}^{2+}$ .**

Profils d'électrophorèse en gels dénaturants de polyacrylamide 17 % / urée 8 M des produits des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanine® Cy5 de 17 nucléotides hybridée à une matrice de 87 nucléotides comportant une base uracile (U), apurinique (AP), 8oxoG ou 8oxoA en position 33 (X) (A), de 200  $\mu\text{M}$  de dNTPs et en présence de 5 mM de cofacteur  $\text{Mg}^{2+}$  (B) ou  $\text{Ca}^{2+}$  (C). Les réactions ont été initiées par l'ajout de 75 nM de PabPoIB *exo+* ou *exo-* comme indiqué et incubées 30 min à 55°C. La piste 17 comporte le marqueur de taille d'oligonucléotides.

## II.2.2. Influence de la variation de concentration de dNTPs sur la synthèse d'ADN erronée par PabPolB

La concentration en désoxyribonucléotide triphosphates est un paramètre limitant de la réaction de synthèse d'ADN. Ainsi, dans les conditions testées *in vitro*, 1  $\mu\text{M}$  de dNTPs sont nécessaires pour obtenir une extension de la longueur totale de l'amorce, soit 87 nucléotides (Figure 18B, piste 5). A 0,1  $\mu\text{M}$  de concentration en dNTPs, un arrêt net de l'extension d'amorce est observé quatre bases en amont du dommage uracile (Figure 18B, piste 12). Lorsque des concentrations plus élevées en dNTPs sont utilisées, un arrêt de la synthèse d'ADN est observé mais des produits d'extension de 87 nucléotides sont synthétisés (Figure 18B, pistes 14 à 16). Ces résultats révèlent que la concentration en dNTPs peut affecter la tolérance des pols au dommage uracile. Si le cofacteur utilisé est  $\text{Ca}^{2+}$ , le taux d'extension de l'amorce est similaire à celui donné par  $\text{Mg}^{2+}$  (comparer pistes 1 à 8, Figures 18B et C). Lorsqu'un dommage uracile est présent, PabPolB effectue aussi un arrêt de la synthèse en position 29. En revanche, la synthèse au-delà du dommage (produits > 29 nt) est amoindrie par rapport à  $\text{Mg}^{2+}$  (pistes 14 à 16, comparer Figure 18B et C). Là encore, ces résultats suggèrent que  $\text{Ca}^{2+}$  en tant que cofacteur préviendrait la synthèse erronée de la pol réplicative.



**Figure 18.** Influence de la concentration en dNTPs sur la réaction de polymérisation par PabPolB avec  $\text{Mg}^{2+}$  ou  $\text{Ca}^{2+}$  en présence d'une matrice ADN saine ou contenant un dommage uracile.

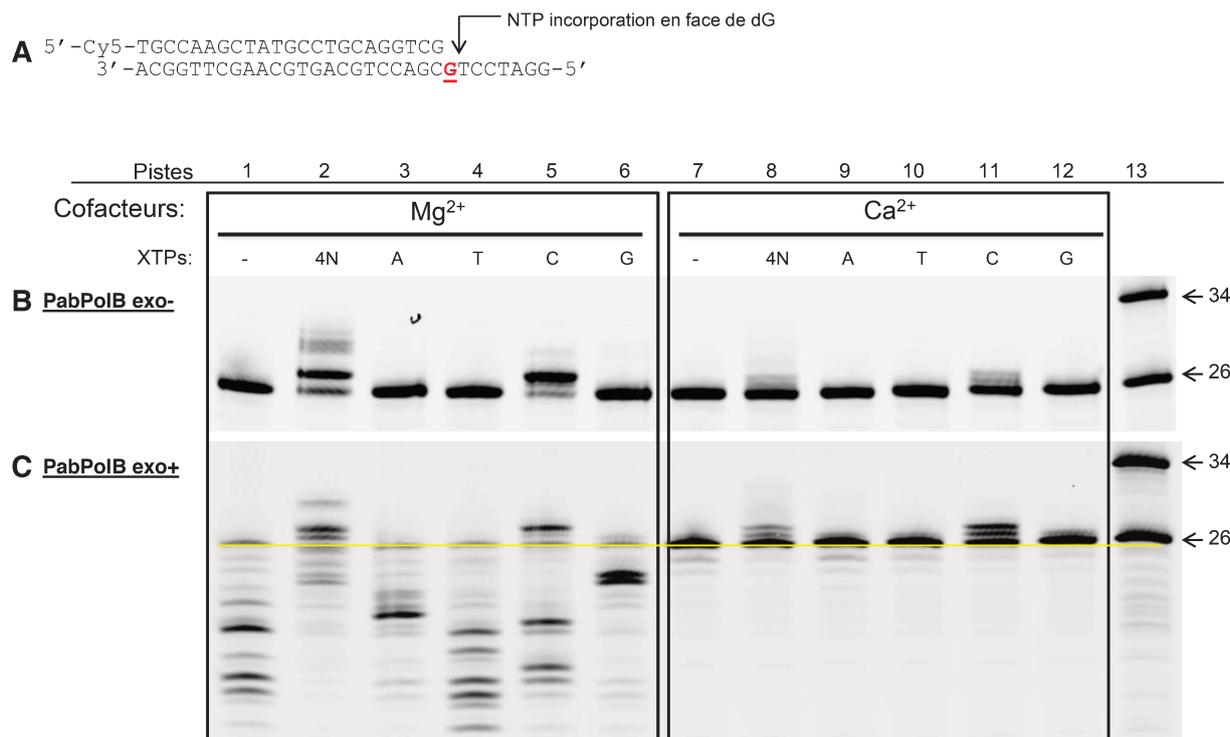
Profils d'électrophorèse en gels dénaturants de polyacrylamide 17 % / urée 8 M des produits des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 17 nucléotides hybridée à une matrice de 87 nucléotides comportant une base cytosine ou uracile en position 33 (X) (A), en présence d'une gamme de concentrations croissantes en dNTPs (indiquées au-dessus des gels) et en présence de 5 mM de cofacteur  $\text{Mg}^{2+}$  (B) ou  $\text{Ca}^{2+}$  (C). Les réactions ont été initiées par l'ajout de 75 nM de PabPolB *exo-* et incubées 30 min à 55°C. La piste 17 comporte le marqueur de taille d'oligonucléotides.

## II.3. Comportement de PabPolB en présence de ribonucléotides (NTPs)

### II.3.1. Incorporation unique de NTP

Les pols translesionnelles de la famille Y, en raison de leur faible spécificité de substrat, peuvent incorporer des NTPs lors de la synthèse d'ADN (Grúz et al., 2003). Plus récemment, malgré la spécificité élevée des pols répliquatives pour leurs substrats, des études ont montré que la Pol  $\delta$  eucaryote incorpore des NTPs au cours de la synthèse d'ADN (Nick McElhinny et al., 2010a).

Dans cette expérience, la capacité de la Pol B archéenne de *P. abyssi* à incorporer des NTPs a été explorée en présence des cofacteurs  $Mg^{2+}$  et  $Ca^{2+}$ . Une structure particulière d'amorce hybridée à une matrice a été utilisée de sorte qu'au cours de l'extension, la première base instructrice à être lue soit la base G. Tout d'abord, les résultats montrent que l'incorporation de ribonucléotides par PabPolB *exo+* et PabPolB *exo-* est possible (Figure 19B et C, pistes 2), en revanche l'incorporation n'est pas suffisante pour obtenir un produit d'extension finale de 34 nt (Figure 19 B et C, pistes 2). Ensuite, les résultats montrent que seule la base CTP est incorporée et respectant ainsi l'appariement Watson-Crick (piste 5), la sélectivité de la base et non du sucre semble donc être conservée. Toutefois, la stimulation intense de l'activité exonucléase lors de l'incorporation du CTP par PabPolB *exo+* (Figure 19C, piste 5) témoigne d'un signal d'alarme. Cette activité de relecture semble prévenir l'incorporation des trois autres bases (Figure 19C, pistes 3 à 6). Lorsque  $Ca^{2+}$  est le cofacteur de la réaction, PabPolB *exo-* incorpore très faiblement le pool des quatre NTPs, une base seulement (Figure 19B, piste 8) de la même façon que la base CTP (Figure 19B, piste 11). Nous avons montré précédemment que l'activité de relecture de PabPolB *exo+* est inhibée lorsque  $Ca^{2+}$  est le cofacteur (article en préparation) l'incorporation de deux bases de CTP par PabPolB *exo+* est tout de même réalisée en présence de  $Ca^{2+}$  mais de façon moindre à celle faite en présence de  $Mg^{2+}$  (comparer pistes 5 et pistes 11, Figure 19B et C).



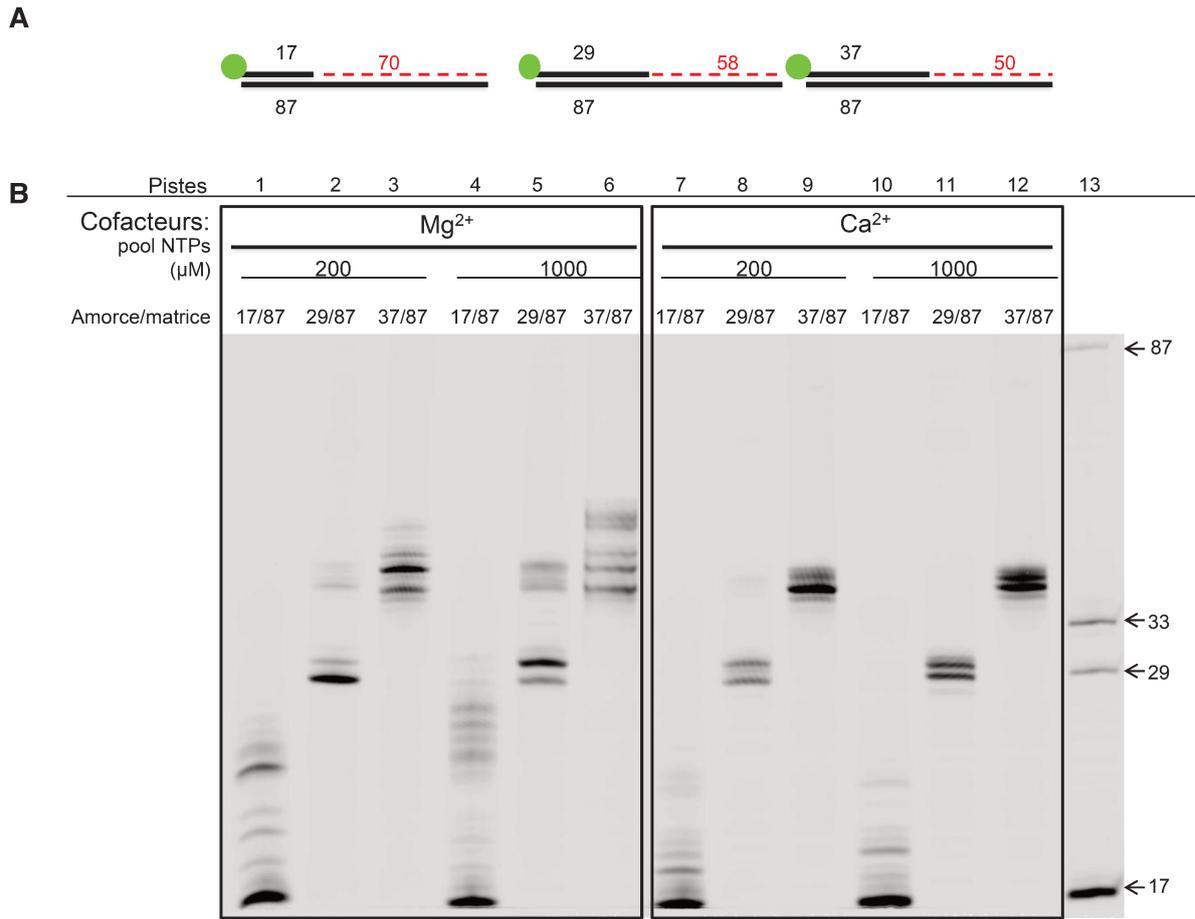
**Figure 19. Réaction d'incorporation préférentielle d'un seul ribonucléotide face à une base guanine par PabPoIB exo- et PabPoIB exo+.**

Profils d'électrophorèse en gels dénaturants de polyacrylamide 17 % / urée 8M des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 26 nucléotides hybridée à une matrice de 34 nt contenant une base guanine en position +1 par rapport à la jonction amorce/matrice (dX) (**A**) et de 200  $\mu$ M de ribonucléotides triphosphates (NTPs) différents indiqués au-dessus des gels. Les réactions ont été initiées par l'ajout de 75 nM de PabPoIB exo- (**B**) et PabPoIB exo+ (**C**) et incubées 30 min à 55°C. Les pistes 1B et C et 7B et C, montrent les résultats de la réaction en absence de substrat nucléotidique. Les pistes 2B et C et 8B et C montrent les résultats en présence des quatre substrats nucléotidiques ATP, TTP, CTP et GTP. Les oligonucléotides marqueurs de références 26 et 34 nucléotides sont indiqués en pistes 13 B et C.

### II.3.2. Extension d'amorce en présence de NTPs libres et en fonction de la longueur de l'amorce

Les résultats obtenus montrent que PabPoIB est capable de faire de la synthèse d'ARN à partir de NTPs et d'une matrice ADN (Figure 20B). Cependant, cette extension est relativement faible et s'effectue sur le tiers de la longueur totale de la matrice. Aucun produit final d'extension (87 nt) n'est observable. De plus, nous pouvons observer qu'en présence de la structure d'ADN amorce/matrice (37/87), seules deux bases ont été incorporées (Figure 20, piste 3). Cette extension semble être dépendante de la longueur de l'amorce utilisée. En effet, plus l'amorce est longue et plus l'élongation ribonucléotidique est inhibée (Figure 20, comparer graduellement pistes 1, 2 et 3). L'augmentation de la concentration en NTP de 200  $\mu$ M à 1 000  $\mu$ M induit une petite augmentation de la longueur du brin amorce étendu (Figure 20, pistes 4 à 6). En présence de Ca<sup>2+</sup>, la synthèse ribonucléotidique est moins efficace, et, de façon intéressante, l'augmentation de la concentration en NTPs ne favorise pas davantage l'extension (Figure 20, comparer pistes 7 à 9 aux pistes 10 à 12).

Globalement, les résultats préliminaires sur l'étude de la sélectivité des ribonucléotides tendent à montrer que le Ca<sup>2+</sup> limite la synthèse erronée de PabPoIB.

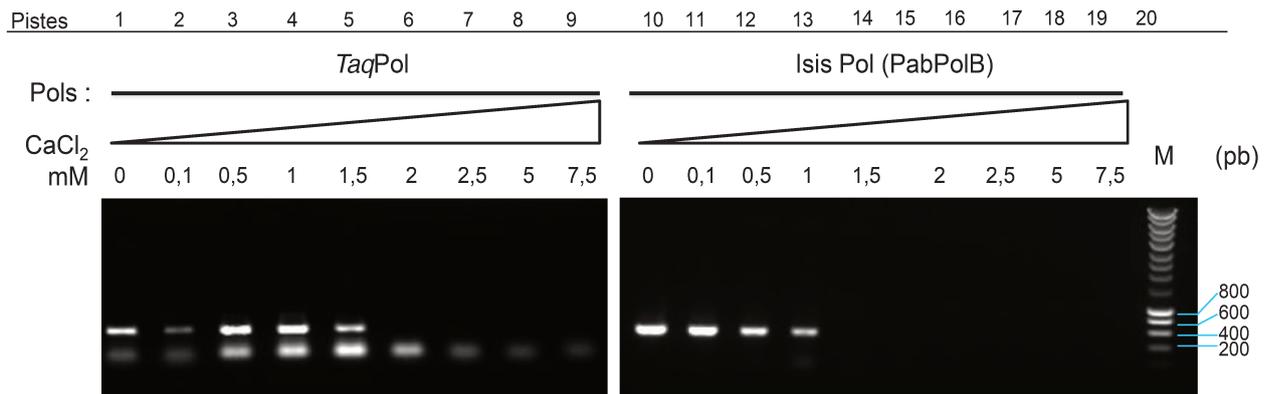


**Figure 20.** Extension d'amorce hybridée à une matrice de 87 nucléotides PabPolB exo- en fonction de la longueur de l'amorce ADN et en présence de Mg<sup>2+</sup> ou de Ca<sup>2+</sup> comme cofacteur.

Profil d'électrophorèse en gel dénaturant de polyacrylamide 17 % / urée 8 M des produits des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 17, 29 ou 37 nucléotides hybridée à une matrice de 87 nucléotides (A), en présence de 200 ou 1 000 μM de NTPs (indiquées au-dessus des gels) et en présence de 5 mM de cofacteur Mg<sup>2+</sup> ou Ca<sup>2+</sup>. Les réactions ont été initiées par l'ajout de 75 nM de PabPolB exo- et incubées 30 min à 55°C. La piste 13 comporte le marqueur de taille d'oligonucléotides.

#### II.4. Effet du calcium sur l'activité de PabPolB au cours de la PCR

Les réactions de PCR actuelle utilisent du  $MgSO_4$  comme cofacteur de PabPolB (commercialisée sous le nom d'Isis DNA pol par MP Biomedicals). En premier lieu, la capacité de PabPolB à utiliser le  $Ca^{2+}$  dans des réactions de PCR a été entrepris et n'a donné aucun résultat positif (résultats non montré). Il semblerait donc que le  $Ca^{2+}$  ne soit pas un cofacteur adapté à l'amplification *in vitro* d'ADN à cause du temps relativement court dédié à l'étape d'élongation de 2 minutes. En second lieu, l'impact du calcium dans les échantillons sur la réaction de PCR effectuée avec le  $Mg^{2+}$  a été étudiée. Ainsi, des expériences d'amplification d'un fragment de 500 kb d'ADN génomique extrait de *P. abyssi* ont été effectuées en présence du cofacteur classique  $Mg^{2+}$  et en présence d'une gamme croissante de  $CaCl_2$ . Les résultats montrent une inhibition nette de la réaction de PCR effectuée par la *Taq* polymérase dès 2 mM de  $CaCl_2$  (Figure 21, piste 6). Chez PabPolB (Isis) cette inhibition par le  $CaCl_2$  apparait dès 1,5 mM (Figure 21, piste 14). Des études ont montré que cette inhibition par les ions calcium de la réaction PCR conduite par la *Taq* polymérase, rencontrée lors d'expériences de biologie moléculaire effectuées dans les aliments lactés, pouvait être levée par l'ajout de grandes quantités compensatrices d'ion magnésium (Bickley et al., 1996).



**Figure 21. Amplification d'ADN génomique de *P. abyssi* par PCR en présence de  $CaCl_2$**

Profil d'électrophorèse en gel d'agarose BET des produits de réactions d'amplification d'un fragment de 500 kb du génome (20 ng/ $\mu$ L) de *P. abyssi* effectuées en présence d' 1  $\mu$ M d'amorces, de 200  $\mu$ M de dNTPs, de 2 mM de  $Mg^{2+}$  et d'une gamme de concentrations croissantes en  $CaCl_2$ . Les réactions ont été initiées par l'ajout de 1 unité de *Taq*Pol ou de Isis™Pol et incubées dans un thermocycleur. La piste 20 comporte le marqueur de taille de fragments (Smart Ladder Long Fragment Eurogentec).



## **CHAPITRE III**

Etude structurale et biochimique de la reconnaissance de dommage à l'ADN (base nucléotidique désaminée) par l'ADN polymérase B de *Pyrococcus abyssi*.

---

Présentation de l'étude

Article

Résultats supplémentaires



# Première partie: Présentation de l'étude

---

Le maintien de l'intégrité du génome de génération en génération est indispensable à la survie des espèces. L'ADN des organismes est constamment soumis à des agressions physiques et chimiques plus ou moins intenses et pouvant varier au cours du temps. Il n'est donc pas rare que des dommages à l'ADN surviennent. Il existe des mécanismes de détection et de réparation des dommages à l'ADN qui, s'ils ne sont pas fidèles, génèrent des lésions persistantes qui pourront être transmises aux générations futures.

Bien que la structure en double hélice de l'ADN offre une protection contre les coupures glycosidiques et les désaminations, les hautes températures participent, entre autres, à l'accélération de la réaction de désamination des bases nucléotidiques de l'ADN adénine et cytosine (Lindahl and Karlstrom, 1973 ; Lindahl and Nyberg, 1974 Lindahl, 1979 ; Shen et al., 1994). L'ADN simple brin est formé transitoirement pendant la transcription, la réplication et la recombinaison. C'est au cours de ces processus que les bases cytosines et adénosine de l'ADN sont particulièrement vulnérables. Une des réactions chimiques courantes est la désamination de la cytosine et de l'adénine qui donne respectivement l'uracile et l'hyoxanthine par hydrolyse d'un groupement amine  $-NH_2$ . L'uracile, possédant une structure chimique proche de celle de la thymine, peut s'apparier à une adénine donnant une transition C:G  $\rightarrow$  T:A à la prochaine génération de cellules. Ce dommage est donc hautement mutagène.

Dans la cellule, deux lignes de défense contre la réaction de désamination de la cytosine ont été identifiés à ce jour. Un des mécanismes consiste en l'élimination de la base uracile. Il est initié par des glycosylases spécifiques de l'uracile (Pearl, 2000) et est ensuite pris en charge par les voies de réparation de l'ADN. En effet, des UDG (Uracil DNA Glycosylases) ont été identifiées chez des organismes thermophiles (Sandigursky and Franklin, 1999 ; Starkuviene and Fritz, 2002). Tout récemment, une UDG a été isolée et caractérisée chez une Archée du genre *Pyrococcus* (Li-bo et al., 2012). Une autre voie existe au sein de la cellule et consiste en l'ajustement de la concentration cellulaire en dUTP (Vértessy et al, 2009). Une troisième ligne de défense spécifique est présente au sein même de certaines ADN polymérases. Il s'agit d'une reconnaissance structurale de l'uracile. Ce mécanisme a été vu en détail dans la deuxième partie du chapitre I de ce manuscrit (page 48), il participe à la haute fidélité des polys répliquatives archéennes. C'est un mécanisme moléculaire de lecture en amont de la base uracile (Greagg et al., 1999). Il induit un arrêt de la polymérisation qui permet ensuite l'intervention des mécanismes de réparation de l'ADN, il s'agit en quelque sorte d'un mécanisme d'alerte précoce, signalant une anomalie de la réplication.

En collaboration avec les cristallographes Marc Delarue et Jérôme Gouge de l'Unité de Dynamique Structurale des Macromolécules de l'Institut Pasteur à Paris, nous avons cherché à comprendre plus en détail le phénomène de reconnaissance de la base endommagée uracile par l'ADN polymérase B de *Pyrococcus abyssi* au cours de la synthèse d'ADN. Par une approche méthodologique complémentaire, à la fois structurale et fonctionnelle, nous avons tenté de comprendre les mécanismes qu'utilise PabPolB, dépourvue de sa fonction exonucléasique, pour discriminer les bases endommagées des bases canoniques en cours de synthèse de l'ADN.

Des cristaux de PabPolB ont été obtenus puis la structure tridimensionnelle de PabPolB en complexe avec une construction amorce/matrice d'ADN contenant une base désaminée uracile a été résolue. La structure cristallographique obtenue, hautement résolutive, a révélé l'existence, au sein du domaine N-terminal, d'une structure en « T-cleft » située sous la poche de reconnaissance de l'uracile. Cette structure implique trois acides aminés qui chélatent un ion métallique divalent par l'intermédiaire de six molécules d'eau, formant ainsi une sphère d'hydratation. La présence de six molécules d'eau autour d'un métal est rare dans les structures cristallographiques des protéines mais plus courant dans les structures d'oligonucléotides. De plus, en raison de sa proximité avec la poche de reconnaissance à l'uracile, l'hypothèse proposée est que ce T-cleft jouerait un rôle dans la liaison à l'ADN et dans la reconnaissance de dommage. Ainsi, deux hypothèses ont été émises quant au rôle de l'ion métallique divalent :

i) Le métal sert de guide à l'ADN pour passer dans le T-cleft.

ii) Le métal permet à PabPolB d'effectuer une reconnaissance de l'uracile, de marquer une pause dans la polymérisation et ainsi d'augmenter la discrimination entre une base lésée et une base canonique. Ces hypothèses, qui reposent sur des résultats de données structurales, ont été étudiées d'un point de vue fonctionnel. Pour cela, des mutations ponctuelles des acides aminés de PabPolB formant le T-cleft impliqué dans la liaison à l'ion métallique ont été effectuées (D4A (acide aspartique → alanine), E251A (acide glutamique → alanine), D343A) rendant ainsi les acides aminés moins hydrophiles, et de fait, abolissant les liaisons avec la sphère d'hydratation.

L'étude fonctionnelle s'est déroulée de la façon suivante :

- Production et purification de PabPolB et de ses mutants (Annexe 1, pages 189-190).
- Contrôle quantitatif (dosage de la concentration) et qualitatif (pureté (Annexe 2, page 190), activité exonucléase) des protéines purifiées.
- Mise au point des conditions expérimentales pour la mesure des activités ADN polymérases et comparaison de la cinétique d'extension d'amorce hybridée à de l'ADN sain ou de l'ADN endommagé par les PabPolBs.
- Mesure de la capacité de liaison des PabPolBs aux matrices saines et endommagées en présence ou non d'ion métallique  $Mg^{2+}$ .

L'ensemble des résultats obtenus a été valorisé sous la forme de publication scientifique dans la revue scientifique « Journal of Molecular Biology », (voir ci-après).

## Deuxième partie: Article

---





## Molecular Recognition of Canonical and Deaminated Bases by *P. abyssi* Family B DNA Polymerase

Jérôme Gouge<sup>1,2</sup>, Céline Ralec<sup>3,4</sup>, Ghislaine Henneke<sup>3,4</sup>  
and Marc Delarue<sup>1\*</sup>

<sup>1</sup>Unité de Dynamique Structurale des Macromolécules, UMR 3528 du CNRS, Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France

<sup>2</sup>Université Paris Diderot, Sorbonne Paris Cité, Cellule Pasteur, Paris, France

<sup>3</sup>Ifremer, UMR 6197 du CNRS, Laboratoire de Microbiologie des Environnements Extrêmes, BP 70, 29280 Plouzané, France

<sup>4</sup>Université de Bretagne Occidentale, UMR 6197 du CNRS, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France

Received 6 June 2012;  
received in revised form  
27 July 2012;  
accepted 30 July 2012  
Available online  
16 August 2012

Edited by R. Huber

### Keywords:

DNA polymerase;  
X-ray crystallography;  
deaminated bases;  
archaea;  
proofreading mechanism

Euryarchaeal polymerase B can recognize deaminated bases on the template strand, effectively stalling the replication fork 4nt downstream the modified base. Using *Pyrococcus abyssi* DNA B family polymerase (PabPolB), we investigated the discrimination between deaminated and natural nucleotide(s) by primer extension assays, electrophoretic mobility shift assays, and X-ray crystallography. Structures of complexes between the protein and DNA duplexes with either a dU or a dH in position +4 were solved at 2.3 Å and 2.9 Å resolution, respectively. The PabPolB is found in the editing mode. A new metal binding site has been uncovered below the base-checking cavity where the +4 base is flipped out; it is fully hydrated in an octahedral fashion and helps guide the strongly kinked template strand. Four other crystal structures with each of the canonical bases were also solved in the editing mode, and the presence of three nucleotides in the exonuclease site caused a shift in the coordination state of its metal A from octahedral to tetrahedral. Surprisingly, we find that all canonical bases also enter the base-checking pocket with very small differences in the binding geometry and in the calculated binding free energy compared to deaminated ones. To explain how this can lead to stalling of the replication fork, the full catalytic pathway and its branches must be taken into account, during which the base is checked several times. Our results strongly suggest a switch from elongation to editing modes right after nucleotide insertion when the modified base is at position +5.

© 2012 Elsevier Ltd. All rights reserved.

\*Corresponding author. Department of Structural Biology and Chemistry, Institut Pasteur, 25 rue du Dr Roux, Paris 75015, France. E-mail address: [marc.delarue@pasteur.fr](mailto:marc.delarue@pasteur.fr).

Abbreviations used: PabPolB, *Pyrococcus abyssi* B family DNA polymerase; P/T, primer–template; PDB, Protein Data Bank; ssDNA, single-stranded DNA; TLS, translesion synthesis; EMSA, electrophoretic mobility shift assay; TgoPolB, *Thermococcus gorgonarius* B family DNA polymerase; PfuPolB, *Pyrococcus furiosus* B family DNA polymerase; dsDNA, double-stranded DNA; Mes, 4-morpholineethanesulfonic acid; EDTA, ethylenediaminetetraacetic acid.

## Introduction

Deamination of nucleotides spontaneously occurs in cells under stress conditions such as high temperature<sup>1</sup> or oxidative stress.<sup>2</sup> Cytosine and adenine are especially sensitive to those damages, resulting in uracil (U) and hypoxanthine (H), respectively. The first round of DNA replication produces the base pairs G:U and T:H. After the second round of replication, the DNA contains a transition G:U→A:T or T:H→C:G. Typically, the cell has two lines of defense against this phenomenon; first, it can try to keep the dUTP concentration low by using dUTPase, or it can remove U from double-stranded DNA (dsDNA) by using uracil-DNA glycosylases that recognize the G:U mismatch, flip it out, and excise uracil by glycosidic cleavage.<sup>3</sup> The DNA repair is further achieved by the base excision repair pathway.<sup>4</sup> Base excision repair is present in every living organism, even in archaea.<sup>5</sup> A third line of defense exists in this particular kingdom, embedded in the DNA polymerase itself and in the progression of the replication fork. Indeed, archaeal family B DNA polymerases possess an N-terminal domain that can recognize and bind deaminated bases 4nt upstream the primer-template (P/T) junction during DNA replication. Upon binding to the damaged base, elongation is stalled and exonuclease activity is concomitantly increased.<sup>6</sup> On the structural level, two crystal structures of *Thermococcus gorgonarius* PolB with either a uracil at +4 [Protein Data Bank (PDB) ID 2VWJ<sup>7</sup>] or a hypoxanthine at +2 (PDB ID 2XHB<sup>7</sup>) confirmed the existence of a specific pocket that binds the deaminated base. In these structures, the polymerase is in the editing mode even if the duplex is perfectly matched. Thus, with the primer 3' end being diverted away from the elongation site, the polymerase is unable to add nucleotides. On the functional level, it has been measured that the base-checking pocket's affinity is 100 times higher for single-stranded DNA (ssDNA) containing a dU instead of a regular dT.<sup>8</sup> However, the implied binding energy difference represents only about 3kcal/mol, corresponding to the energy necessary to break about two to three hydrogen bonds, and thus can hardly account for the stalling observed. A second paradox concerns the reparation process: the deaminated base being bound and buried in the protein, the DNA repair machinery does not really have access to the lesion to process the damage. In order to further document this phenomenon, here we present structural and functional data on another PolB from the same family, namely, *Pyrococcus abyssi* PolB, where it was possible to crystallize a complex with DNA containing either deaminated bases or canonical ones in the base-checking pocket. While structural and functional data continue to grow on

DNA polymerase at a rapid pace (see, e.g., Polbase<sup>9</sup>), structural information on archaeal PolB used in PCR is still relatively scarce, despite obvious biotechnological applications.

*P. abyssi* is a hyperthermophilic archaeon that grows at an optimum temperature of 95°C.<sup>10</sup> Previous studies have shown that PCR applications<sup>11</sup> for the B family polymerase (PabPolB) are possible. Due to its living conditions in hydrothermal vents, *P. abyssi* has to face environmental fluctuations that may damage its genome. Genetic maintenance mechanisms therefore play a key role for this euryarchaeota to survive. For example, in front of an abasic site, PabPolB is able to achieve translesion synthesis (TLS) by preferentially adding dATP.<sup>12</sup> Here, we report the activity of Pab polymerase B on P/T containing a modified base dU on the template strand. Crystal structures of PabPolB in complex with a dU or a dH in position +4, solved at a resolution of 2.3 and 2.9 Å, respectively, were used to study the specific recognition of deaminated nucleotides. In order to compare with canonical bases, we crystallized PabPolB in the editing mode: the crystal structures, solved between 2.15 and 3.2 Å resolution, revealed that the natural bases may be accommodated in the same base-checking pocket but are not as stabilized as deaminated bases. The capacity to bind each nucleotide was estimated with electrophoretic mobility shift assay (EMSA) experiments and *in silico* docking programs. In addition, a new metal binding site has been found in the T-cleft. This metal is fully hydrated in an octahedral fashion and seems to be involved in guiding the 5' end of the template strand below the cavity. A triple mutant was constructed to abolish this metal binding site, and its structure was solved at 2.9 Å resolution. Its functional properties were also characterized.

## Results

### Activity of PabPolB on primer-templates containing a deaminated base

As many other members of the archaeal family B DNA polymerases, PabPolB contains an NH<sub>2</sub>-terminal domain that can recognize and bind deaminated bases on the template strand 4nt upstream the P/T junction. The residues involved in the recognition are highly conserved in euryarchaeal and crenarchaeal family B DNA polymerases.<sup>13,14</sup> We have tested the activity of PabPolB (exo+ and exo-) on primer-templates containing a single dU on the template strand at position +37 (Fig. 1). Extension assays show that PabPolB stops to elongate the primer up to position +33, exactly 4nt before the lesion (Fig. 1a). We also observe that full-length

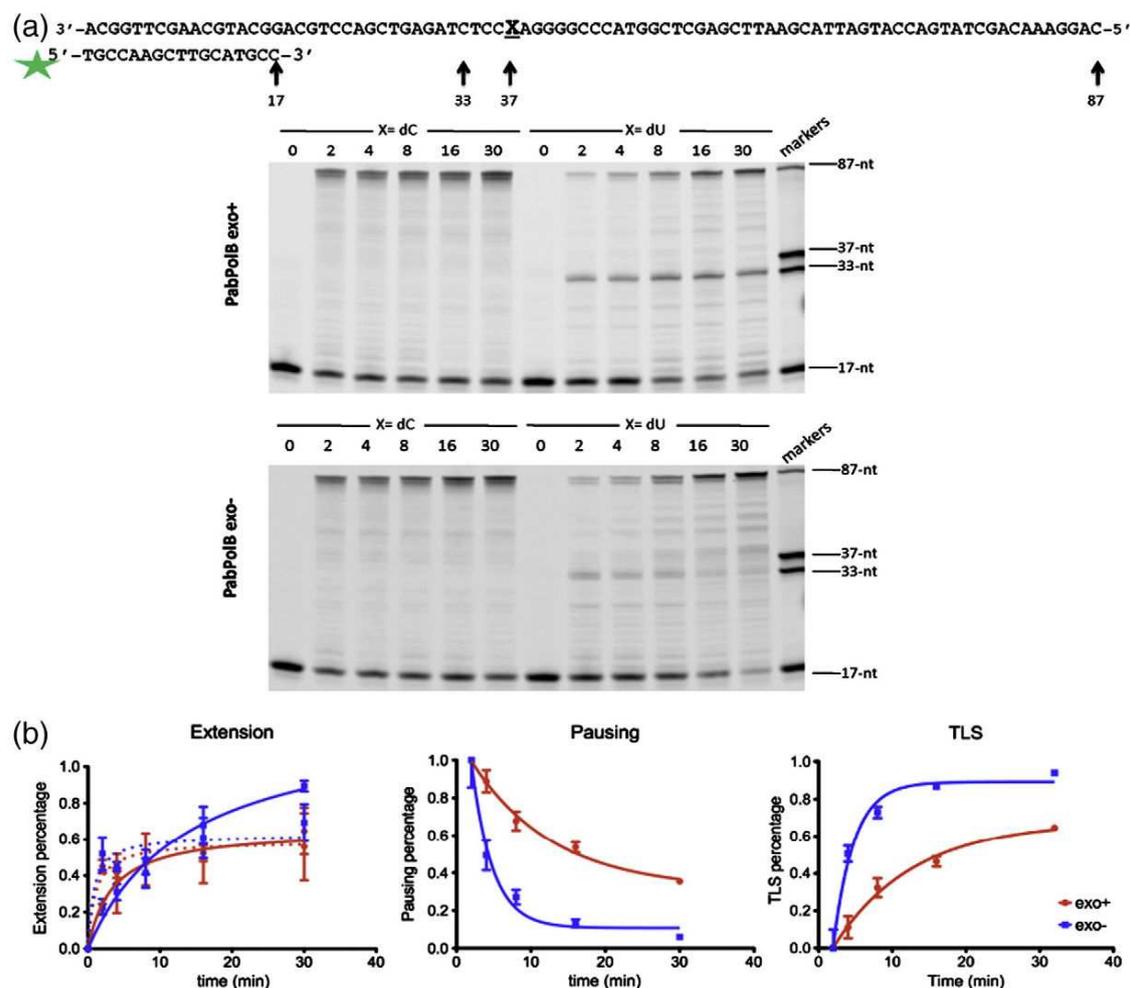
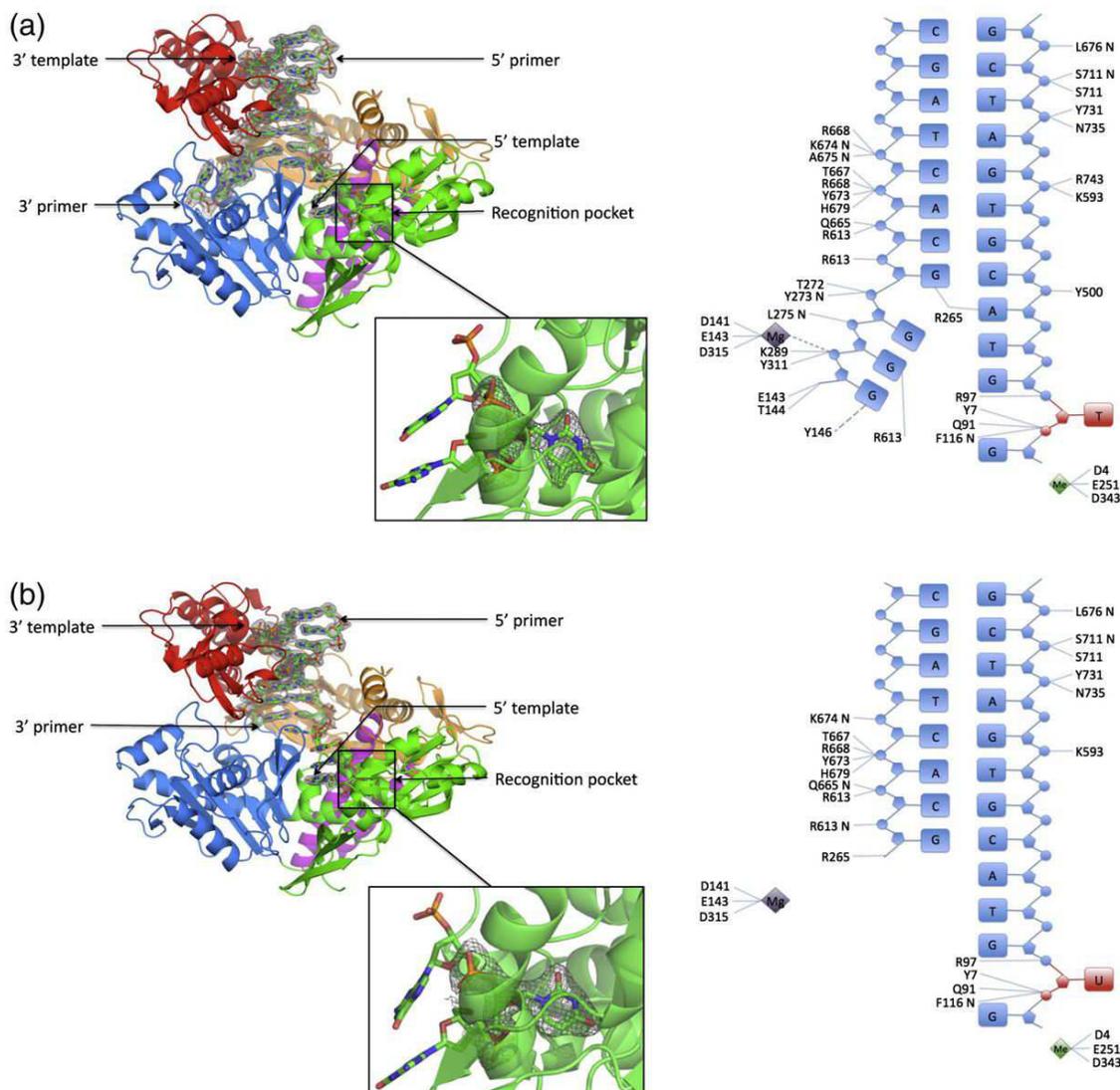


Fig. 1. PabPolB-catalyzed extension of a dsDNA with and without a deaminated base in the template strand. (a) Extension assays. P/T duplexes (25nM, labeled with Cy5 at the 5' end of the primer) with either a dU or a dC located 16nt upstream the P/T junction were extended by 75 nM PabPolB, either exo+ (top) or exo- (bottom). For both constructions, the polymerase stops before the lesion but is also able to proceed further and to bypass the deaminated nucleotide, the so-called TLS activity. (b) Elongation analyses. Intensities of bands were measured and plotted as a function of time to illustrate elongation, stalling, and TLS. The broken lines correspond to extension assays with canonical template, whereas the continuous lines correspond to the extension assays with a damaged template. The blue curve corresponds to PabPolB exo-, while the red one corresponds to PabPolB exo+. The extension, pausing, and translesion percentages are defined in Materials and Methods. The primer extension percentage increases rapidly and reaches a plateau with a canonical template. When a dU is present in the template strand, the slope of extension percentage is lower at a short incubation time. At 30 min of incubation, the extension percentage by PabPolB exo+ is in the range of the extension rate on a canonical template (about 60%), whereas PabPolB exo- extends more primer. The pausing percentage rapidly decreases for PabPolB exo-, whereas pausing is stronger for PabPolB exo+. A consequence is that the TLS rate is always higher for PabPolB exo- than for PabPolB exo+.

extension product accumulates at longer incubation times when the template strand contains a dU. These observations are similar to those observed for *Pyrococcus furiosus* PolB.<sup>15,16</sup> Therefore, PabPolB is able to recognize the deaminated base, thus inducing prominent uracil-dependent pause sites during DNA synthesis.

Comparison of the extension assays with dC and dU at the same position shows a clear reduced rate

of extension at short incubation times ( $t < 10$  min) when uracil is present on the template strand. However, at longer incubation times, PabPolB exo- extends more primer when a deaminated base is present ( $89 \pm 2\%$  for exo- versus  $56 \pm 18\%$  for the exo+ construct). A possible explanation for this phenomenon could simply be a higher turnover of the exo- polymerases. The pausing for PabPolB exo+ is stronger and longer than that for PabPolB



**Fig. 2.** Structures of 13T\_11 and 13U\_8: Canonical bases also enter the base-checking pocket. (a) Overall view of PabPolB complexed with a duplex containing three mismatches (13T\_11), which ensures that the polymerase is in the editing mode. The NH2 domain (residues 1-130 and 327-369) is depicted in green, the exonuclease domain (residues 131-326) is colored blue, the palm domain (residues 369-449 and 500-569) is in orange, the finger domain (residues 450-499) is depicted in pink, and the thumb domain (residues 586-757) is colored red. The inset shows a thymine bound in the recognition pocket. The contouring of  $2F_o - F_c$  electron density map is at  $1.2\sigma$ . The right part of the figure shows the interactions between PabPolB and the dsDNA. The base located in the cavity has been flipped and is colored red. In the exonuclease site, metal A, in deep purple, is visible in all the structures and interacts with the phosphodiester bond to be cleaved for structures containing the primer in the E-cleft. In all our structures, a second metal is bound in the T-cleft, in green. For the highest-resolution structures, six water molecules arranged in octahedron coordinate the metal and all metal-protein interactions are indirect. (b) Overall view of PabPolB complexed with a dU in position +4 (13U\_8). The polymerase is in the editing mode. The  $2F_o - F_c$  electron density map of the bound dU is contoured at  $1.2\sigma$ . The right part of the figure shows the interactions between PabPolB and the P/T duplex.

exo- (Fig. 1b). This suggests that, as seen for *T. gorgonarius* B family DNA polymerase (TgoPolB), the polymerase needs to degrade the primer strand to achieve an efficient stalling. Consistently, when PabPolB exo+ was incubated in primer extension assays with uracil at +4 position, primer degrada-

tion was observed (data not shown). TLS is always observed for all the constructions tested as seen with *Pyrococcus furiosus* B family DNA polymerase (PfuPolB).<sup>6</sup> At a finer level of analysis, it is apparent that the rate of TLS is higher for the exo-polymerase, the stalling being less effective. In the

case of PabPolB exo+, the TLS is slower because the polymerase is able to digest the extended primer, thereby locking the uracil in the recognition pocket.

### Structures of PabPolB in editing mode

The different crystal structures presented in this article are designated by  $nX_m$ , where the first number ( $n$ ) corresponds to the length of the template; the letter ( $X$ ) identifies the base bound in the pocket; the last number ( $m$ ) corresponds to the length of the primer strand (Fig. 2). Structures 13T\_11 and 13U\_8 will be taken as (respectively)

representative of PabPolB in the exonuclease mode (13A\_11, 13C\_11, 13G\_11, 13T\_11) or bound to a deaminated base (13H\_8, 13U\_8, 15H\_8, 15U\_8) and will be discussed in greater detail. These structures are chosen as references because of their higher resolution. Data collection and refinement statistics are listed in Table 1. In all cases, the DNA is clearly visible in  $2F_o - F_c$  maps as well as in omit maps.

The overall superimposition of all C $^\alpha$  atoms for 13U\_8 and 13T\_11 (Fig. S1) reveals that PabPolB bound to uracil adopts a conformation compatible with a 3'-5' exonuclease activity (i.e., editing mode) even if the DNA duplex is perfectly matched (rmsd,

**Table 1.** Data collection and refinement statistics

	13A_11	13C_11	13G_11	13T_11	13U_8 soaked in MnCl <sub>2</sub>
PDB ID	4FLT	4FLU	4FLV	4FLW	4FM1
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>				
Cell parameters					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	69.21, 114.86, 129.17	69.05, 114.63, 128.74	69.03, 115.0, 129.21	69.03, 114.49, 127.80	68.73, 115.02, 127.26
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution shell	47–2.9	44–3.1	47–2.7	44–2.15	44–3.0
No. of unique reflections	22,297 (3248)	18,522 (2703)	28,989 (4136)	55,498 (7839)	505,777 (7266)
Multiplicity	4	4	4.9	5.8	2.6
Completeness	97.2 (98.7)	97.7 (98.8)	100 (100)	99.5 (98.1)	92.7 (95.1)
$R_{\text{merge}}$ (%)	13.4 (51.3)	15.8 (45.6)	13.2 (66.1)	9.4 (74.6)	14.7 (50.6)
$I/\sigma(I)$	11.2 (3.1)	10 (3.0)	13 (2.4)	14.2 (2.3)	6.9 (2.1)
$R$ (%)	16.4	16.3	17.1	16.7	16.5
$R_{\text{free}}$ (%)	20.6	21.7	21.5	21.7	22.6
No. of protein atoms	6107	6130	6102	6115	6043
No. of DNA atoms	484	492	494	493	426
No. of water molecules	313	212	363	671	256
$B$ -factor overall (Å <sup>2</sup> )	30.2	32.8	42.9	43.3	36.6
$B$ -factor DNA (Å <sup>2</sup> )	55.5	52.2	46.2	52.0	53.0
Ramachandran outliers ( $n$ )	0	0	0	0	0
RMSD bond lengths (Å)	0.01	0.01	0.01	0.01	0.01
RMSD bond angles (°)	1.11	1.09	1.08	1.04	1.12

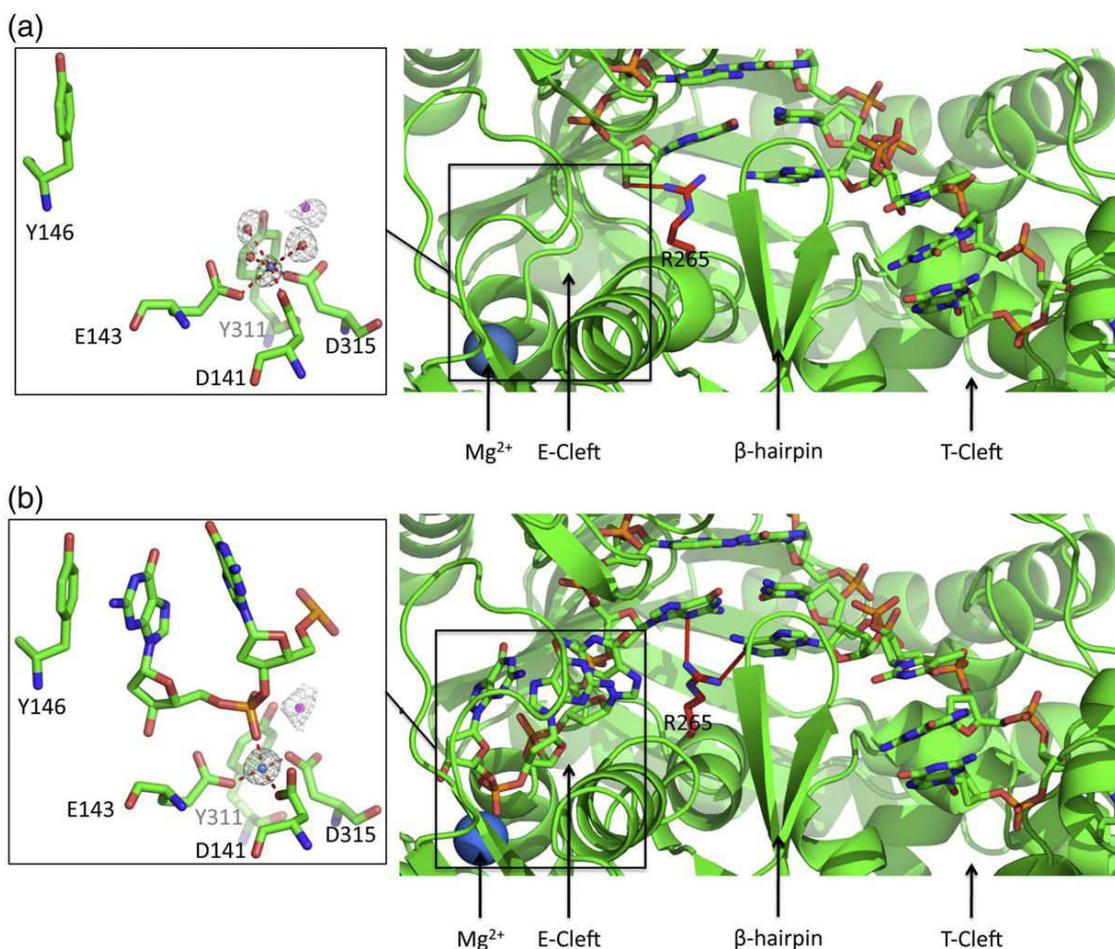
	13H_8	13U_8	15H_8	15U_8	Triple mutant 15U_8
PDB ID	4FLX	4FLY	4FLZ	4FM0	4FM2
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>				
Cell parameters					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	69.21, 114.53, 128.50	69.10, 114.34, 128.58	69.22, 115.11, 128.71	69.61, 114.72, 128.09	68.73, 115.02, 127.26
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution shell	44–2.9	44–2.3	44–3.2	43–3.1	44–2.9
No. of unique reflections	22,777 (3297)	44,918 (6401)	17,526 (2491)	18,760 (2699)	21,903 (3183)
Multiplicity	3.5	4.1	4.0	3.9	3.7
Completeness	98.8 (99.5)	97.8 (97.3)	99.9 (100)	99.5 (99.8)	97.7 (98.8)
$R_{\text{merge}}$ (%)	15.1 (51.2)	14.1 (66.1)	13.0 (64)	13.1 (66.2)	10.4 (86)
$I/\sigma(I)$	8.8 (2.4)	10.7 (2.6)	12.3 (2.1)	8.4 (2.4)	11 (2.0)
$R$ (%)	16.3	17.6	17.5	17.7	18.4
$R_{\text{free}}$ (%)	21.1	21.9	20.4	20.8	22.1
No. of protein atoms	6144	6003	6001	5894	5937
No. of DNA atoms	428	426	470	462	470
No. of water molecules	440	563	66	87	139
$B$ -factor overall (Å <sup>2</sup> )	33.4	32.8	65.7	69.1	74.5
$B$ -factor DNA (Å <sup>2</sup> )	44.6	42.5	98.7	98.4	103.2
Ramachandran outliers ( $n$ )	0	0	0	0	0
RMSD bond lengths (Å)	0.01	0.01	0.01	0.01	0.01
RMSD bond angles (°)	1.12	1.08	1.13	1.11	1.12

Values in parentheses refer to the highest-resolution shell.

0.36 Å; 4830 atoms). Comparison of PabPolB structures with 2VWJ<sup>13</sup> and 2XHB<sup>7</sup> (*T. gorgonarius* PolB with U+4 and H+2, respectively) shows a similar conformation, but the thumb subdomain of TgoPolB is shifted by about 7 Å for 2VWJ and about 3 Å for 2XHB (Fig. S1). These differences may be explained by either packing interactions specific to each protein or a slightly distorted conformation of DNA molecule in the hairpin construct used for TgoPolB. Here, PabPolB was crystallized with a DNA that consists of two separated strands. Compared to other structures of archaeal *apo* PolBs (Pfu,<sup>17</sup> Tgo,<sup>18</sup> and Kod<sup>19</sup>), the thumb subdomain is

shifted toward the exonuclease domain, thereby opening a channel that allows the primer to enter in the editing cleft (E-cleft). Our own unpublished structure of PabPolB apo, solved at 3.2 Å resolution, shows yet another slightly different orientation for the thumb subdomain (rigid-body movement), emphasizing the intrinsic mobility of the C-terminal domain.

In the 13T\_11 structure, the three mismatches inserted at the 3' end of the primer induce a melting of the dsDNA at the P/T junction. Thus, three bases now lie in the E-cleft. Comparison of archaeal PolB in the apo form and in editing mode shows that the  $\beta$ -hairpin motif (residues 240–251) does not adopt



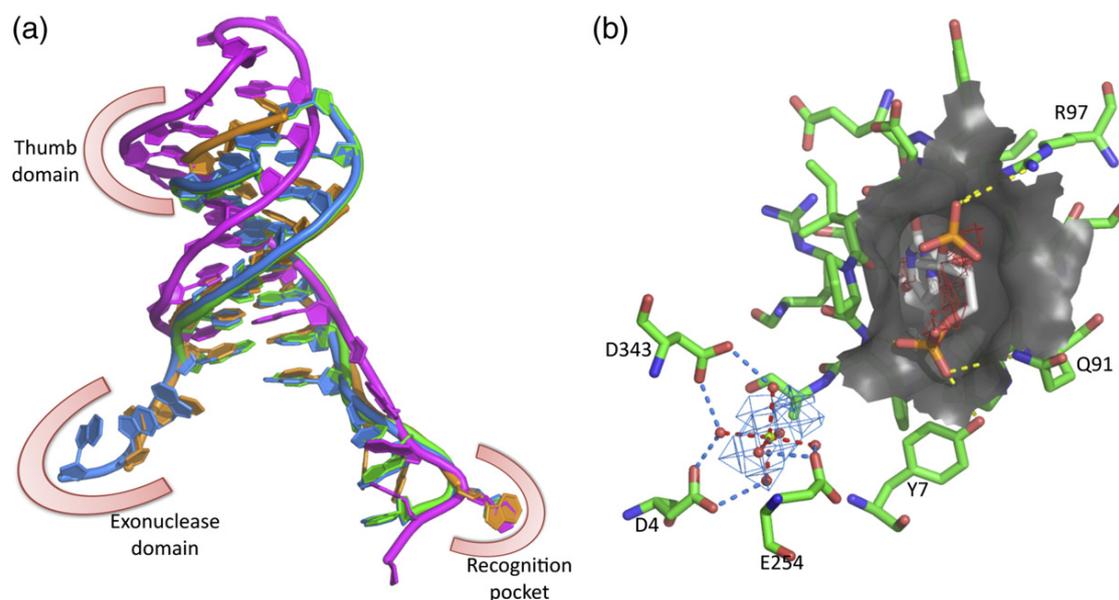
**Fig. 3.** Interactions between PabPolB and DNA duplexes at the P/T junction and in the exonuclease catalytic site. (a) P/T junction in 13U\_8. R265 (depicted in red) makes a direct interaction with the 3' terminus of the primer. In the exonuclease site, only metal A (blue sphere) is present because of the mutation D215A. The coordination sphere of the metal adopts an octahedral conformation when the E-cleft is apo, highlighted with red broken lines. Three acidic residues (D141, E143, and D315) take part in the coordination sphere that is completed by three water molecules. The attacking hydroxide or water molecule is depicted in pink.  $2F_o - F_c$  maps were contoured at  $3\sigma$  for the metal and  $1.5\sigma$  for the water molecules. (b) P/T junction in 13T\_11. R265 (depicted in red) interacts with the last hybridized base of the primer strand and the first non-hybridized of the template, thereby stabilizing the P/T junction if a mismatch is present. In the exonuclease site, metal A (blue sphere) binds the last phosphodiester bond and has switched to a tetrahedral coordination, highlighted with red broken lines. Y146 is stacked against the 3' terminus of the primer. The attacking hydroxide or water molecule is depicted in pink.  $2F_o - F_c$  maps were contoured at  $3\sigma$  for the metal and  $1.5\sigma$  for the water molecule.

an “up” or “down” conformation as seen in the recent structure of RB69 phage DNA polymerase.<sup>20</sup> The tip of this particular motif does not contain the Y255–G256 motif of T4 PolB (Y257–G258 for RB69 PolB) that has been shown to be required for DNA replication fidelity.<sup>21</sup> Nevertheless, the  $\beta$ -hairpin loop and arginine 265 block the P/T junction from two different sides (Fig. 3). This latter residue interacts via salt bridges with both the last base of the hybridized primer and the first base of the single-stranded template, thereby stabilizing the duplex when the primer is present in the E-cleft. When it is absent and the polymerase keeps its editing conformation (13U\_8), R265 makes a hydrogen bond with the 3'-OH of the primer (Fig. 3).

### Primer strand interactions

The mechanism of exonuclease activity normally involves two metal ions as described for PolA<sup>22,23</sup> and PolB,<sup>24</sup> but here, only metal A is seen because the residue involved in stabilizing metal B has been

mutated (D215A). Metal A is coordinated by D141, E143, and D315, in addition to either one of the oxygen phosphate of the last base (13T\_11) or three water molecules (13U\_8). Without DNA in the active site, metal A is in an octahedral coordination state but is switched to a tetrahedral one when the primer is present (Fig. 3). To our knowledge, this is the first experimental characterization of a coordination switch of metal A upon binding of the primer in the exonuclease domain. These observations are compatible with previous experiments. Beese and Steitz have shown that the exonuclease domain of the Klenow fragment is functional in the presence of  $\text{Zn}^{2+}$ ,  $\text{Mn}^{2+}$ , or  $\text{Mg}^{2+}$  but exhibits higher activity when metal A can adopt a tetrahedral conformation.<sup>22</sup> Brautigam *et al.* have demonstrated that in the presence of  $\text{Zn}^{2+}$  and  $\text{Mg}^{2+}$ , metal A is preferentially a  $\text{Zn}^{2+}$ , probably because the energetic penalty of a change from an octahedral to a tetrahedral coordination for a  $\text{Mg}^{2+}$  is greater than that for  $\text{Zn}^{2+}$ .<sup>25,26</sup> The attacking water molecule (or the hydroxide ion) is chelated by Y311, D315, and



**Fig. 4.** Base-flipping and base-checking pocket overview. (a) Superimposition of dsDNA complexed with archaeal PolB. Structures 13U\_8 (green), 13T\_11 (blue), TgoPolB U+4 (magenta, PDB ID 2VWJ), and TgoPolB H+2 (orange, PDB ID 2XHB) have been superimposed. For clarity reasons, only the oligonucleotides have been represented in the sketch. The double-stranded part of 13U\_8 and 13T\_11 is almost superimposable. However, the comparison with TgoPolB's DNAs shows that the double-stranded part does not superimpose well, probably because the thumb domains are slightly shifted (see Fig. S1). The last three bases of the primer in 13T\_11 and the last two bases of TgoPolB H+2 lie in the exonuclease domain. In front of the recognition pocket, the template strand is kinked by about  $80^\circ$  to flip the base out in order to check its identity. High  $B$ -factors are observed on the template strand from the recognition pocket to the P/T junction part for all structures. (b) Base-checking pocket and T-cleft. The nucleotide in front of the recognition (here, a dU in white sticks) is flipped in the cavity (represented with a dark surface). The phosphate immediately 5' is bound by Y7, Q91, and F116N (not represented here), whereas the 3' phosphate makes a salt bridge with R97. Right below the cavity, a fully hydrated divalent metal ion is bound to the protein via five out of six of its water molecules. The map represents the contouring of the predicted cation density at 8.0M as calculated by the electrostatic program AQUASOL<sup>27,28</sup> on the 13U\_8 structure without DNA or metal ions. The positive potential is in red and located at the entrance of the binding pocket. The negative potential is in blue and located around the hydrated metal.

one of the oxygen phosphate at the 3' end of the primer strand. Interestingly, the electron density of this water molecule is present in both structures with or without DNA bound in the E-cleft. At the end of the E-cleft, the 3' end nucleotide is nicely stacked against Y146. As one could expect, no sequence-specific interactions are seen between the polymerase and the ssDNA.

### Template strand interactions

The electronic density of the template strand between the recognition pocket and the P/T junction is poorly defined as reflected by the greater *B*-factors in this region. Some of the bases were only partially modeled (i.e., #4 for 13A\_11, #1 and #4 for 13G\_11). There are indications that the third base (+3 from the P/T junction) may adopt two distinct conformations. In 13T\_11 and 13C\_11, the base is stacked over the fourth base (+2 from the P/T junction) and points to the  $\beta$ -hairpin motif. In the second conformation (13A\_11 and 13G\_11), the base points toward the NH2-terminal domain. In the other structures (i.e., with dU or dH bound), the third base is mainly stacked under the preceding one, but residual electronic density in the  $2F_o - F_c$  map indicates that the base can also adopt the alternative conformation. Interactions of the phosphodiester backbone with the protein are seen at the entrance of the base recognition pocket. The phosphate 5' to the flipped-out base (+4) is hydrogen bonded by Y7, Q91, and F116, whereas its 3' phosphate makes a salt bridge with R97

(Fig. 4). A kink angle of about 80° is observed on the template strand direction in front of the binding pocket. This bending allows the base being checked to bind the recognition pocket in an extra-helical (flipped out) fashion. The electrostatic potential around cavity shows a clear positive peak at the entrance of the cavity (i.e., where the phosphates groups lie) and a negative peak below the pocket, where an additional fully hydrated metal is bound (see below and Fig. 4b).

### The base-checking pocket binds deaminated bases but can also accommodate canonical bases

Superimposition of 13U\_8 PabPolB base-checking pocket with TgoPolB's structure (apo, 1TGO; H+2, 2XHB; U+4, 2VWJ) clearly shows that the binding sites are similar and open even in the apo form (Fig. S1). Moreover, four prolines (P36, P90, P94, and P115) take part in the formation of the pocket, possibly contributing to the cavity's rigidity. In 13A\_11, 13G\_11, 13T\_11, and 13\_C\_11 structures, all protein residues around this cavity occupy the same position except for a few side chains that point outside the cavity and for F116 in 13T\_11 that is slightly shifted to accommodate the 5-methyl group of thymine. The bases are held in place via a network of H bonds that involves mainly backbone amide and carbonyl groups. The interactions between the protein and the recognized base are summarized in Table 2. At the structural level, the network clearly helps stabilize the deaminated bases, whereas

**Table 2.** Summary of the interactions between PabPolB and the base bound in the base-checking pocket

	dU		dC		dT	
	Favorable	Unfavorable	Favorable	Unfavorable	Favorable	Unfavorable
Y37N	O4			N4	O4	
Y37CO	N3, water		N4, N3		N3, water	
P90						C5
V93	Aromatic ring		Aromatic ring		Aromatic ring	
E111CO	water				water	
I114N	O2		O2		O2	
F116						C5
R119	Water				Water	
water	N3				N3	
	dH		dA		dG	
	Favorable	Unfavorable	Favorable	Unfavorable	Favorable	Unfavorable
Y37N	O6			N6	O6	
Y37CO	N1, water		N6, water		N1, water	
V93	Aromatic ring		Aromatic ring		Aromatic ring	
E111CO	Water		Water		Water	N2
I114N	N3				N3	
R119	Water		Water		Water	
Water	N1		N6, N7		N1	

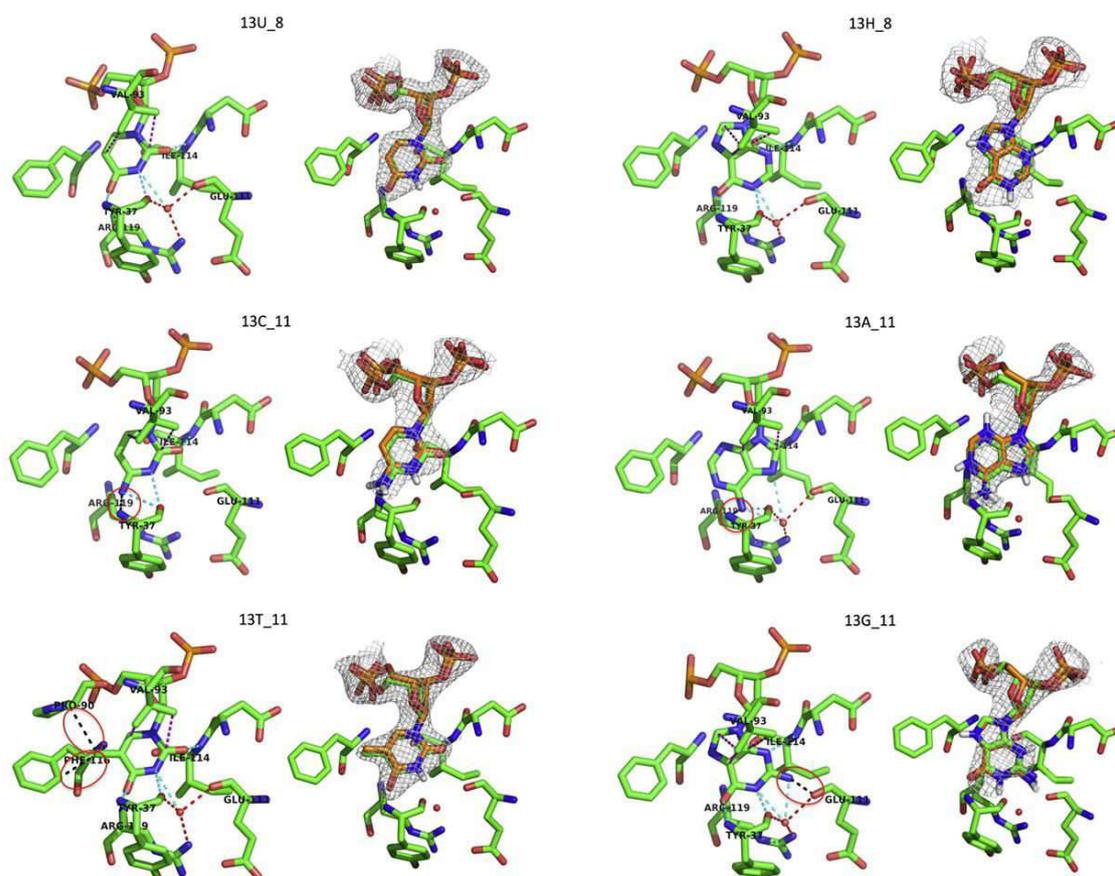
Deaminated bases are stabilized by an extensive H-bond network. For canonical bases, this network is weaker and may display one unfavorable contact from the so-called selectivity filter (Y37 N-atom for dC and dA) or too close a contact (E111 C=O for dG and F116 for dT).

canonical ones can be bound but with at least one less favorable interaction and/or one more unfavorable ones.

Inside the recognition pocket, all the six bases' aromatic rings are strictly superimposable, but the exact arrangement of H-bonds is slightly different (Fig. 5). Since the pocket's walls are rigid, the bases have in each case slightly adapted their position in order to fit in the cavity. A striking example is the adenine's conformation that is bound in *syn* whereas all the other bases are in *anti*. The side chain of V93 points toward the base aromatic ring. A water molecule held by Y37 CO, E111 CO, and R119 side chain contacts all possible bases except for the cytosine where there is no clear electron density for this water in the  $2F_o - F_c$  map. The backbone amide

group of Y37 is used to identify nucleotides having a carbonyl (i.e., O6 in hypoxanthine and guanine, O4 of uracil and thymine). The distance between Y37N and N6 of adenine or N4 of cytosine is quite short (3.1 Å and 3.2 Å, respectively), and this probably induces some strain that impairs strong binding. Moreover, a steric hindrance prevents guanine to fit snugly in the pocket because E111 CO comes close to the N2 atom (2.3 Å). The methyl group of thymine, the only structural difference with uracil, comes in close vicinity of H<sup>β</sup> of P90 and of F116. The side chain of this latter residue has been shifted by about 0.8 Å compared to other structures (Fig. 5).

We then tested the ability of the nucleotides to bind in the recognition pocket using the docking software AutoDock.<sup>29,30</sup> The recognition pocket of



**Fig. 5.** Deaminated and canonical bases in the base-checking pocket. Interactions in the recognition pocket and results of *in silico* docking solutions are represented on the left and right sketches, respectively. Hydrogen bonds between PabPolB and the nucleotide are colored light blue. A water molecule is held in the recognition pocket by hydrogen bonds that are colored red. This water has clear density in all solved structures except 13C\_11. Canonical nucleotides display less tight interactions with the pocket than canonical ones, namely, (i) lack of stabilization by the backbone amide of Y37 (13A\_11 and 13C\_11) and (ii) steric clashes (13G\_11 and 13T\_11). These interactions are depicted with black lines and highlighted with a red circle. Hydrophobic interactions between V93 and the aromatic nucleobase are depicted with purple broken lines. For each base bound in the pocket, the electronic density of  $2F_o - F_c$  map is shown with a contouring at  $1\sigma$  on the right side. The best docking solution is colored orange and in general superimposes well with the crystallographic structure with an rmsd less than 1 Å.

13U\_8 was used as the “receptor” and 10 poses were selected for visual inspection for each base. The best solution exhibits an rmsd lower than 1 Å compared to the crystallographic data (Fig. 5). Lower binding free energies were observed for deaminated bases (dH:  $-8.8 \text{ kcal mol}^{-1}$ , dU:  $-8.4 \text{ kcal mol}^{-1}$ ) than for canonical nucleotides (dC:  $-8.1 \text{ kcal mol}^{-1}$ , dT:  $-7.4 \text{ kcal mol}^{-1}$ , dA:  $-7.6 \text{ kcal mol}^{-1}$ , dG:  $-7.1 \text{ kcal mol}^{-1}$ ). These results are consistent with the crystallographic structures where the interactions between the protein and the bound base are less extensive for the canonical bases than for the deaminated ones.

However, these results cannot explain the higher affinity for deaminated bases observed experimentally with PfuPolB in complex with ssDNA containing dU or dH (9 nM for dU, 31 nM for dH compared to 3000 nM for dT<sup>31</sup>) or with a fully matched dsDNA (4.4 nM for a dU, 182 μM for dT<sup>7</sup>). A factor of 100 in  $K_d$  would require  $\Delta\Delta G = 2.8 \text{ kcal mol}^{-1}$  instead of the  $1 \text{ kcal mol}^{-1}$  suggested by the docking experiments. This difference may arise from the different binding modes for the DNA moiety. In the binding experiments in solution, it is not known if the ssDNA is bound as a template strand (in the T-cleft) or as a primer strand (in E-cleft or in the elongation site), while it is not known for certain in the dsDNA experiments if the polymerase adopts an editing conformation or an elongation conformation. In order to quantify the binding energy difference for the deaminated *versus* natural nucleotide, in situations where only the bound base differs, EMSA experiments have been performed with PabPolB bound to a P/T duplex in editing mode (three mismatches at the 3' end) and containing a dC, a dT, or a dU in +4. Analysis of the proportion of bound DNA showed that a canonical template (dC or dT) is less effective than one containing a deaminated base dU (Fig. 6a). We estimate the difference in binding energy to be about  $1 \text{ kcal mol}^{-1}$ , consistent with the docking experiments and the new structural results presented here.

The lesser stability of natural bases in the recognition pocket is reflected by their greater *B*-factors (Fig. 6b). Single-stranded templates (bases 1 to 5) containing canonical bases, except dT, exhibit the highest *B*-factor deviations. The *B*-factor plot for the template strand clearly exhibits two regimes: for 13A\_11 and 13G\_11 for instance, the *B*-factor deviations are high in the single-stranded part (bases 1 to 5) and then decreases in the double-stranded part (bases 6 to 13). dU- and dH-containing template strands exhibit lower *B*-factors in the single-stranded part of the template strand, especially for bases 1 to 3. Purines bound to the base-checking pocket have the highest deviations of the *B*-factors. Bound pyrimidines have an intermediate behavior.

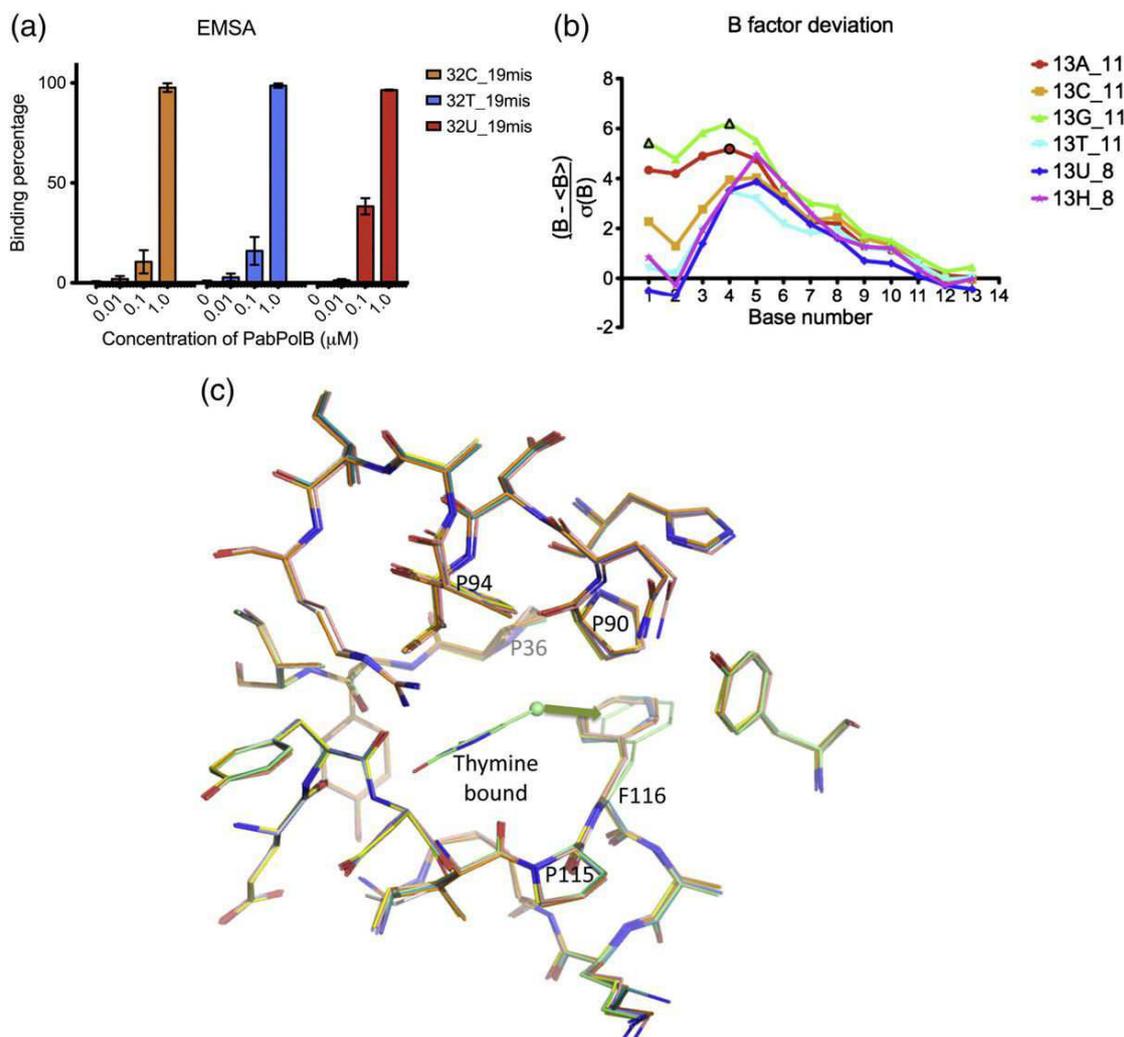
The most mobile one is adenine because of its N6 that is not recognized by Y37. Guanine is the second

most mobile classical base in the pocket: the clash with E111 carbonyl oxygen probably tends to destabilize binding. Cytosine fits better in the pocket, but Y37, which is instrumental in recognizing deaminated bases, brings one bad contact. Thymine has a standard deviation of *B*-factors in the same range as that of uracil or hypoxanthine. Thus, this base is not as mobile as other canonical bases. An explanation for this observation may come from the fact that the room left in the cavity (Fig. 6c) is not big enough to allow thymine to move, consequently lowering its *B*-factor, but in a position leading to an unfavorable binding energy.

### A divalent metal bound to DNA in close proximity to the base-checking pocket

Right below the cavity near the 5' end of the template strand, there is strong density for a metal and its coordination shell, namely, six water molecules arranged in an octahedral geometry. The template strand enters in the protein roughly (Fig. 7) at this point. All the structures presented here exhibit a metal at this position. In the case of our highest-resolution structures (13U\_8 and 13T\_11), the mean distance between the metal and the water molecules is 2.15 Å with a standard deviation of 0.1 Å. The six coordinating water molecules are in contact with protein residues (D4, A5 CO, E251, and D343) and the sulfate moiety of a 4-morpholineethanesulfonic acid (Mes) molecule (Fig. 7). The cyclic part of the Mes molecule is stacked below the first base of the template strand. To investigate the role of this additional metal, we also solved the structure of PabPolB bound to a longer template strand (15 nt instead of 13 nt). The diffraction resolution dropped below 3 Å (Table 1). The first two bases on the template strand adopt an unusual conformation due to the packing interaction with a symmetrical molecule. However, the third base is located in front of the metal, exactly where the Mes molecule was. Because of the lower resolution, the hydration shell of the metal is not as well defined as in the structures with the shorter template. However, the distances between the metal and the acidic residues are unchanged; thus, we may assume that the hydration shell is conserved, except that the density on the phosphate side indicates loss of a ligand water molecule and now the phosphorus atom of the phosphate of the fourth base is situated at 4.6 Å from the metal. Thus, the metal probably interacts directly with the DNA, when the latter is sufficiently long (Fig. 7), while it interacts only indirectly with the protein through its hydration shell.

To test the nature of this ion, we recorded X-ray fluorescence spectra between 5 and 15 keV on a 13U\_8 crystal. No fluorescence signals specific for chromium, manganese, iron, cobalt, nickel, copper,



**Fig. 6.** Deaminated and canonical bases, and functional and structural differences. (a) EMSA analysis of PabPolB binding on DNA duplexes. The primer-templates contain three mismatches at the 3' end of the primer in order to lock the polymerase in editing conformation. The percentage of bound DNA is higher for P/T duplexes containing a deaminated base than for canonical ones. The binding capacity with template strands with a dC or dT in position +4 (in orange and blue, respectively) is about fivefold lower than for a dU (in red). (b) *B*-factor deviations of each base on the template strand compared to the mean *B*-factor of the structure. *B*-factor deviations have been expressed as the number of  $\sigma$  above the mean value of the overall *B*-factor distribution. The bound base for all structures is #2. Due to poor electronic density in the  $2F_o - F_c$ , some of the bases are partially modeled (points depicted with black borders). Templates containing dU or dH have lower *B*-factor deviations on their single-stranded part (bases 1 to 5), whereas bound canonical bases, except thymine, exhibit higher intrinsic thermal agitation. The double-stranded part (bases 6 to 13) has a lower *B*-factor deviation because of the hybridization with the primer strand. (c) Superimpositions of the different pockets of PabPolB in complex with different dsDNAs (yellow, 13U\_8; salmon, 13H\_8; gray, 13A\_11; blue, 13C\_11; orange, 13G\_11; green, 13T\_11). All residues are almost superimposable except for F116 that has been shifted by about 0.8 Å in 13T\_11 (green) to accommodate the 5-methyl group of thymine represented with a green sphere. The aromatic part of the base is shown in lines in the cavity.

or zinc were visible. However, an anomalous map obtained after soaking the crystals of 13U\_8 overnight in a solution containing 2 mM  $MnCl_2$  exhibits two positive peaks larger than  $8\sigma$ : one in the E-cleft (which serves as positive control<sup>24</sup>) and one in the T-cleft, suggesting that the new metal is exchangeable

with manganese (Fig. 7). Taking into account this observation plus the presence of six water molecules in the inner shell and their respective distances ( $2.2 \text{ \AA} \pm 0.21 \text{ \AA}$  from Metal Sites in Proteins database<sup>32</sup> for  $Mg^{2+}$  compared to  $2.15 \text{ \AA} \pm 0.1 \text{ \AA}$  for PabPolB), we suggest that this metal is indeed a magnesium ion.

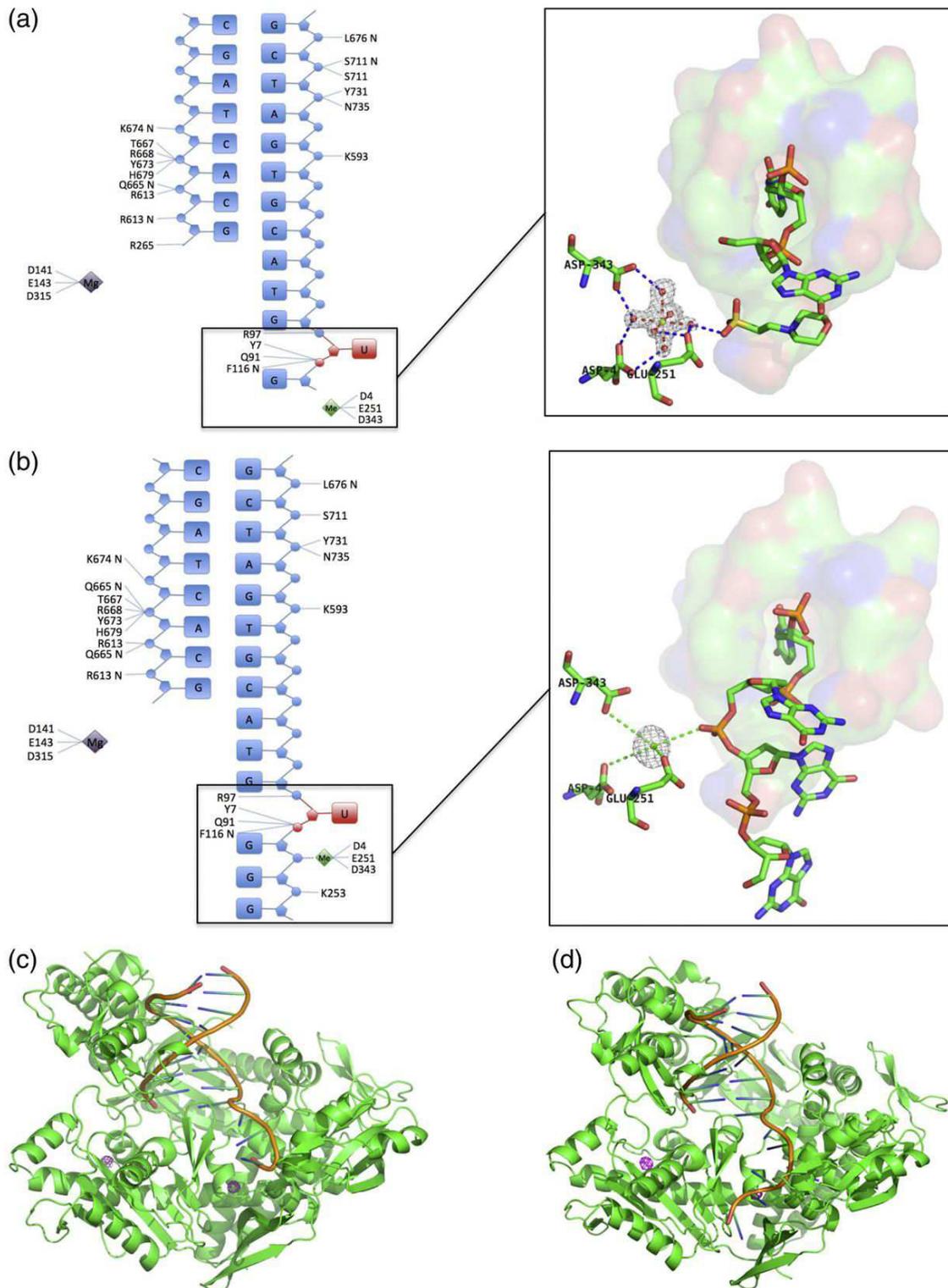


Fig. 7 (legend on next page)

Several structures of other archaeal B family polymerases have been solved. However, to our knowledge, only one contains a metal at this place (PfuPolB, PDB ID 2JGU<sup>16</sup>). Other structures contain a strong peak in the electron density that has been assigned to a water molecule at the same place (TgoPolB PDB ID 1TGO water A/955,<sup>18</sup> *Desulfurococcus* strain Tok B family DNA polymerase PDB ID 1QQC water A/2002<sup>33</sup>).

### Structural and functional characterization of a triple mutant abolishing the extra divalent metal

To investigate the role of this metal in binding the template strand, we have constructed a triple mutant in which all the residues involved in the coordination of the ion by water molecules have been mutated to alanine (D4A, E251A, and E343A). The structure of the triple mutant was solved and indeed showed the absence of the metal in the T-cleft (data not shown). The triple mutant structure superimposes extremely well on the other PabPolB structures: the rmsd using all atoms is 0.266 Å with 15U\_8 (5018 atoms) and 0.286 Å with 13U\_8 (4929 atoms).

We tested the ability of the polymerase lacking the metal to extend a primed DNA with or without a dU on the template strand (Fig. 8). It was observed that the mutant polymerase retains the ability to pause DNA synthesis four bases upstream the P/T junction in the presence of a dU (Fig. 8a). The rate of TLS is similar to that observed for PabPolB exo- (Fig. 8b). Thus, the metal is not mandatory in the polymerase stalling when the template strand is damaged or in the bypass of the lesion. However, in the same reaction conditions used for PabPolB exo+ and PabPolB exo-, the extension rate of the triple mutant is highly reduced (Fig. 8b).

## Discussion

### The additional divalent ion

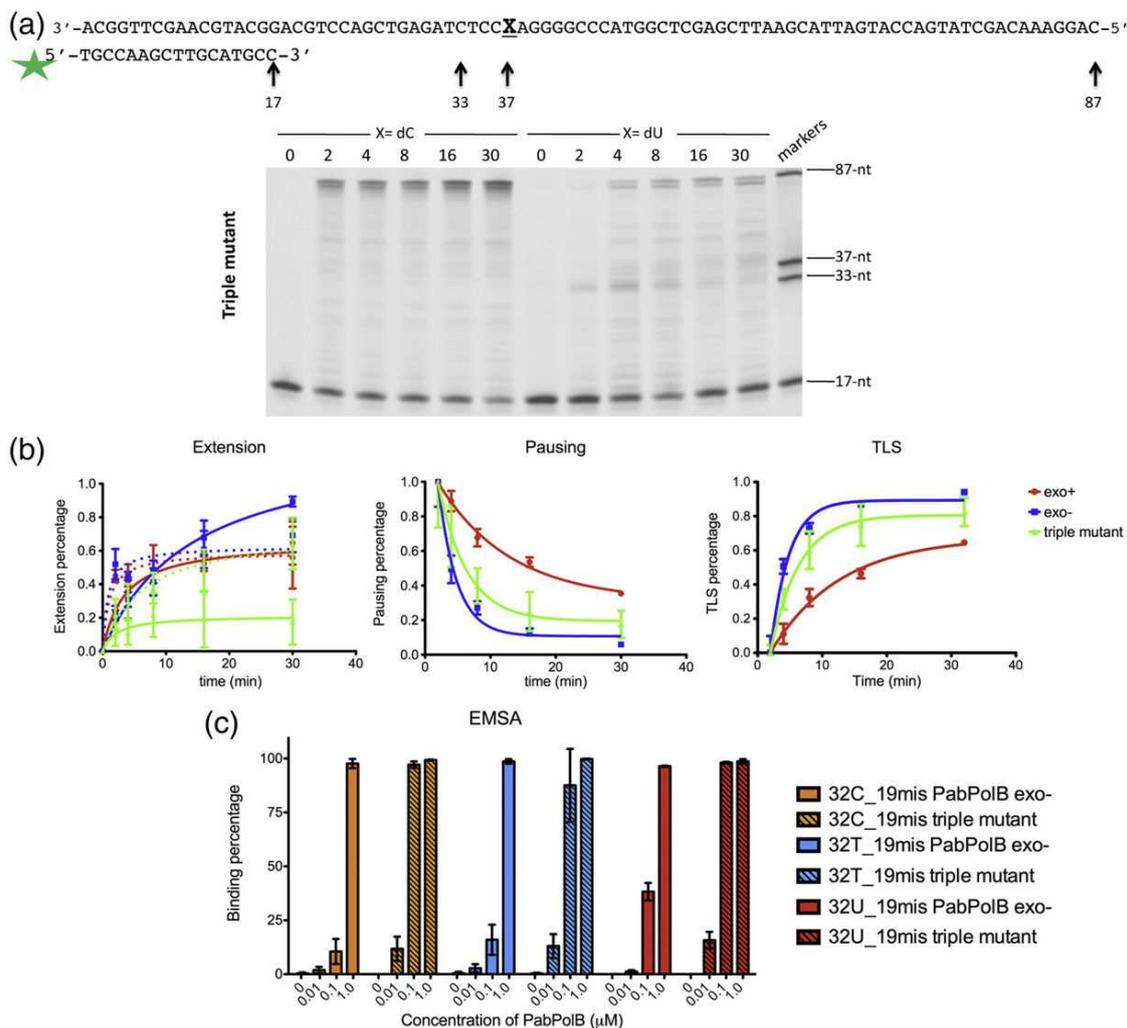
The newly described divalent ion, located in the T-cleft, is apparently conserved in other known

archaeal structures because its side-chain ligands are strictly conserved. To our knowledge, this metal has been described only in the (apo) PfuPolB structure but has been assigned to a water molecule in other crystallographic studies of similar archaeal DNA polymerases. In all our structures, the geometry of the coordination sphere is most compatible with a Mg<sup>2+</sup>. Also, the refined B-factor, which is similar to nearby protein atoms, indicates that the refined ion (Mg<sup>2+</sup>) has the correct number of electrons. The six water molecules surrounding the metal are arranged in a perfect octahedral coordination sphere. This first hydration shell is in contact with three acidic residues of the protein (D4, E251, and D343). In the presence of the template strand of sufficient length, distances between the DNA and the metal indicate a possible direct interaction.

Usually, Mg<sup>2+</sup> interact indirectly (i.e., through their hydration shell) with nucleic acids,<sup>34,35</sup> whereas they are directly coordinated in proteins, for instance, when they participate in catalytic processes. In PabPolB structures presented here, the metal located in the T-cleft is bound in an indirect manner to the protein and so we would predict that it does not have a catalytic activity. We decided to look for a nuclease activity anyway, but cleavage assays of a labeled template strand by the triple mutant or PabPolB exo- construct did not show any degradation (data not shown). Superimposition of the triple mutant structure on other PabPolB structures did not exhibit any differences. Therefore, the metal does not catalyze template strand cleavage or have an influence on PabPolB conformation, at least in the editing mode.

The metal interacts with the template strand at the level of the phosphate in position +5 in the template strand. In Fig. 7c and d, it is ideally placed to help the template strand to find its way in the protein just upstream the base-checking cavity that otherwise induces a sharp kink in the phosphodiester pathway adopted by the DNA. It turns out that PabPolB polymerases lacking this ion are less effective to synthesize DNA but present a higher binding capacity for DNA. At first sight, it might seem surprising that a higher binding constant would lead to a lower extension rate. However, EMSA

Fig. 7. A fully hydrated divalent metal close to the T-cleft. (a) Interactions between 13U\_8 and PabPolB are depicted in the left panel. In the T-cleft, the metal is fully hydrated in all short template structures (i.e., 13 nt). The first hydration shell is in direct contact with the protein (D4, E251, D343). The electronic density in the  $2F_o - F_c$  map around the metal and its hydration shell has been contoured at  $1.2\sigma$ . (b) Interactions between 15U\_8 and PabPolB are depicted in the left panel. In the T-cleft of longer template structures (i.e., 15 nt), the distance between the metal and the DNA is 3.4 Å. Due to a lower resolution, the density around the metal impairs the accurate building of the hydration shell, but the protein-metal distances are unchanged. The electronic density in the  $2F_o - F_c$  map around the metal has been contoured at  $2.0\sigma$ . (c) Anomalous map after 16 h of soaking in stabilization solution containing 2 mM MnCl<sub>2</sub>. Two positive peaks are present (in magenta), one in the E-cleft (on the left) and one in the T-cleft (on the right). Contours of the anomalous electronic density are shown at  $7\sigma$ . (d) The anomalous map calculated after the divalent metal exchange experiment has been superimposed on the structure 15U\_8. The template strand enters in the protein roughly at this point.



**Fig. 8.** Biochemical characterization of the triple mutant lacking the metal near the T-cleft. (a) Extension assays. Primer-templates (25 nM, labeled with Cy5 at the 5' end of the primer) with either a dU or a dC located 16 nt upstream the P/T junction were extended by 75 nM of PabPolB lacking the metal in the T-cleft. The triple mutant stops before the lesion but is also able to extend the primer after the deaminated nucleotide (TLS). (b) Elongation analyses. Intensities of bands were measured and plotted as a function of time to illustrate elongation, stalling, and TLS. The broken lines correspond to extension assays with a canonical template, whereas the continuous lines correspond to the extension assays with a template containing a deaminated base. The blue curve corresponds to PabPolB exo-, while the red one corresponds to PabPolB exo+; the triple mutant is depicted in green. Analyses were carried out as in Fig. 1. The slope of the extension percentage by the triple mutant on a canonical template is reduced compared to PabPolB exo+ and exo-, but the presence of a dU highly impairs the extension. The percentage of pausing of the triple mutant rapidly decreases and is in the range of PabPolB exo-. (c) EMSA analyses of the triple mutant binding to various dsDNA. An increase in binding capacity is observed when the metal is missing in the T-cleft (dashed histograms) compared to the exo- construct (solid histograms). The amount of bound dsDNA is similar whether the bound base is canonical (here dC or dT) or damaged (dU). Moreover, it is higher than the one of the exo- construct for a dU.

experiments assess the equilibrium binding of the protein on DNA whereas the extension assays measure the kinetics of polymerization. The catalytic cycle of processive DNA polymerase implies translocation of the double strand, an event that might be impaired if the binding capacity is too high. These observations lead us to propose that the metal

facilitates the sliding of the polymerase on DNA by reducing direct protein-DNA interactions in the T-cleft and/or guiding the DNA out of the base-checking pocket, on the 5' side. This is in contrast with most known protein-DNA structures where a metal at the protein-DNA interface produces a tighter binding. Interestingly, this magnesium ion is

situated in the close vicinity of the  $\beta$ -hairpin loop involved in stabilizing the editing mode, and E251, one of its ligands, is part of this motif. This ion may therefore be one of the sensors of the template strand upstream region and may participate in the partitioning of the polymerase's activity into elongation or exonuclease active sites.

### The base-checking binding pocket

We confirm the existence of a base recognition pocket located in position +4 from the P/T junction when the polymerase is in the exonuclease mode, as described by Connolly.<sup>36</sup> Previous studies have shown that the presence of this cavity allows PolB to stall DNA replication. This specificity occurs only in archaea and not in the other kingdoms (i.e., bacteria, eukaryotes,<sup>37</sup> or phage<sup>15</sup>). When the pocket is filled by a larger side chain (V93Q), the polymerase loses this capacity.<sup>14,38</sup> The structures of PabPolB in complex with a damaged template strand confirm the switch from the elongation mode to the editing mode upon binding of deaminated bases. Previous studies pointed to a read-ahead mechanism.<sup>15</sup> The binding pocket is rigid and preformed even in the apo form of archaeal PolB (Fig. S1).

Here, we show that the structure of PabPolB in editing mode can also accommodate canonical bases in the rigid binding site. Of course, one may question the lifetime in solution of such a complex. The presence of the three mismatches on the primer strand and the inability to degrade it may have forced the polymerase in an otherwise unstable state. Also, crystallization may have captured complexes that exist in solution only transiently; namely, the base will only shortly visit the cavity and will not stay for long in this conformation in solution. Nevertheless, it is known that switching to the exonuclease mode is one of the two branches of the classical proposed pathway (along with translocation) that the polymerase will follow after every insertion step (see below).

Interactions between the template strand and the protein are of three types: (i) in the T-cleft via the metal described above, (ii) at the entrance of the recognition pocket via the two phosphate groups binding, and (iii) in the thumb subdomain via salt bridges. As a result, a sharp bend of about 80° is generated in front of the cavity. This "bend" is necessary to flip the base in the cavity. The same mechanism is also observed in DNA repair enzymes (i.e., DNA glycosylases<sup>39</sup>) or DNA-modifying enzymes (i.e., methyltransferases<sup>40</sup>). These enzymes need an access to a specific nucleotide buried in a double-stranded context. The following base-flipping features are conserved: the strand containing the base to be checked, repaired, or modified is bent; the phosphodiester dihedral angles flanking the nucleotide are distorted; and a bulky residue takes the place

of the flipped base.<sup>41</sup> In archaeal PolB, the base to be read is in a (different) ssDNA context; therefore, the void left by the extruded base does not have to be filled; on the other hand, the path of the phosphodiester backbone is guided by an additional Mg<sup>2+</sup>. The cavity, rigid, preformed, and located on the path of the template strand, is therefore most probably an obligatory checkpoint for any base before it enters in the polymerase site.

### Modeling the template strand in the elongating mode

There is still one important part of the puzzle that would be needed to fully understand the efficient recognition of deaminated bases: How does the polymerase bind canonical bases in the recognition pocket in elongation mode? Does every single base go through the base-checking pocket in this mode too? And if so, what is its structure? Structural data of archaeal PolB in complex with dsDNA in the polymerase site and a long 5' end template strand would be needed to elucidate this point.

Until recently, the only possible template for modeling the elongation state of a PolB was the bacteriophage RB69 DNA polymerase.<sup>42</sup> However, this polymerase does not possess the base-checking pocket seen in archaeal PolBs. Several variants of the ternary complex have been published, and the one with the longest 5' end of template strand is 3SJJ.<sup>43</sup> Very recently, the binary complex of dsDNA and PfuPolB in the elongation mode (S. A. Wynne *et al.*, manuscript in preparation) has been made available in the PDB (4AIL). The latter one contains a base-checking pocket but no 5' overhang of the template strand. After superimposition of residues 1–300, it can be seen that (i) 4AIL and 3SSJ are indeed very similar and show the same path for their common dsDNA (Fig. S3a) with a large reorientation of the DNA with respect to the complex in the exonuclease mode and (ii) the deaminated base could be made to enter the base-checking pocket only if it were in the +5 position, when the polymerase is in elongation mode. This would effectively represent the *pre-catalytic state* (Fig. S3b). It can therefore be inferred that the difference between a deaminated and normal base manifests itself only after chemical addition of one nucleotide; according to the established classical kinetic pathway of polymerases, this coincides with a choice between two branches, namely, translocating or going to the exonuclease mode. The simplest hypothesis is that dU or dH (but not normal bases) switches the polymerase in exonuclease mode because of some friction in the pocket that prevents translocation from occurring rapidly. This kinetic scheme is well known and documented.<sup>44,45</sup>

This is compatible with a possible mechanism for the switch between the two modes of the

polymerase that was proposed earlier from single-molecule studies showing that a tension on the template strand tends to displace the equilibrium from the polymerase mode to the editing mode.<sup>46</sup> In the case of archaeal PolB, a base still sitting in the pocket may induce strain or friction in the template strand when the polymerase tries to translocate the DNA. This tension may be the signal for the PolB to switch from the elongation mode to the editing mode. Quite logically, the transition between the two modes may therefore constitute the first event to stall the replication fork.<sup>44</sup>

Purely structural data point to the fact that natural bases are more mobile than unnatural ones when engaged in the discrimination pocket. Consistent with this idea, functional data have shown that deaminated nucleotides are slightly more stable in the pocket than canonical ones. However, it is clear that structural data alone are not enough to explain the full net effect of dU-containing template strands, namely, the stalling of the replication fork. To do so, we need to integrate the different kinetic pathways of the DNA polymerase.

### Possible kinetic proofreading

In the docking and EMSA experiments with PabPolB, we have shown that the difference of binding free energy between a canonical and a deaminated base is about  $1 \text{ kcal mol}^{-1}$ . Typically, this value corresponds to about one hydrogen bond and can hardly account for stalling. An "idling" process has been proposed earlier for archaeal DNA polymerases in which the polymerase may switch from polymerase to editing mode upon binding of the deaminated base<sup>6</sup> and then degrade part of the primer strand. As observed in the 13N<sub>11</sub> structures, 3nt of the primer must be melted in order to reach the editing active site. Thus, degradation will start when the deaminated base is in position +1, precisely avoiding that the dU base reaches the templating position.

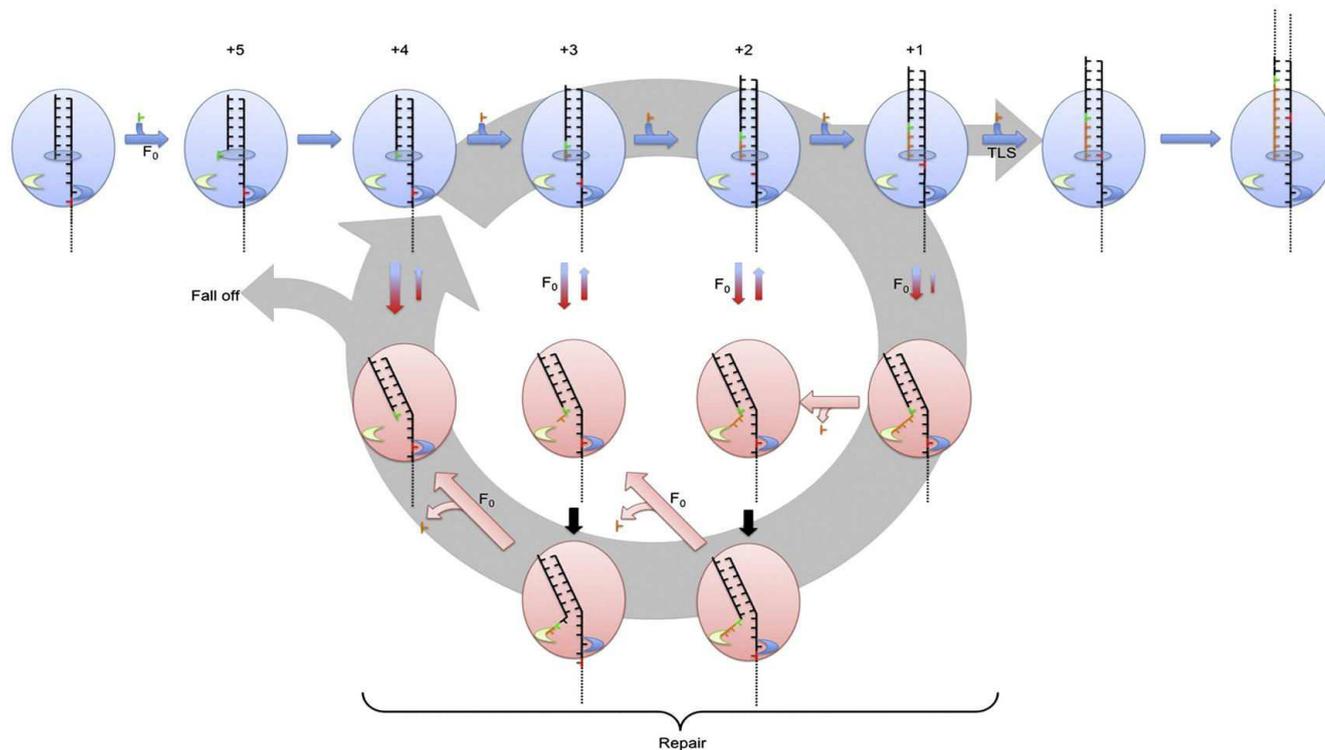
Based on the discussion above, a more detailed idling process can now be proposed (Fig. 9), which takes into account all the known branches of a polymerase catalytic cycle. Comparison of the structure of PfuPolB bound to a fully matched dsDNA (S. A. Wynne *et al.*, manuscript in preparation) and RB69 DNA polymerase with a long 5' end template strand<sup>43</sup> indicates that the deaminated base is first bound when in position +5. After insertion of the incoming nucleotide, this base is now in position +4. Probably due to inefficient translocation, the polymerase will switch to the editing mode, instead of translocating the dsDNA. This event is the first step that allows stalling of the replication. The primer being too short to reach the exonuclease catalytic site, no degradation occurs.

However, the protein may also switch back to the elongation mode because we do observe TLS.

Several factors may be involved in this process, such as thermal activation, elevated local nucleotide concentration,<sup>44</sup> intrinsic rate constant for mode switching,<sup>44,45,47</sup> or "leaking"/escape of the bound damaged base due to a low binding energy difference. This means that, whenever the protein binds a nucleotide in the base-checking pocket, it can be expected to oscillate back and forth between the elongation and the editing mode (idling mode). The rate constant of this switching depends on the polymerase. The ease of the conversion between the proofreading and elongating modes can be estimated by the analysis of the contribution of the lowest-frequency normal modes derived from the simple Elastic Network Model to this conformational transition, as derived by the web server NOMAD-Ref:<sup>48</sup> it turns out that the first 15–20 lowest-frequency normal modes are sufficient to explain more than 90% of this transition in the structurally aligned PabPolB and PfuPolB (data not shown), a recurrent theme in DNA polymerases,<sup>49</sup> implying that this transition occurs easily and frequently in solution. This is also true for the two known conformations of RB69 DNA polymerase in the editing mode (1CLQ) and in the elongation mode (1IG9), where the overlap coefficient between the lowest-frequency mode and the difference vectors between the two forms is very high (0.77).

The second event that allows the polymerase to stall the replication is the possibility to switch to editing mode when the deaminated base is in the positions ranging from +1 to +3.<sup>6,7,13,15</sup> Thus, the protein can achieve a backward movement and can once again oscillate between the editing and elongation modes. The exonuclease activity is enhanced when the damaged nucleotide comes closer to the P/T junction<sup>6,15</sup> but becomes fully effective in position +1 because the primer then physically reaches the exonuclease catalytic site. The polymerase may remove one or several nucleotides in a row as seen for T4<sup>50</sup> and RB69<sup>51</sup> DNA polymerases. Degradation of the primer strand allows the polymerase to reset the position of the deaminated base (from +1 to +4). The whole catalytic cycle may start again. The net incorporation is null and corresponds to the stalling observed experimentally. The deaminated base is not tightly bound and buried in the recognition pocket but rather can move in and out, allowing repair enzymes to handle the damage. The idling process in itself, if it lasts long enough, might be a stress signal for the cell that may trigger DNA repair. However, the precise repair pathway in a replicative context in archaea is unknown.

Experimentally, we observe only one strong pausing at position +4, probably because this is the most stable position of the cycle. When the dU is in position +3 to +1, the protein needs to melt dsDNA in order to bind the deaminated base in the



**Fig. 9.** Idling process and kinetic proofreading mechanism. The polymerase in elongation or in edition mode is depicted in blue and red, respectively. The newly added base is colored orange and the damaged one is in red. The blue arrows represent a nucleotide insertion whereas the pink ones correspond to an excision. The Pol site is represented by the blue oval, the recognition pocket is the “blue moon”, and the exo site is depicted with the “green moon”. The polymerase checks the nucleotide quality in position +5, but the discrimination occurs after addition of the incoming nucleotide, represented here in green. The bound base then is in position +4. The translocation default would tend to switch the polymerase to the edition mode. At each position between +4 and +1, the polymerase can oscillate back and forth between the elongation and the edition mode (bicolor arrow). At position +1, the primer terminus reaches the exonuclease site and can be degraded. The polymerase can then be reset at position +4. Each time the polymerase can bind the deaminated base, a discrimination factor  $F_0$  is observed. For one full cycle, the polymerase discriminates by a factor  $(F_0)^5$  the deaminated base over canonical ones. If the degradation does not occur, TLS is observed. The overall reset of the polymerase corresponds to the idling process, represented by the gray arrow. The net extension of the primer is equal to zero because the polymerase degrades the primer in order to bind the deaminated base at position +4.

dedicated pocket. The energetic penalty for unwinding the dsDNA increases when the deaminated base comes closer to the P/T junction. Once in editing mode with the primer in the E-cleft, the perfectly matched strands can anneal more easily, thereby switching the protein to elongation mode. The intermediate states +3 to +1 are therefore transitory states between the lesion bypass (i.e., *exo-* constructs) and the reset (i.e., *exo+* constructs).

At the purely discrimination level, one can observe that, during one cycle, the polymerase has at least six opportunities to bind the damaged nucleotide (see Fig. 9). If we call  $F_0$  the discrimination factor between the deaminated base and the canonical one, which would be the ratio of binding constant of a dU to any other canonical base,  $K_{dU}/K_{dN}$ , one full cycle may produce a discrimination of  $(F_0)^6$ . In the end, this process is typically a kinetic proofreading mechanism in which several binding events (even at a low discrimination factor) produce a strong macroscopic effect.<sup>52,53</sup> This process allows a better discrimination between bases because the polymerase may go through several idling cycles before translesion occurs and get several chances of putting the deaminated base through the filter of its dedicated pocket; therefore, this phenomenon can stall DNA replication more efficiently and perhaps trigger DNA repair more often.

## Materials and Methods

### Production and purification of PabPolB

PabPolB gene WT has been inserted into a pET28 vector (Novagen) between NdeI and BamHI sites. All the mutations have been achieved using the QuikChange Mutagenesis XL kit from Agilent Technologies. The residue D215 has been mutated to alanine in order to abolish exonuclease activity,<sup>54</sup> subsequently referred to as *exo-* construct. The triple mutant where the additional Mg<sup>2+</sup> binding site was abolished also contains the D215A mutation. For expression, *Escherichia coli* Rosetta 2 cells were transformed with the corresponding plasmid for each of the constructs.

The protocol of production and purification is the same for all constructions. The bacterial cultures were grown to an OD<sub>600</sub> (optical density at 600nm) of 0.6. After addition of 1mM IPTG, cultures were incubated at 30°C overnight (16h) and harvested by centrifugation (4000rpm, 20min, 4°C).

The pellet was resuspended in a solution containing 50mM Tris-HCl, pH7.5, 200mM NaCl, and 9mM imidazole; supplemented with tablets of a protease inhibitor cocktail (Complete, EDTA-free, Roche); lysed by sonication on ice; heated at 55°C for 10min; and centrifuged at 15,000rpm for 30min at 4°C. The supernatant was filtered through a 0.22- $\mu$ m filter and loaded on a 5-mL nickel-affinity resin (HisTrapHP, GE Healthcare). Elution was achieved with 500mM imidazole gradient.

Fractions containing the protein were pooled, dialyzed overnight in buffer A (50mM Tris-HCl, pH7.5, 10mM NaCl, and 2mM MgCl<sub>2</sub>) and then loaded on a 5-mL HiTrap Heparin HP column. Elution was achieved with 2M NaCl gradient. Fractions containing the protein were pooled; dialyzed overnight in a solution containing 50mM Tris-HCl, pH7.5, 500mM NaCl, and 2mM MgCl<sub>2</sub>; and applied to a Superdex™ 200 10/300 GL (GE Healthcare). Fractions containing the proteins were pooled and dialyzed in the final buffer (50mM Tris-HCl, pH7.5, and 2mM MgCl<sub>2</sub>).

Oligonucleotides for mutagenesis and enzymatic activities were purchased from Eurogentec (Seraing, Belgium). The unlabeled and 5'-Cy5 indocarbocyanin®-labeled short oligonucleotides (<87nt) were purified by RP-cartridge or RP-HPLC. The undamaged oligonucleotide (>87nt) or that containing one uracil was purified by PAGE. The oligonucleotide templates were resuspended in TE buffer and annealed to 5'-end-labeled oligonucleotide primers in the presence of 10mM Tris-HCl, pH8.0, and 50mM NaCl at a 1:1 molar ratio by heating to 95°C for 5min and cooling to room temperature. Sequences were checked by GATC Biotech. For crystallization, they were purchased from the same supplier but further purified on RP-HPLC and resuspended in 25mM Tris-HCl, pH7.5, 10mM MgSO<sub>4</sub>, and 0.5mM ethylenediaminetetraacetic acid (EDTA).

### Primer extension assays

For primer extension assays, a 5'-Cy5-labeled 3' 17-nt 5'-TGCCAAGCTTGCATGCC-3' was annealed to an 87-nt template 5'-CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTXTAGATCGACCTGCAGGCATGCAAGCTTGGCA-3', where X is either a dC or a dU located 16nt upstream the P/T junction. Reactions were carried out in 10 $\mu$ L containing 75nM protein, 25nM duplex, 50mM Tris, pH8.8, 1mM DTT, 10mM KCl, 2mM MgCl<sub>2</sub>, and 200 $\mu$ M dNTPs at 55°C at different times indicated on the figure and quenched by adding 15 $\mu$ L of stop buffer (84% formamide, 10mM EDTA, 0.01N NaOH, and 1 $\mu$ M competitor oligonucleotide, the exact complement of the template under study). The reaction products were separated by gel electrophoresis on an 18% polyacrylamide 19:1, 8M urea, analyzed with a Mode Imager Typhoon 9400 (GE Healthcare). Three bands were quantified with ImageQuant 5.2 software and corresponds to A, the band of the nonelongated primer (at 17nt); B, the band of the pause products (at 33nt); and C, the band of the full-length products (at 87nt). Extension ratio is defined as the ratio of extended primer to total products [i.e., (B+C)/(A+B+C)]. Pausing is defined as the ratio of bands of paused products to elongated products [i.e., B/(B+C)]. The ratio is then scaled with respect to the highest proportion of stalled products, namely, the ratio at  $t=2$ min. Translesion ratio is defined as the ratio of extended products past the lesion, thus corresponding to 1-pausing ratio. Experiments were carried out three times for the polymerase, which is able to degrade the primer (*exo+* construct) and the D215A mutant (*exo-* construct), and eight times for the triple mutant. Plots were fitted using Prism5 to a one-step association model for elongation and TLS, and to a one-phase decay for pausing.

### Electrophoretic mobility shift assays

For EMSA, a 5'-Cy5-labeled, 22-nt 5'-GCAGTCCTA-GACGATCACGGGG-3' (19mis) was annealed to a 32-nt template 5'-CATCCGTGGXGTACGTGATCGTCTAG-GACTGC-3', where X is a dC, a dT, or a dU located four bases upstream of the P/T junction. Binding was carried out with 25 nM duplexes and increasing concentrations of exonuclease-deficient PabPolB (indicated on the figure) for 20 min at 25 °C in a buffer (10  $\mu$ L) containing 50 mM Tris-HCl, pH 8, 20 mM NaCl, 2 mM MgCl<sub>2</sub>, 5 mM EDTA, and 0.05 mg/mL bovine serum albumin. The reaction mixture was then loaded with 8% Ficoll type 400 on native gel containing 8% polyacrylamide 29:1 and run in 0.5 $\times$  TB buffer (5.4 g/L Tris-base and 2.75 g/L boric acid) supplemented with 2 mM MgCl<sub>2</sub>. Gels were analyzed with a Mode Imager Typhoon 9400 (GE Healthcare) and processed with ImageQuant 5.2 software. The percentage of protein-DNA complexes corresponds to the ratio of the intensity shifted bands to the intensity of all bands (shifted and non-shifted band). Experiments were realized in replicate:  $n=3$  for the exo- construct,  $n=4$  for the triple mutant.

### Crystallization

For deaminated-base-containing template, the 8-nt-long primer 5'-CGATCACG-3' was annealed to a 13-nt-long template, 5'-GXGTACGTGATCG-3', where X is either a dU or a dH. For a longer template, the same primer sequence was used but was annealed to 15-nt 5'-GGGUGTACGTGATCG-3' or 5'-AAGHGATCGT-GATCG-3'. For the structures with a canonical base in the recognition pocket, the same primer was used (11 nt, 5'-CGATCACGGGG-3') and annealed to a 13-nt template strand (5'-GXGTACGTGATCG-3', where X is dA, dC, dG, or dT).

The protein was incubated with the duplex at a ratio of 1:1.2 for 20 min at 37 °C. Crystals of 13-nt-long matrices grew in 2 days by mixing 2  $\mu$ L of the complexes at 10 mg mL<sup>-1</sup> and 1  $\mu$ L of the crystallization solution (7–12% polyethylene glycol 20000 and 100 mM Mes, pH 6.5). Seeds of these crystals were used to crystallize the complexes of a 15-nt-long template after 16 h of equilibration. Crystals were subsequently cryoprotected by transferring them shortly to a mother liquor solution containing 30% glycerol before flash cooling in liquid nitrogen. Metal exchange assays were carried out by transferring the crystals grown in 2 mM MgCl<sub>2</sub> into a stabilization solution containing 2 mM MnCl<sub>2</sub> for 16 h.

### Diffraction data collection and structure determination

Diffraction data were collected with an ADSC Q210 CCD at beamline ID 14eh1 (European Synchrotron Research Facility, Grenoble, France) and on an ADSC Q315r at beamline Proxima 1 (Synchrotron SOLEIL, Gif-sur-Yvette, France). Data sets were collected at 100 K,  $\lambda=0.93$  Å for 13A\_11, 13C\_11, 13G\_11, 13H\_8, and 15H\_8;  $\lambda=0.98$  Å for 13T\_11, 13U\_8, and 15U\_8; and  $\lambda=1.79$  Å for the divalent metal exchange experiment. All the data were processed using XDS<sup>55</sup> and POINTLESS<sup>56</sup> and were

scaled and merged with SCALA.<sup>57</sup> Data collection statistics are presented in Table 1. Five percent of the reflections were removed from the refinement and kept aside to calculate the  $R_{free}$ .

Molecular replacement was achieved with Phaser<sup>58</sup> using 2VWJ apo as a search model for 13T\_11. Molecular replacement was carried out with this latter structure. Manual building was achieved with Coot.<sup>59</sup> BUSTER-TNT<sup>60</sup> was used to refine the model until convergence of the  $R$ -factors. TLS were chosen with TLSMD<sup>61</sup> and included in the last stages of refinement; five TLS groups were selected for the protein, two TLS groups for the template strand, and one or two for the primer strand (structures with bound dU/dH or bound canonical base, respectively). Anomalous maps were calculated with programs available in CCP4 suite 6.1.1.<sup>62</sup> For each base recognized in the pocket, an omit map was calculated in order to correctly place the aromatic ring. All the refined models were validated with MolProbity.<sup>63</sup> Superimpositions and pictures have been realized with PyMOL.<sup>64</sup>

### Docking experiments and pocket visualization

The recognition pocket of 13U\_8 was chosen as the target receptor, without DNA, using all atoms within a sphere of radius 10 Å around V93 C<sup>B</sup>. Side chains were not allowed to move. Ten poses were calculated for each of the four bases using AutoDock;<sup>29,30</sup> all of them were visually examined after superimposition and the best solution was selected. The semiempirical free energy contains both an enthalpic term and an entropic term. The enthalpic term is made of six different terms (differences between ligand-ligand, protein-protein, and protein-ligand interactions) that include van der Waals, hydrogen bonding, electrostatic, and solvation interactions.

The solvent-accessible surface has been calculated with the program HOLLOW 1.2<sup>65</sup> using a probe radius of 1.4 Å.

### Poisson-Boltzmann-Langevin electrostatics calculations

The program AQUASOL<sup>27,28</sup> was used to compute the solvent, cation, and anion density maps around the structure 13U\_8 without DNA, water, and metals. PDB2PQR<sup>66</sup> and the CHARMM force field were used to assign charges and atom radii to every atom (including experimentally invisible hydrogens), respectively. The pH used for the calculation was set to 7.

The resulting pqr file was given as an input for AQUASOL. The solute was embedded in a cubic grid of 129 points in each direction equally spaced by 1 Å. A trilinear interpolation protocol was used to spread the charges of atoms over the grid. The solvent region was defined as being outside the solvent-accessible surface that has been computed by rolling a probe of radius 1.4 Å on the atoms of the structure. Water is defined by its dipolar moment (set to 2.35 D) and its bulk concentration of 55 M. Monovalent cation and anion concentrations were set to a bulk concentration of 200 mM. Their radius was set to 1.4 Å. The electrostatic potential was set to zero on the borders of the grid. AQUASOL self-consistently solved the nonlinear Poisson-Boltzmann-Langevin equation.<sup>28</sup> Maps were calculated for the electrostatic potential, ion,

and water densities. Subsequent anion and cation density maps were extracted from a combination of ion density and ion valency.

#### Normal mode calculations and overlap coefficient

Overlap coefficients were calculated as previously described<sup>49</sup> using the elastic network and CA-only models. The cutoff radius used for spring definition was 10 Å. All calculations were made using the NOMAD-Ref web server.<sup>48</sup>

#### Accession numbers

The accession numbers of the structures are the following: 13A\_11, 4FLT; 13C\_11, 4FLU; 13G\_11, 4FLV; 13T\_11, 4FLW; 13H\_8, 4FLX; 13U\_8, 4FLY; 15H\_8, 4FLZ; 15U\_8, 4FM0; 13U\_8 soaked in MnCl<sub>2</sub>, 4FM1; triple mutant, 4FM2. These accession numbers are listed in Table 1.

#### Acknowledgements

We thank the staff of Proxima1 in Soleil (Orsay) for help in data collection, as well as the staff from the European Synchrotron Research Facility (Grenoble). We thank Patrice Koehl for help in energy calculation and minimization. We thank Frederic Poitevin for help with AutoDock and AQUASOL (electrostatic) calculations. Thanks are due to Pierre Beguin for help in cloning and to Ahmed Haouz and PF6 (IP) for help in crystallization experiments. C.R. was financially supported by Ifremer and the Brittany Regional Council. G.H. was supported by grant ANR-10-JCJC-1501 01 from the National Research Agency.

#### Supplementary Data

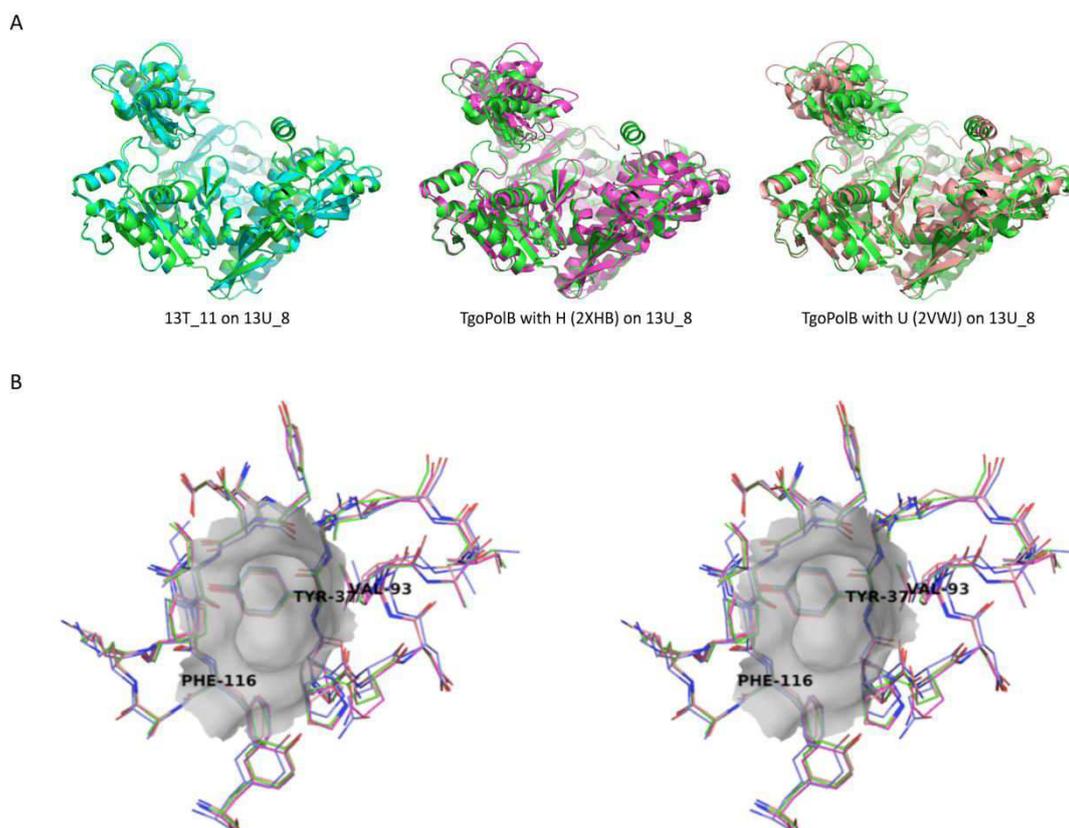
Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2012.07.025>

#### References

- Lindahl, T. & Nyberg, B. (1974). Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405–3410.
- Kow, Y. W. (2002). Repair of deaminated bases in DNA. *Free Radic. Biol. Med.* **33**, 886–893.
- Krokan, H. E., Standal, R. & Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. *Biochem. J.* **325**, 1–16.
- Seeberg, E., Eide, L. & Bjørås, M. (1995). The base excision repair pathway. *Trends Biochem. Sci.* **20**, 391–397.
- Kiyonari, S., Tahara, S., Uchimura, M., Shirai, T., Ishino, S. & Ishino, Y. (2009). Studies on the base excision repair (BER) complex in *Pyrococcus furiosus*. *Biochem. Soc. Trans.* **37**, 79.
- Russell, H. J., Richardson, T. T., Emptage, K. & Connolly, B. A. (2009). The 3′–5′ proofreading exonuclease of archaeal family-B DNA polymerase hinders the copying of template strand deaminated bases. *Nucleic Acids Res.* **37**, 7603–7611.
- Killelea, T., Ghosh, S., Tan, S. S., Heslop, P., Firbank, S. J., Kool, E. T. & Connolly, B. A. (2010). Probing the interaction of archaeal DNA polymerases with deaminated bases using x-ray crystallography and non-hydrogen bonding isosteric base analogues. *Biochemistry*, **49**, 5772–5781.
- Shuttleworth, G., Fogg, M. J., Kurpiewski, M. R., Jen-Jacobson, L. & Connolly, B. A. (2004). Recognition of the pro-mutagenic base uracil by Family B DNA polymerases from archaea. *J. Mol. Biol.* **337**, 621–634.
- Langhorst, B. W., Jack, W. E., Reha-Krantz, L. & Nichols, N. M. (2012). Polbase: a repository of biochemical, genetic and structural information about DNA polymerases. *Nucleic Acids Res.* **40**, D381–D387.
- Erauso, G., Reysenbach, A.-L., Godfroy, A., Meunier, J.-R., Crump, B., Partensky, F. *et al.* (1993). *Pyrococcus abyssi* sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**, 338–349.
- Dietrich, J., Schmitt, P., Zieger, M., Preve, B., Rolland, J.-L., Chaabihi, H. & Gueguen, Y. (2002). PCR performance of the highly thermostable proof-reading B-type DNA polymerase from *Pyrococcus abyssi*. *FEMS Microbiol. Lett.* **217**, 89–94.
- Palud, A., Villani, G., L'Haridon, S., Querellou, J., Raffin, J.-P. & Henneke, G. (2008). Intrinsic properties of the two replicative DNA polymerases of *Pyrococcus abyssi* in replicating abasic sites: possible role in DNA damage tolerance? *Mol. Microbiol.* **70**, 746–761.
- Firbank, S. J., Wardle, J., Heslop, P., Lewis, R. J. & Connolly, B. A. (2008). Uracil recognition in archaeal DNA polymerases captured by x-ray crystallography. *J. Mol. Biol.* **381**, 529–539.
- Fogg, M. J., Pearl, L. H. & Connolly, B. A. (2002). Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat. Struct. Mol. Biol.* **9**, 922–927.
- Greagg, M. A., Fogg, M. J., Panayotou, G., Evans, S. J., Connolly, B. A. & Pearl, L. H. (1999). A read-ahead function in archaeal DNA polymerases detects pro-mutagenic template-strand uracil. *Proc. Natl Acad. Sci. USA*, **96**, 9045–9050.
- Emptage, K., O'Neill, R., Solovyova, A. & Connolly, B. (2008). Interplay between DNA polymerase and proliferating cell nuclear antigen switches off base excision repair of uracil and hypoxanthine during replication in archaea. *J. Mol. Biol.* **383**, 762–771.
- Kim, S. W., Kim, D.-U., Kim, J. K., Kang, L.-W. & Cho, H.-S. (2008). Crystal structure of Pfu, the high fidelity DNA polymerase from *Pyrococcus furiosus*. *Int. J. Biol. Macromol.* **42**, 356–361.
- Hopfner, K.-P., Eichinger, A., Engh, R. A., Laue, F., Ankenbauer, W., Huber, R. & Angerer, B. (1999). Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl Acad. Sci. USA*, **96**, 3600–3605.
- Kuroita, T., Matsumura, H., Yokota, N., Kitabayashi, M., Hashimoto, H., Inoue, T. *et al.* (2005). Structural mechanism for coordination of proofreading and

- polymerase activities in archaeal DNA polymerases. *J. Mol. Biol.* **351**, 291–298.
20. Aller, P., Duclos, S., Wallace, S. S. & Doublé, S. (2011). A crystallographic study of the role of sequence context in thymine glycol bypass by a replicative DNA polymerase serendipitously sheds light on the exonuclease complex. *J. Mol. Biol.* **412**, 22–34.
  21. Subuddhi, U., Hogg, M. & Reha-Krantz, L. J. (2008). Use of 2-aminopurine fluorescence to study the role of the beta hairpin in the proofreading pathway catalyzed by the phage T4 and RB69 DNA polymerases. *Biochemistry*, **47**, 6130–6137.
  22. Beese, L. S. & Steitz, T. A. (1991). Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* **10**, 25–33.
  23. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R. & Steitz, T. A. (1988). Cocystal structure of an editing complex of Klenow fragment with DNA. *Proc. Natl Acad. Sci. USA*, **85**, 8924–8928.
  24. Shamoo, Y. & Steitz, T. A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell*, **99**, 155–166.
  25. Brautigam, C. A. & Steitz, T. A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **8**, 54–63.
  26. Brautigam, C. A., Sun, S., Piccirilli, J. A. & Steitz, T. A. (1999). Structures of normal single-stranded DNA and deoxyribo-3′-S-phosphorothiolates bound to the 3′–5′ exonucleolytic active site of DNA polymerase I from *Escherichia coli*. *Biochemistry*, **38**, 696–704.
  27. Azuara, C., Orland, H., Bon, M., Koehl, P. & Delarue, M. (2008). Incorporating dipolar solvents with variable density in Poisson–Boltzmann electrostatics. *Biophys. J.* **95**, 5587–5605.
  28. Koehl, P. & Delarue, M. (2010). AQUASOL: an efficient solver for the dipolar Poisson–Boltzmann–Langevin equation. *J. Chem. Phys.* **132**, 064101.
  29. Trott, O. & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461.
  30. Seeliger, D. & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and AutoDock/Vina. *J. Comput.-Aided Mol. Des.* **24**, 417–422.
  31. Gill, S., O’Neill, R., Lewis, R. J. & Connolly, B. A. (2007). Interaction of the family-B DNA polymerase from the Archaeon *Pyrococcus furiosus* with deaminated bases. *J. Mol. Biol.* **372**, 855–863.
  32. Hsin, K., Sheng, Y., Harding, M. M., Taylor, P. & Walkinshaw, M. D. (2008). MESPEUS: a database of the geometry of metal sites in proteins. *J. Appl. Crystallogr.* **41**, 963–968.
  33. Zhao, Y., Jeruzalmi, D., Moarefi, I., Leighton, L., Lasken, R. & Kuriyan, J. (1999). Crystal structure of an archaeobacterial DNA polymerase. *Structure*, **7**, 1189–1199.
  34. Misra, V. K. & Draper, D. E. (1998). On the role of magnesium ions in RNA stability. *Biopolymers*, **48**, 113–135.
  35. Bandyopadhyay, D. & Bhattacharyya, D. (2003). Different modes of interaction between hydrated magnesium ion and DNA functional groups: database analysis and ab initio studies. *J. Biomol. Struct. Dyn.* **21**, 447–458.
  36. Connolly, B. A. (2009). Recognition of deaminated bases by archaeal family-B DNA polymerases. *Biochem. Soc. Trans.* **37**, 65.
  37. Wardle, J., Burgers, P. M. J., Cann, I. K. O., Darley, K., Heslop, P., Johansson, E. *et al.* (2008). Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res.* **36**, 705–711.
  38. Horváth, A. & Vértessy, B. G. (2010). A one-step method for quantitative determination of uracil in DNA by real-time PCR. *Nucleic Acids Res.* **38**, e196–e.
  39. Banerjee, A., Yang, W., Karplus, M. & Verdine, G. L. (2005). Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature*, **434**, 612–618.
  40. Cheng, X. & Roberts, R. J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Res.* **29**, 3784–3795.
  41. Stivers, J. T. (2004). Site-specific DNA damage recognition by enzyme-induced base flipping. *Prog. Nucleic Acid Res. Mol. Biol.* **77**, 37–65.
  42. Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997). Crystal structure of a pol [alpha] family replication DNA polymerase from bacteriophage RB69. *Cell*, **89**, 1087–1099.
  43. Xia, S., Wang, M., Blaha, G., Konigsberg, W. H. & Wang, J. (2011). Structural insights into complete metal ion coordination from ternary complexes of B family RB69 DNA polymerase. *Biochemistry*, **50**, 9114–9124.
  44. Kunkel, T. A. & Bebenek, K. (2000). DNA replication fidelity. *Annu. Rev. Biochem.* **69**, 497–529.
  45. Reha-Krantz, L. J. (2010). DNA polymerase proofreading: multiple roles maintain genome stability. *Biochim. Biophys. Acta*, **1804**, 1049–1063.
  46. Ibarra, B., Chemla, Y. R., Plyasunov, S., Smith, S. B., Lázaro, J. M., Salas, M. & Bustamante, C. (2009). Proofreading dynamics of a processive DNA polymerase. *EMBO J.* **28**, 2794–2802.
  47. Fidalgo da Silva, E. & Reha-Krantz, L. J. (2007). DNA polymerase proofreading: active site switching catalyzed by the bacteriophage T4 DNA polymerase. *Nucleic Acids Res.* **35**, 5452–5463.
  48. Lindahl, E., Azuara, C., Koehl, P. & Delarue, M. (2006). NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res.* **34**, W52–W56.
  49. Delarue, M. & Sanejouand, Y.-H. (2002). Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. *J. Mol. Biol.* **320**, 1011–1024.
  50. Reddy, M. K., Weitzel, S. E. & von Hippel, P. H. (1992). Processive proofreading is intrinsic to T4 DNA polymerase. *J. Biol. Chem.* **267**, 14157–14166.
  51. Hogg, M., Aller, P., Konigsberg, W., Wallace, S. S. & Doublé, S. (2007). Structural and biochemical investigation of the role in proofreading of a beta hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the B family. *J. Biol. Chem.* **282**, 1432–1444.

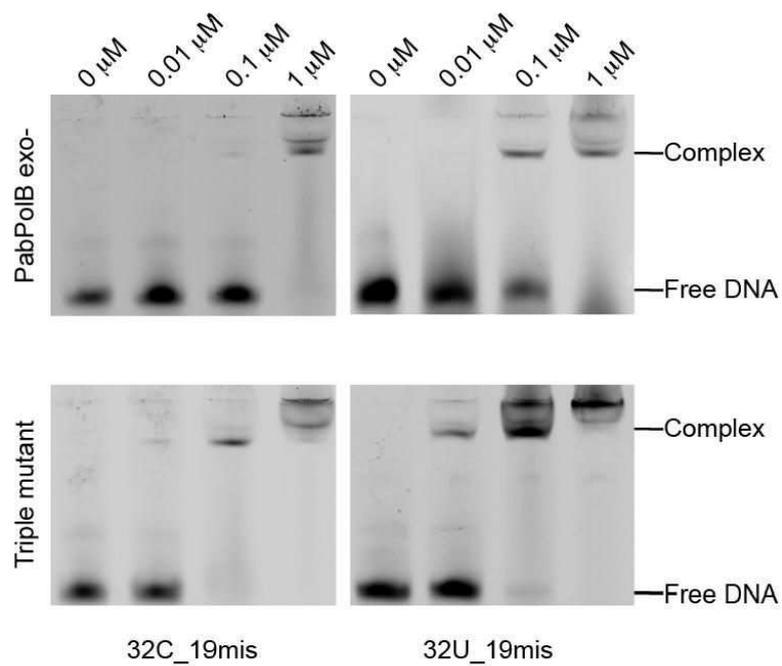
52. Hopfield, J. J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl Acad. Sci. USA*, **71**, 4135–4139.
53. Ninio, J. (1975). Kinetic amplification of enzyme discrimination. *Biochimie*, **57**, 587–595.
54. Evans, S. J., Fogg, M. J., Mamone, A., Davis, M., Pearl, L. H. & Connolly, B. A. (2000). Improving dideoxynucleotide-triphosphate utilisation by the hyperthermophilic DNA polymerase from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* **28**, 1059–1066.
55. Kabsch, W. (2010). XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 125–132.
56. Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 282–292.
57. Evans, P. (2005). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
58. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
59. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
60. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W. *et al.* (2011). BUSTER version 2.11.2 Global Phasing Ltd., Cambridge, United Kingdom.
61. Painter, J. & Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 439–450.
62. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50** (1994), 760–763.
63. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J. *et al.* (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 12–21.
64. DeLano, W.L. (2002). The PyMOL Molecular Graphics System. [www.pymol.org](http://www.pymol.org).
65. Ho, B. K. & Gruswitz, F. (2008). HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct. Biol.* **8**, 49.
66. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667.



**Figure S1: Superimposition of PabPolB complexes with other archaeal PolB.**

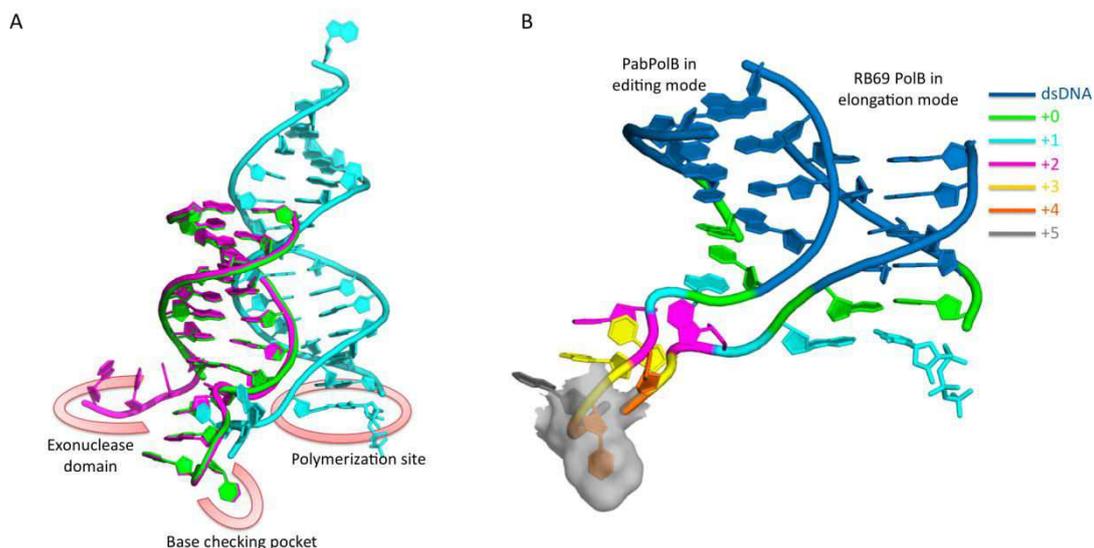
A. Superimposition of 13U\_8 (green) with 13T\_11 (blue), Tgo PolB with dH in position +2 (2XHB, magenta) or dU+4 (2VWJ, salmon). For clarity the DNA has been omitted. Whereas the thumb subdomain is almost superimposable for all structures concerning PabPolB it is found to be shifted by about 3 Å for TgoPolB bound to dH and 7 Å for TgoPolB bound to dU.

B. Stereo view of the base-checking pocket of archaeal PolBs. The recognition sites have been superimposed. The sketch is a superimposition of 13U\_8 (green), Tgo in apo conformation (pdb id 1TGO, dark blue) bound to uracil (pdb id 2VWJ, salmon) or to hypoxanthine (pdb id 2XHB, magenta). The pockets have been highlighted using a surface representation. This clearly shows that the pocket is rigid, even in the apo form. Principal residues are numbered on the figure.



**Figure S2: Examples of EMSA gels**

All duplexes used in binding assays contain a mismatch of 3 bases at the 3'-end of the primer in order to lock the polymerase in editing conformation. 25 nM of each duplexes have been mixed with increasing concentrations of either PabPolB exo- (upper panel) or the triple mutant lacking the metal in the T-cleft (lower panel). The binding capacity of the triple mutant is higher than for the exo- construct.



**Figure S3: Comparison of DNA involved in binary complexes with PolB**

A. PabPolB structures in edition mode (13U\_8 in green, 13T\_11 in magenta) have been superimposed on PfuPolB and RB69 in elongation mode (pdb id 4AIL in dark blue, pdb id 3SJJ in cyan respectively). For clarity reasons only the DNAs have been represented. The localization of DNA is different if the polymerase is in elongation mode or in editing mode.

B. 13U\_8 structure has been superimposed on RB69 polB in polymerization mode (pdb id 3SJJ). For clarity reasons only the DNAs are represented here and the upper part of the duplexes have been omitted. Each base has been colored depending on its position from the P/T junction as depicted in the upper right panel. Before the reaction occurs the incoming nucleotide is hybridized with the base in position +1 (depicted in cyan in the RB69 PolB structure). Therefore the base-checking pocket would bind the nucleotide in position +5 (not seen here because RB69 PolB lacks the cavity). After the reaction, the bound nucleotide is in position +4. If the polymerase can't translocate the DNA, it switches in editing mode (depicted with the structure of PabPolB here). The base sitting in the cavity is therefore in position +4 as observed structurally and functionally.



# Troisième partie: Résultats supplémentaires

---

Le rôle de l'ion métallique divalent tout récemment identifié par cristallographie au sein du domaine N-terminal de l'ADN polymérase répliquative de la famille B de *Pyrococcus abyssi* reste aujourd'hui encore mal compris. Afin de déceler une différence de propriétés fonctionnelles de PabPolB triple mutée par rapport à PabPolB sauvage, diverses réactions enzymatiques supplémentaires ont été effectuées.

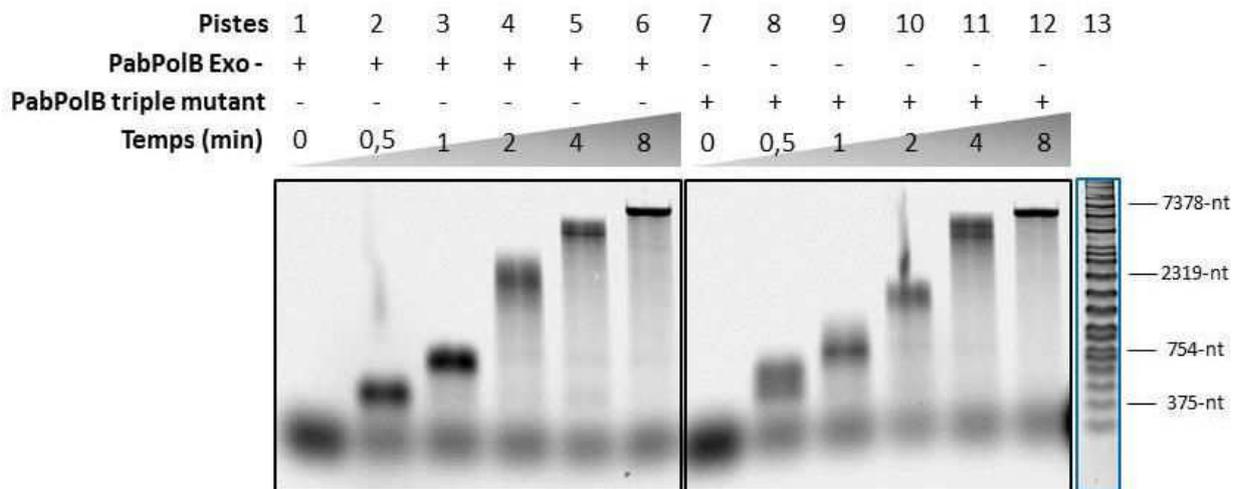
## **I. Matériels et méthodes supplémentaires : Réaction de processivité**

L'ADN M13mp18(+) circulaire simple brin de 7 250 bases a été utilisé en tant que matrice. La matrice a été hybridée à une amorce de 32 nucléotides marquée en 5' avec l'indocarbocyanin® Cy-5. Les réactions d'extension ont été effectuées *in vitro* en présence de 7 nM d'amorce /matrice (32/M13mp18) et de 200 nM d'enzyme PabPolB dans son tampon réactionnel (50 mM Tris HCl pH 8,8, 1 mM dithiothreitol, 10 mM KCl, 5 mM MgCl<sub>2</sub> ou CaCl<sub>2</sub> en fonction des expériences) et initiées par l'ajout d'un mélange de dNTPs (200 μM) et de « trap » héparine (de 10<sup>-2</sup> à 10<sup>-3</sup> mg/mL). Le mélange réactionnel a été incubé à 60°C pendant différents temps. Pour les temps supérieurs à 30 min d'incubation, de l'huile minérale a été déposée au-dessus du mélange réactionnel afin de pallier à l'évaporation de l'échantillon. La réaction d'extension a été arrêtée sur glace par addition de 10 μL de tampon inhibiteur (96% formamide, 15 mM EDTA). Dans le but de piéger la pol lors de son détachement pendant un cycle d'extension, l'héparine a été préincubée avec la pol pendant 8 min à 60°C avant l'ajout du mélange réactionnel. Une réaction de contrôle sans piégeage de l'enzyme a été réalisée en l'absence d'héparine. Les produits d'extension obtenus ont été dénaturés à 100°C pendant 5 min. Ils ont été ensuite séparés en fonction de leur taille par électrophorèse en gel d'agarose à 1 % (p/V) alcalin, dans un tampon de migration NaOH (0,05 N) et EDTA (1 mM) pH 8 sous l'influence d'un champ électrique (30 V constant) pendant 14 à 17 h à 4°C. L'intensité de fluorescence émise par les produits d'extension marqués, après excitation à la longueur d'onde de 633 nm, a été mesurée à 670 nm par Mode Imager Typhoon 9400 (GE Healthcare). En parallèle, un marqueur de taille d'ADN (Raoul™, 200 - 48 000 pb, Qbiogene) a été déposé sur le gel afin de déterminer la taille des fragments d'ADN obtenus puis révélé séparément dans un bain tamponné à pH 8 de SYBRGold (Invitrogen) pendant 30 min sous agitation à 25°C. Son intensité de fluorescence, après excitation à la longueur d'onde de 488 nm, a été mesurée à 532 nm.

## II. Résultats supplémentaires

### II.1. Efficacité d'extension de PabPolB triple mutée

Des réactions d'extension d'amorces hybridées à une longue matrice d'ADN de 7 250 nucléotides, soit environ 80 fois plus longue que la matrice d'ADN utilisée dans les expériences précédentes, ont été effectuées dans le but d'apprécier une différence de capacité d'extension lors de temps courts d'incubation. Nous pouvons remarquer que PabPolB triple mutée (D4A, E251A, E343A) a le même profil de cinétique d'extension d'une longue matrice que PabPolB sauvage (Figure 22). Ainsi, à 8 min d'incubation, les produits d'extension de longueur maximale, c'est-à-dire de 7 250 nucléotides, apparaissent déjà et ceci pour les deux polymérases (Figure 22, pistes 6 et 12). Ces résultats suggèrent que l'ion métallique divalent nouvellement identifié n'interviendrait pas dans la rapidité de PabPolB à polymériser de longs fragments d'ADN.



**Figure 22. Cinétique d'extension d'amorce hybridée à une matrice ADN de 7 250 nucléotides par PabPolB sauvage et PabPolB triple mutée.**

Profil d'électrophorèse en gel dénaturant d'agarose à 1 % alcalin des produits des réactions enzymatiques effectuées en présence de 7 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 hybridée à la matrice M13mp18(+), de 200  $\mu$ M de dNTPs et de 200 nM de chaque PabPolB indiquée. Les mélanges réactionnels ont été incubés à différents temps indiqués au-dessus du gel à 60°C. La piste 13 comporte le marqueur de taille d'ADN

## II.2. Tolérance du dommage uracile et contexte nucléotidique

PabPolB *exo-* a été mise en présence de deux types de matrices de 87 nucléotides (nt) contenant une base uracile à des positions différentes. Les environnements nucléotidiques des duplexes amorce/matrice (17/87) formés, notés dU +16 et dU +20 sont respectivement les suivants :

dU +16 : 5'-CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTUTAGAGTCGACCTGCAGG  
CATGCAAGCTTGGCA-3'

dU +20 : 5'-CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGAUCCTCTAGAGTCGACCTGCAGG  
CATGCAAGCTTGGCA-3'

(La base notée en gras soulignée correspond à la base lue quatre bases en amont du dommage uracile, la base en rouge est l'uracile).

Il est intéressant de noter ici que l'intensité de la pause de polymérisation par PabPolB, quatre bases avant le dommage uracile, située respectivement en position 33 pour le dommage dU +16 et en position 37 pour le dommage dU +20, est différente selon la matrice d'ADN utilisée (Figure 23, comparer par exemple bande 29, piste 6 et bande 33, piste 12) et plus particulièrement le contexte nucléotidique qui environne le dommage ponctuel uracile. Ainsi, PabPolB semble effectuer une pause dans la synthèse d'ADN quatre bases en amont du dommage de façon plus forte dans le contexte dU +16, et de ce fait, freine la synthèse complète du brin complémentaire (bande de 87-nt moins intense, comparer la piste 6 à la piste 12).

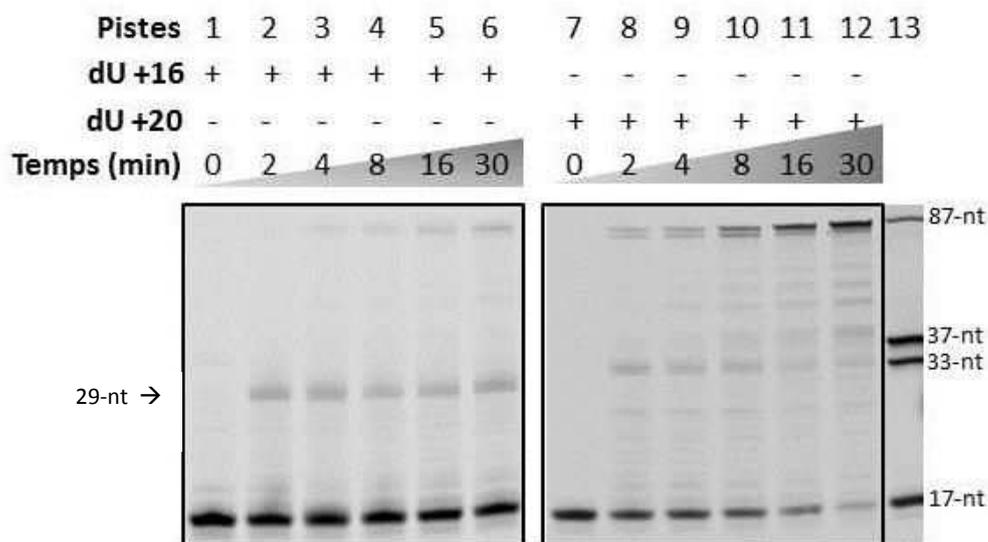
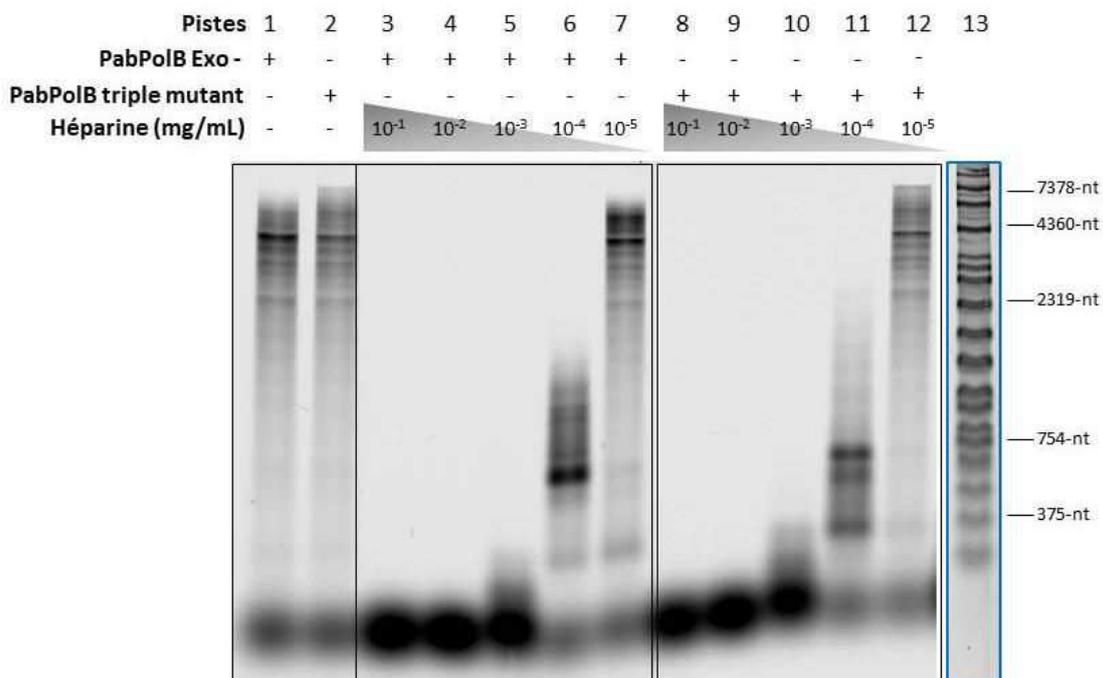


Figure 23. Cinétique d'extension d'amorce hybridée à une matrice de 87 nucléotides par PabPolB *exo-*.

Profils d'électrophorèse en gels dénaturants de polyacrylamide 17 % / urée 8 M des produits des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 17 nucléotides hybridée à une matrice de 87 nucléotides comportant une base uracile en position 29 (pistes 1 à 6) ou en position 33 (pistes 7 à 12), de 200  $\mu$ M de dNTPs et de 75 nM de PolB *exo-*. Les mélanges réactionnels ont été incubés à 55°C aux temps indiqués au-dessus des gels. La piste 13 comporte le marqueur de taille d'oligonucléotides.

### II.3. Processivité de PabPoIB triple mutée

La possibilité d'action de l'ion catalytique divalent sur certaines propriétés de PabPoIB a été explorée. A savoir, le profil d'extension d'amorce, la vitesse de réaction de polymérisation et la capacité de liaison à l'ADN sain et à l'ADN endommagé. Nous avons par la suite étudié le rôle de cet ion dans la processivité de l'enzyme. Pour l'ADN polymérase, la processivité est définie comme le nombre de nucléotides incorporés par l'enzyme par unité de temps au cours d'un cycle d'attachement à l'ADN. Pour étudier cette propriété intrinsèque, il est nécessaire d'utiliser une longue matrice d'ADN, ici le phage circulaire M13mp18 de 7 250 bases a été choisi ; et d'utiliser une molécule qui empêche la pol d'entamer un second cycle d'attachement lorsqu'elle se décroche de l'ADN en cours de réplication, ici, l'héparine a été utilisée. Il a d'abord été nécessaire d'estimer la concentration optimale en héparine à introduire pour obtenir une inhibition efficace de la polymérisation lors de la réaction d'extension d'amorce. Pour cela, l'inhibition de l'activité d'extension d'amorce de PabPoIB *exo-* et de PabPoIB triple mutée a été mesurée en fonction d'une gamme de concentration croissante en héparine. Dans les conditions d'expériences pratiquées,  $10^{-3}$  mg/mL (Figure 24, pistes 5 et 10) semblent être la concentration en héparine optimale pour piéger efficacement PabPoIB *exo-* et PabPoIB triple mutée au cours des réactions d'extension d'amorce.

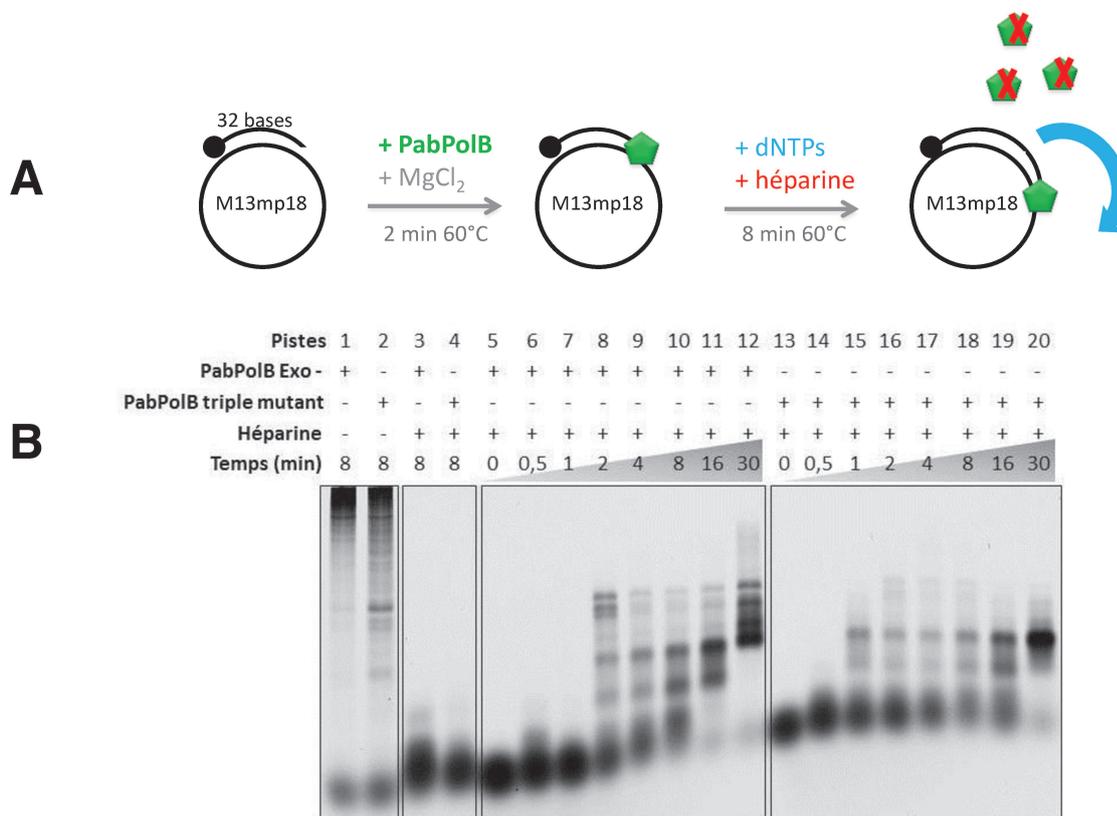


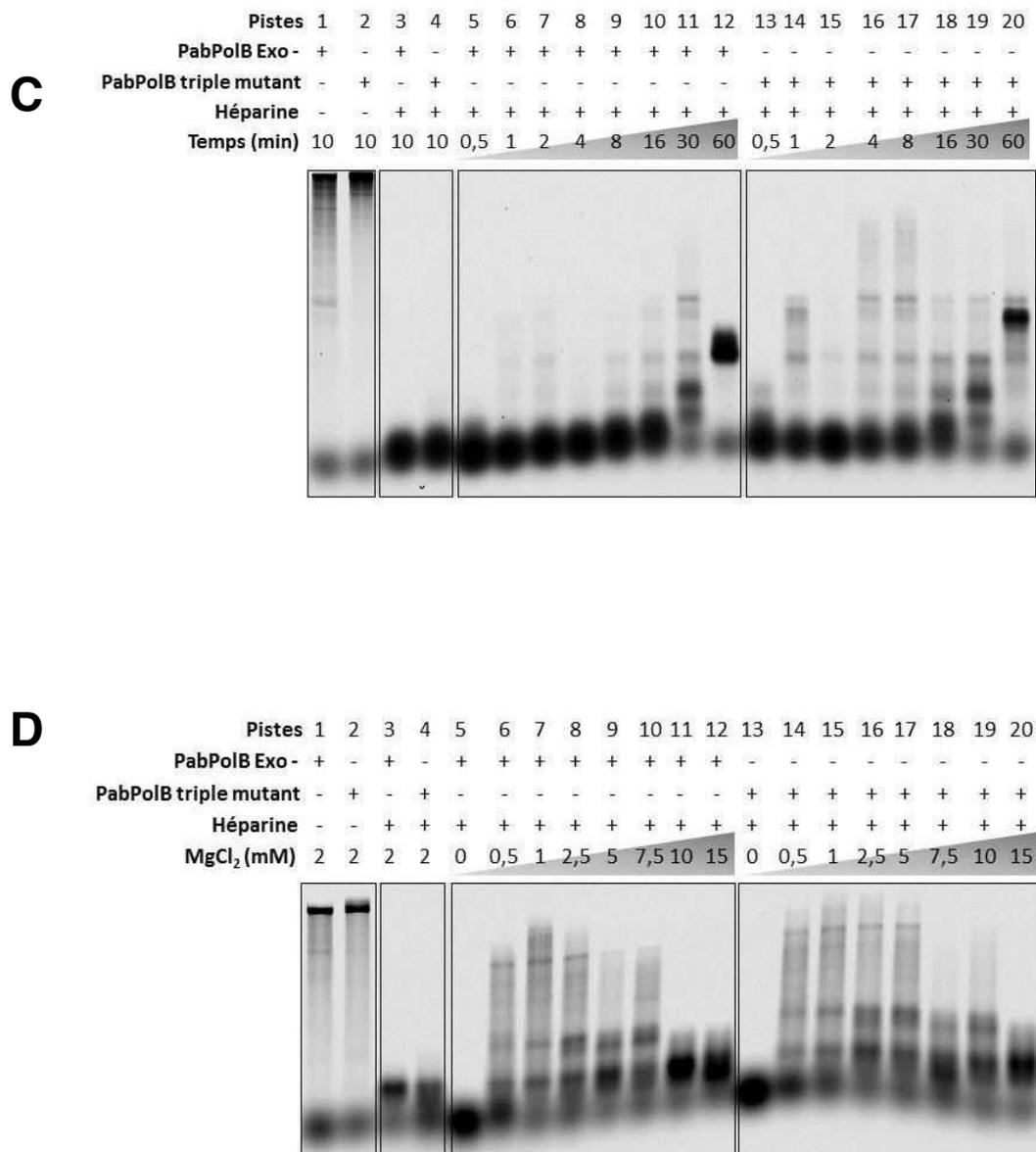
**Figure 24. Réaction d'extension d'amorce hybridée au plasmide M13mp18 de 7250 nucléotides par PabPoIB *exo-* et PabPoIB triple mutée en présence de quantités décroissantes d'héparine.**

Profils électrophorétiques en gel dénaturant d'agarose alcalin à 1 % des produits des réactions enzymatiques effectuées en présence de 7 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 hybridée à la matrice M13mp18(+), de 200  $\mu$ M de dNTPs, de 200 nM de chaque PabPoIB indiquées et de quantités décroissantes d'héparine indiquées au dessus du gel. Les mélanges réactionnels ont été incubés 8 min à 60°C.

La seule différence observée en cristallographie entre PabPolB *exo-* et PabPolB triple mutée *exo-* étant la capacité à fixer un ion divalent dans la structure T-cleft nouvellement identifiée au niveau du domaine N-terminal, la processivité de PabPolB *exo-* et de PabPolB triple mutant *exo-* a été mesurée. La processivité de PabPolB *exo-* et de PabPolB triple mutant *exo-* en présence de  $10^{-3}$  mg/mL d'héparine a été mesurée en fonction du temps d'incubation (Figure 25). La progression de PabPolB *exo-* le long de la matrice semble être plus élevée que celle de PabPolB triple mutant (Figure 25). Alors que PabPolB triple mutée interrompt sa progression le long de la matrice ADN à environ 750 nt après 30 min d'incubation (Figure 25B, piste 20), PabPolB *exo-* poursuit l'extension bien au-delà, à environ 2 320 nt (Figure 25B, piste 12). Il semblerait que PabPolB *exo-* soit plus processive, et donc que l'ion facilite la translocation de l'enzyme d'une base à l'autre le long de la matrice ADN. Cependant, l'expérience semble ne pas être reproductible. En effet, lors d'expériences suivantes en présence d'une quantité moindre en héparine de  $10^{-2}$  mg/mL, le résultat apparaît inversé (Figure 25C). Ces expériences ont donné des résultats incertains et remettraient en cause le choix du temps d'incubation et l'effet du piège héparine. Afin de pouvoir conclure quant à la différence de processivité des deux PabPolBs, une nouvelle mise au point des conditions expérimentales est requise.

Par la suite, l'influence de la concentration en  $MgCl_2$  sur la processivité de PabPolB *exo-* et PabPolB triple mutée *exo-* a été mesurée (Figure D). La progression des deux PabPolBs le long de la matrice M13mp18 durant une incubation de 8 min et en présence de  $10^{-3}$  mg/mL d'héparine a été évaluée. Les deux PabPolBs semblent présenter le même profil de processivité avec un optimum d'activité proche se situant entre 1 et 2,5 mM (Figure 25D, pistes 7 et 8 et pistes 15 et 16).



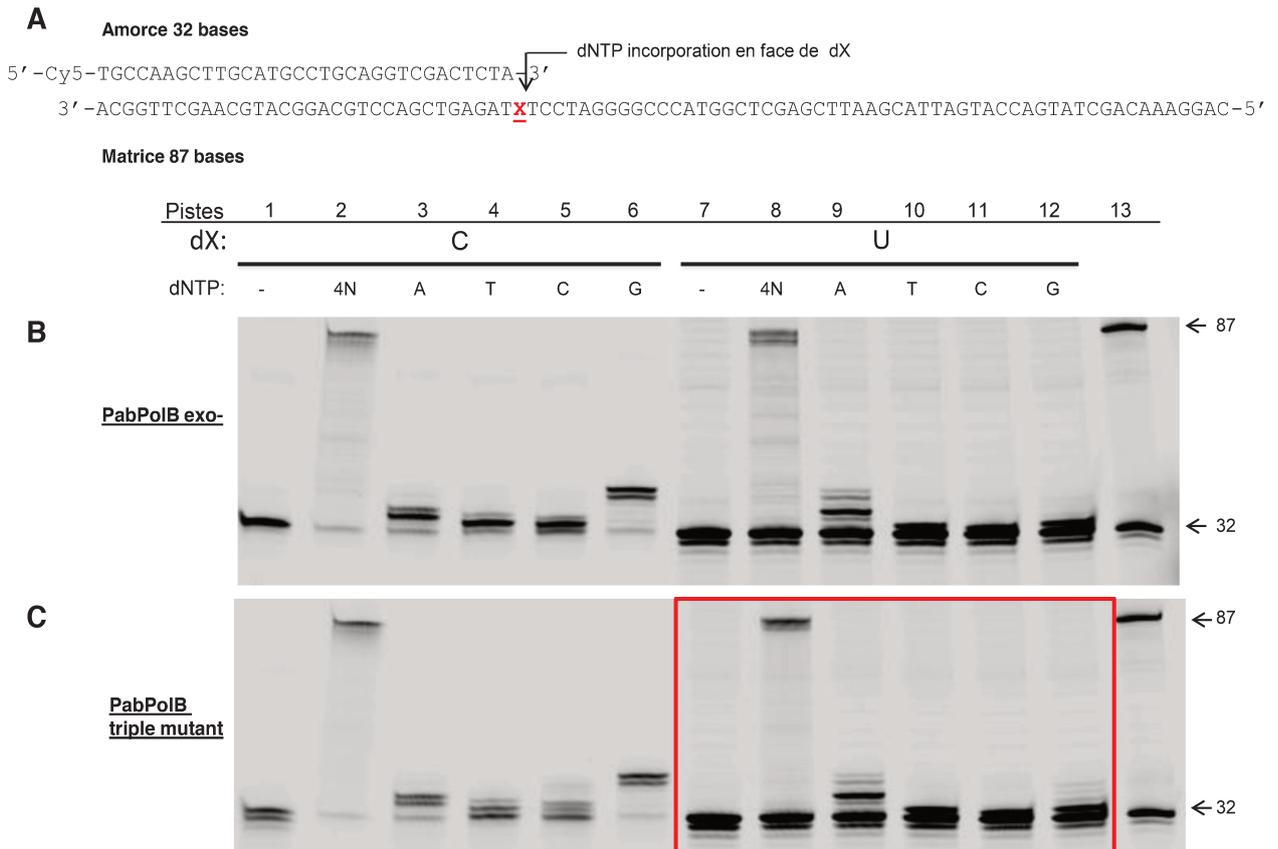


**Figure 25. Réactions de processivité de PabPoIB exo- et PabPoIB triple mutée.**

Profils d'électrophorèse en gel dénaturant d'agarose alcalin à 1 % des produits des réactions enzymatiques effectuées en présence de 7nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 32 nucléotides hybridée à la matrice circulaire de 7 250 nucléotides M13mp18(+), 200 nM de chaque PabPoIB indiquées. Les réactions ont été initiées par l'ajout simultané des dNTPs (200 µM) et de l'héparine (**A**). Les différents mélanges réactionnels ont été incubés à 60°C à différents temps d'incubation indiqués au dessus des gels. (**B**) Réaction de processivité de 0 à 30 min en présence de 10<sup>-3</sup> mg/mL d'héparine. (**C**) Réaction de processivité en fonction du temps de réaction de 0 à 60 min en présence de 10<sup>-2</sup> mg/mL d'héparine. (**D**) Réaction de processivité en fonction d'une gamme de concentration croissante d'ion cofacteur Mg<sup>2+</sup> (indiquées au-dessus du gel) de 8 min en présence de 10<sup>-3</sup> mg/mL d'héparine. Les pistes 1 B,C et D et 2 B,C et D montre les résultats de la réaction en absence de trap et correspond à l'extension maximale de l'amorce par PabPol B exo – (pistes 1) et PabPoIB triple mutée (pistes 2). Les pistes 3 B,C et D et 4 B,C et D montrent les résultats d'extension en présence d'héparine et reflètent l'efficacité du piège pour PabPol B exo – (piste 3) et PabPoIB triple mutée (piste 4).

#### II.4. Etude préliminaire de la sélectivité de PabPolB triple mutée

Cette étude a été effectuée par la méthode d'incorporation préférentielle d'un seul nucléotide face à une base canonique ou une base désaminée (vue dans l'article en préparation présenté dans le Chapitre II de ce manuscrit, pages 71 à 90). Cette réaction consiste à apprécier lequel des quatre dNTPs va être préférentiellement insérés par la Pol B face à une base instructrice donnée par la matrice ADN. Dans l'expérience ci-après, une amorce marquée de 32 nt a été hybridée à une matrice de 87 nt de séquence particulière de sorte que la première base instructrice à être lue par la pol soit une cytosine (C) ou une uracile (U) (Figure 26A). Un des quatre nucléotides libres est mis en présence de la pol dans chaque tube. Le témoin en présence des 4 dNTPs « 4N », montre la capacité de l'enzyme à incorporer l'ensemble des dNTPs et de produire un brin complémentaire de même longueur que la matrice. Ce résultat permet de valider les conditions expérimentales (Figure 26B et C, pistes 2 et 8). Face à la base C, PabPol B triple mutée insère préférentiellement une base G (Figure 26C, piste 6), créant un appariement légitime « Watson-Crick », tout comme PabPolB exo- (Figure 26B, piste 6). On notera que les bases de A, T et C sont incorporées faiblement, ceci est dû au fait que les deux pols étudiées sont dépourvues de la fonction exonucléasique de correction des erreurs d'insertion. Lorsqu'une base endommagée uracile présente dans la matrice ADN est lue par PabPolB triple mutée, celle-ci insère préférentiellement une base adénine en face (Figure 26C, piste 9) et dans ce cas, la base « légitime » G n'est pas préférentiellement incorporée. Le même résultat est observé pour PabPolB exo-. Ainsi, PabPolB exo- et PabPolB triple mutée ont le même profil d'incorporation préférentiel. C'est-à-dire qu'elles incorporent une base légitime face à un C témoignant de leur sélectivité élevée, et une base A face à un dommage uracile. Ainsi, l'ion métallique (absent chez PabPolB triple mutée) ne semble pas intervenir dans la sélectivité élevée de PabPolB. Ces résultats ne sont que préliminaires, une conclusion plus rigoureuse nécessiterait la détermination des paramètres d'incorporation cinétique  $K_m$  et  $k_{cat}$  pour chaque substrat nucléotidique, comme effectué dans l'étude montrée dans l'article en cours de préparation en chapitre II.



**Figure 26. Réaction d'incorporation préférentielle d'un seul nucléotide face à une base cytosine ou une base uracile par PabPoIB exo- et triple mutée.**

Profil d'électrophorèse en gel dénaturant de polyacrylamide 17 % / urée 8M des réactions enzymatiques effectuées en présence de 25 nM d'amorce marquée en 5' par l'indocarbocyanin® Cy5 de 32 nucléotides hybridée à une matrice de 87 nt contenant une base cytosine ou une base uracile en position +1 par rapport à la jonction amorce/matrice (dX) (A) et de 200 µM de dNTP différents indiqués au-dessus des gels. Les réactions ont été initiées par l'ajout de 75 nM de PabPoIB exo- (B) et PabPoIB triple mutée (C) et incubées 30 min à 55°C. Les pistes 1B et C et 7B et C, montrent les résultats de la réaction en absence de substrat nucléotidique. Les pistes 2B et C et 8B et C montrent les résultats en présence des quatre substrats nucléotidiques dATP, dTTP, dCTP et dGTP. Les oligonucleotides marqueurs de références 32 et 87 nucléotides sont indiqués en pistes 13 B et C.

### III. Conclusion – Perspectives

Ces travaux préliminaires sur le rôle fonctionnel de l'ion métallique divalent nouvellement identifié au sein de l'ADN polymérase B archéenne, suggèrent une implication dans l'augmentation de la capacité d'extension d'amorce lorsqu'une base uracile est présente dans l'ADN, par le biais d'une diminution du temps de pause de l'ADN polymérase au dommage au cours de la synthèse d'ADN. En outre, cet ion permettrait d'augmenter la capacité de discrimination de la Pol B entre une base canonique de type cytosine et sa forme désaminée, l'uracile. De façon plus générale, l'ion métallique divalent permettrait à PabPoIB de faciliter son glissement le long de l'ADN de façon indépendante à la présence d'un dommage uracile.

Par la suite, il serait intéressant d'étudier l'affinité de la liaison de PabPoIB en mode polymérase à une matrice d'ADN contenant un dommage uracile. De plus, la propriété fonctionnelle que pourrait conférer cet ion à la processivité de PabPoIB le long d'une matrice d'ADN comportant une base endommagée uracile pourrait être explorée. En outre, il serait intéressant de comprendre le rôle de cet ion nouvellement identifié lorsque PabPoIB est capable de corriger les erreurs d'incorporation en produisant notamment PabPoIB triple mutée exo +. Une étude comparative des propriétés intrinsèques biochimiques entre PabPoIB triple mutée exo – (décrite dans cette étude) et PabPoIB triple mutée exo + (à produire) est donc à mener.



## **CHAPITRE IV**

Essais préliminaires pour la mesure de la concentration intracellulaire « pool » de désoxyribonucléotides et de ribonucléotides chez l'Archaea *Pyrococcus abyssi*.

---

Présentation de l'étude

Matériels et méthodes

Résultats



# Première partie: Présentation de l'étude

---

Les quatre désoxyribonucléotides triphosphate (dNTPs), désoxyadénosine triphosphate, désoxythymidine triphosphate, désoxycytidine triphosphate, désoxyguanosine constituent les briques élémentaires de l'ADN et sont les précurseurs de la synthèse de l'ADN. Leur polymérisation s'effectue au cours de la réplication de l'ADN et est assurée par les acteurs protéiques de la réplication (hélicase, topoisomérase, PCNA, primase, SSB) et l'ADN polymérase.

La réplication fidèle de l'ADN est un des éléments essentiels pour la maintenance de l'intégrité du génome. Elle est à la fois mise en jeu par l'ADN polymérase (qui va sélectionner de façon stéréochimique le dNTP à insérer de façon complémentaire au brin matrice) et par le « pool » de dNTPs lui-même. En effet, la concentration de chaque désoxyribonucléotide triphosphate est finement régulée (Reichard, 1988). Une balance correcte entre les quatre dNTPs et une concentration correcte est requise pour une réplication fidèle. La régulation est effectuée en partie par la ribonucléotide réductase, qui est l'enzyme intervenant dans la biosynthèse des dNTPs. Un déséquilibre dans le « pool » de dNTPs est mutagène. Une augmentation de la concentration d'un des dNTPs entraîne généralement la diminution de la concentration d'un autre, ce qui mène à l'inhibition de la réplication de l'ADN.

Les dNTPs sont également requis lors de la réparation de l'ADN. Des études récentes ont souligné qu'il existe une corrélation entre une induction de dommage à l'ADN et une augmentation de la concentration cellulaire du pool de dNTPs. En effet, chez l'eucaryote *Saccharomyces cerevisiae*, il a été montré que la concentration en dNTPs augmente de 6 à 8 fois en réponse à une agression génotoxique. Cette augmentation est associée à une tolérance aux dommages à l'ADN suggérant que la synthèse translésionnelle d'ADN est plus efficace lorsque la concentration en dNTPs est élevée (Chabes et al., 2003). Au niveau cellulaire, les dommages à l'ADN provoquent des réponses multiples qui incluent l'arrêt du cycle de division et l'activation de la transcription de gènes impliqués dans la réparation de l'ADN (Weinert et al., 1998).

*Pyrococcus abyssi* colonise un milieu relativement hostile (très hautes températures, variations de pH, présence de métaux lourds, etc) et, de ce fait, doit faire face aux attaques génotoxiques. Il a été montré que *Pab* en culture est capable de réparer ses chromosomes fragmentés suite à une irradiation aux rayonnements ionisants gamma et également de prévenir l'accumulation d'erreur génétique par exportation de l'ADN endommagé (Jolivet et al., 2003) ; suggérant ainsi que *Pab* possède un système de réparation de l'ADN hautement efficace.

Par des réactions biochimiques *in vitro*, nous avons observé que la concentration du « pool » total de dNTP influence l'activité enzymatique de l'ADN polymérase B de *Pab*, (Figure 27). En effet, lorsque nous appliquons une augmentation croissante de la concentration totale en dNTPs, nous mettons en évidence un passage progressif de PabPolB du mode exonucléase au mode polymérase. De plus, nous observons que les hautes concentrations testées (500  $\mu$ M) favorisent la synthèse erronée lorsque PabPolB est confrontée au dommage uracile, (Figure 27 pistes 15 et 16). Ce phénomène de translésion a également été montré chez la polymérase répliquative  $\epsilon$  de la levure *S. cerevisiae*, lorsqu'elle est confrontée à une lésion à l'ADN de type oxydative (8-oxoG) (Sabouri et al., 2008).

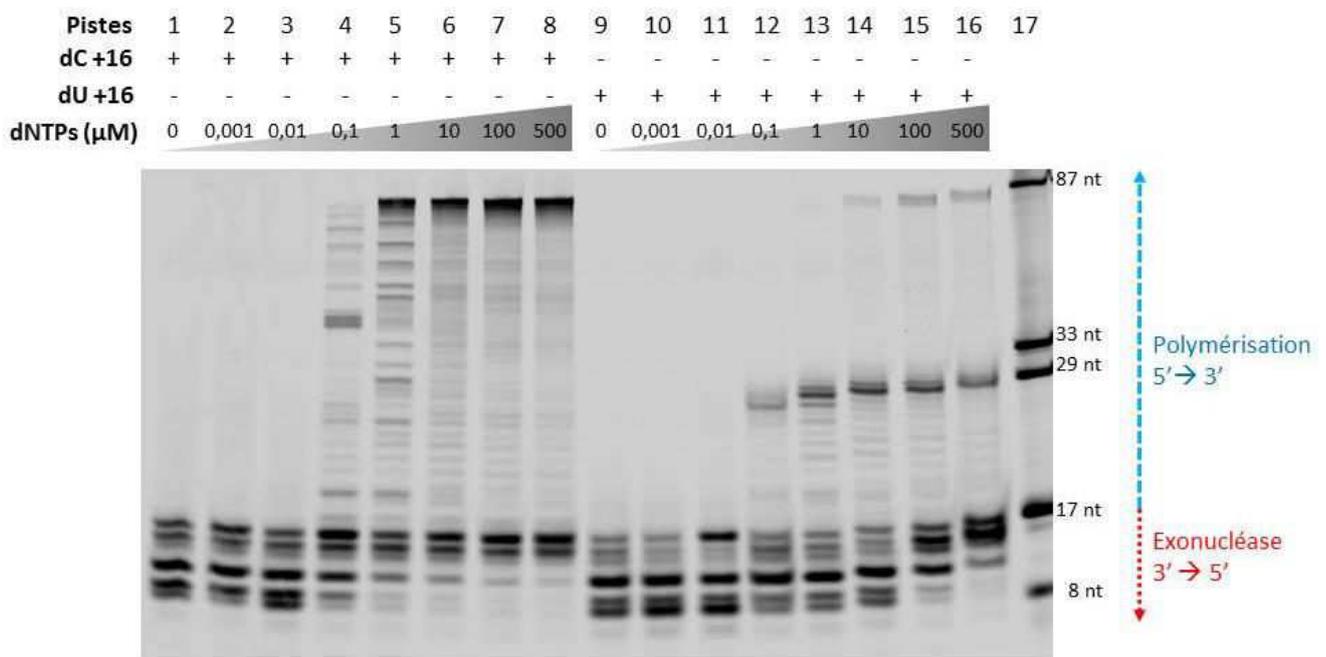


Figure 27. Influence du pool de dNTPs sur la réaction d'extension d'amorce par PabPolB *exo +*.

Gels dénaturants de polyacrylamide 17 % / urée 8 M des réactions d'extension d'amorce de 17 nucléotides marqué en 5' par l'indocarbocyanin® Cy5 hybridée à une matrice de 87 nucléotides comportant en position +16 par rapport à la jonction matrice/amorce une base canonique cytosine (piste 1 à 8) ou une base désaminée uracile (pistes 9 à 16), effectuées en présence de 75 nM de PabPolB *exo +* et d'une gamme de concentrations croissantes de dNTPs, (indiquée au-dessus des gels). Les mélanges réactionnels ont été incubés 30 min à pH 8,8 à 55°C. La piste 17 comporte le marqueur de taille d'oligonucléotides.

Aucune mesure du « pool » de dNTPs dans le domaine Archaea n'a été décrit jusqu'à présent. Etant donné l'importance du rôle joué par les dNTPs au niveau cellulaire et leur implication dans la maintenance génomique, nous avons souhaité mesurer la concentration cellulaire de chaque dNTPs et de chaque NTPs chez *Pab* en conditions normales de croissance et en réponse à un stress à l'ADN induit chimiquement.

# Deuxième partie : Matériels et méthodes

---

## **I. Mesure du pool intracellulaire des dNTPs et NTPs chez *Pyrococcus abyssi***

### **I.1. Culture de *Pyrococcus abyssi*, archée hyperthermophile, en fiole anaérobie**

Des cultures de la souche GE5 de *Pab* (souchothèque de Bretagne) ont été réalisées en mode batch (sans apport de milieu frais), en anaérobiose dans des fioles pénicilline scellées hermétiquement. Deux milieux de culture destinés à la croissance des Thermococcales ont été testés : le milieu SME modifié (Annexe 4 pages 192-193) (Stetter et al., 1983 ; Raven et al., 1992 ; Godfroy et al., 2006) et le milieu minimum « Zobell Marine Broth 2216 » noté « 2216 » initialement élaboré pour la culture de bactéries marines (Zobell, 1941) et par la suite, adapté pour la croissance des Archaea hyperthermophiles sulfato-réductrices (Belkin and Jannasch, 1985 ; Godfroy et al., 1996) dont la composition est donnée en Annexe 3 (page 191). La présence de tampon Pipes permet le maintien du pH du milieu dans une gamme de 6,1 à 7,5 ; le soufre élémentaire (S<sup>0</sup>) joue le rôle d'accepteur final d'électrons et le sulfure de sodium constitue l'agent réducteur permettant de maintenir l'anaérobiose. Une fois stérilisé par tyndallisation, un volume de 800 mL de milieux a été inoculé à raison de 2,5 % par des pré-cultures de *Pab* située en phase exponentielle de croissance avec des seringues stériles. Les cultures ont été incubées à 95 °C.

### **I.2. Suivi de croissance de *Pyrococcus abyssi***

Afin de déterminer la concentration cellulaire en fonction du temps de culture, un dénombrement cellulaire a été effectué toutes les 15-30 min, en triplicat, par observation des cellules placées en hématimètre de Thoma (profondeur 0,02 mm) au microscope photonique. Les cellules de *Pab* en culture ont été prélevées à 5 h de croissance (phase exponentielle) pour l'extraction des nucléotides.

### **I.3. Dosage des dNTPs et NTPs**

La technique de dosage des dNTPs et NTPs utilisée dans cette étude suit un protocole mis au point pour un dosage chez les cellules eucaryotes de levure *Saccharomyces cerevisiae* (Chabes et al., 2003). Ce protocole nécessite une récupération après culture des cellules sur filtre de 0,8 µm. Ces levures ont un diamètre compris entre 4 et 7 µm et sont donc bien plus volumineuses que les cellules de *P. abyssi* de 0,8 à 1 µm. La présence de cristaux de soufre élémentaire est un paramètre à prendre en compte lors de la collecte des cellules. Ainsi, deux techniques de récolte ont été effectuées pour mettre au point le protocole d'extraction des nucléotides à partir des cellules de *P. abyssi*.

### **I.3.1. Collecte des cellules de *Pyrococcus abyssi***

#### **I.3.1.1. Collecte des cellules par filtration**

Toutes les étapes suivantes du protocole ont été effectuées sur glace et au sein d'une enceinte anaérobie. Les cellules de *Pab* en croissance ont été collectées par filtration à une densité de  $2.10^9$  cellules sur des filtres de nitrocellulose (AAWP 0,22  $\mu\text{m}$ , 25 mm, Millipore). Les filtres ont été immergés dans 700  $\mu\text{L}$  de solution d'extraction [12 % acide trichloroacétique (TCA), 15 mM chlorure de magnésium ( $\text{MgCl}_2$ )] en tube Eppendorf, puis congelés dans l'azote liquide et mis en attente à  $-80^\circ\text{C}$ . Le jour de l'extraction, les échantillons ont été agités par retournements brefs et successifs à l'aide d'un mixeur Intelli RM-2 pendant 15 min.

#### **I.3.1.2. Collecte des cellules par centrifugation**

Toutes les étapes suivantes du protocole ont été effectuées au sein d'une enceinte anaérobie. Les centrifugations ont été réalisées dans des pots hermétiques à l'air. Après culture durant un temps donné, les cellules en phase exponentielle de croissance ont été centrifugées 2 min à 200 g dans le but de culotter le soufre élémentaire. Puis la suspension a été centrifugée 60 min à 15 000 g dans le but de culotter un maximum de cellules de *P. abyssi*. Puis, les cellules ont été rincées dans du milieu 2216 + solution C sans soufre et ont été à nouveau centrifugées. Le culot a ensuite été repris dans un volume donné de solution de TCA/ $\text{MgCl}_2$  et l'ensemble du volume a été réparti dans des tubes Eppendorf de façon à avoir approximativement  $1,25.10^{10}$  cellules/tube. Les tubes ont ensuite été congelés dans l'azote liquide et mis en attente à  $-80^\circ\text{C}$ .

### **I.3.2. Extraction des nucléotides totaux par le fréon**

Cette étape a été adaptée du protocole mis au point par Khym (Khym, 1975). Cette étape fait suite aux deux méthodes de récolte des cellules de *Pab* décrites ci-avant. Les cellules recueillies ont été désintégrées par aspiration et refoulement dans les 700  $\mu\text{L}$  d'acide trichloroacétique (0,6 M) /  $\text{MgCl}_2$  (15 mM). Puis, les cellules en suspension ont été récupérées par centrifugation de 1 min à 10000 rpm à  $4^\circ\text{C}$ . Le surnageant a été extrait deux fois avec 1,1 x le volume du mélange 1,1,2-trichloro-1,2,2-trifluoroéthane (fréon) (78 % v / v) / trioctylamine (22 % v / v) puis vortexé et centrifugé 1 min à 10 000 rpm. Un dixième de l'échantillon a été évaporé à sec par un lyophilisateur et dissous dans 100  $\mu\text{L}$  d'eau. Cette fraction a été utilisée pour la quantification de chaque ribonucléotide triphosphate du « pool » NTPs.

### **I.3.3. Purification des désoxyribonucléotides sur colonne d'affinité boronate**

Cette étape du protocole a été adaptée de Shewach et al. (Shewach et al., 1992). Les 90 % restants de l'échantillon (475  $\mu\text{L}$ ) ont été ajustés à pH 8,9 avec 25  $\mu\text{L}$  de carbonate d'ammonium (1 M) et les désoxyribonucléotides ont été séparés des ribonucléotides par chromatographie d'affinité au boronate (Affi-Gel® 601 Boronate Affinity Gel, BioRad). L'éluat (2 mL), contenant des désoxyribonucléotides, a été ajusté à pH 3,4 avec environ 20  $\mu\text{L}$  d'acide chlorhydrique (3 M) et vortexé pendant 30 s pour laisser échapper le  $\text{CO}_2$  en constante formation. Les échantillons ont

ensuite été déposés sur les colonnes de boronate (un échantillon par colonne) et élués par 2,5 mL de tampon "Ambic" [(MgCl<sub>2</sub> (15 mM) + carbonate d'ammonium (50 mM)]. Les éluats sont ensuite évaporés pendant 24 h en lyophilisateur et finalement séparés par HPLC (dans les mêmes conditions que celles décrites pour les ribonucléotides triphosphates).

#### I.3.4. Séparation et quantification des ribonucléotides et des déoxyribonucléotides par HPLC et détection UV

Cette étape a été adaptée du protocole élaboré par Hofer et al. (Hofer et al., 1998). Les échantillons de NTPs et dNTPs ont été respectivement analysés par HPLC par injection sur une colonne SAX-5 Partisphere (4,6 × 125 mm, PolyLC Inc., Columbia, MD, USA) et élués à 1,5 mL / min isocratiquement en tampon acétonitrile 2,5 % (v / v) et phosphate de potassium 0,35 M à pH 3,4. La détection a été effectuée à 260 nm par un détecteur UV-2075 Plus detector (Jasco, Molndal, Sweden). Les nucléotides ont été quantifiés par la mesure de la hauteur des pics déduite de la courbe standard de NTPs et de dNTPs injectés à 100 pmol.

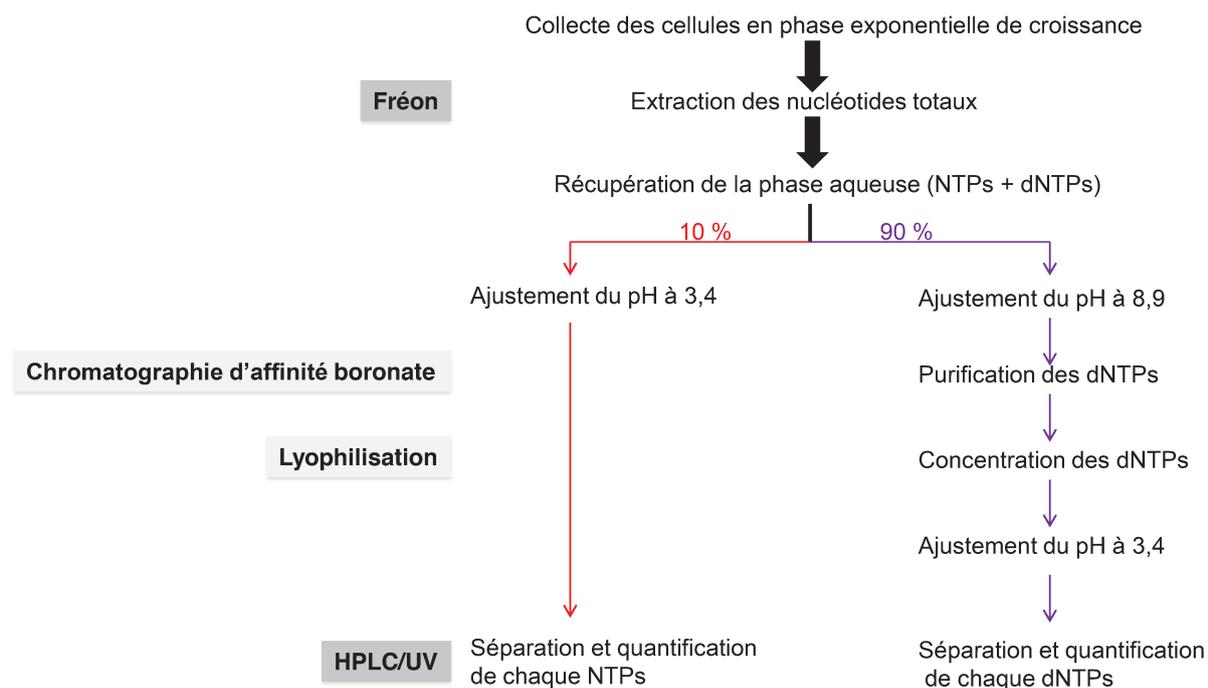


Figure 28. Schéma montrant les différentes étapes du protocole de mesure des concentrations intracellulaires des NTPs d'une part et des dNTPs d'autre part.

## **II. Mesure du pool intracellulaire de dNTPs et NTPs de *Pyrococcus abyssi* ayant subi un stress physico-chimique**

### **II.1. Culture de *Pyrococcus abyssi*, en bioréacteur « gas-lift »**

Des cultures en continu de la souche GE5 de *Pab* ont été effectuées en bioréacteur « gas-lift » de 2 L (Chapitre I, Figure 4D, page 33). C'est un fermenteur qui a été conçu pour cultiver des organismes hyperthermophiles réduisant le soufre (Raven et al., 1992 ; Godfroy et al., 2006) et dont la principale caractéristique est une agitation par un balayage gazeux stérile (azote) en association avec une structure cylindrique de circulation interne des gaz. Cette circulation permet de maintenir l'anaérobiose, d'homogénéiser le milieu de culture et d'évacuer les gaz toxiques (hydrogène et hydrogène sulfuré) produits par le métabolisme de la souche. Le système permet de réguler le pH et la température en temps réel d'après les consignes préalablement enregistrées dans le système analogique de contrôle relié à un ordinateur. Des cultures de *Pab* ont ainsi été réalisées en milieu SME modifié (Annexe 4, pages 192-193) qui est dépourvu de Na<sub>2</sub>S et de S<sup>0</sup> mais qui contient de la L-cystéine comme accepteur final d'électrons, à 90°C, à pH 6,8, en continu, à un taux de dilution de 0,1 h<sup>-1</sup> et avec un balayage à l'azote gazeux à 400 mL/min. Des suivis de croissance ont été effectués, pour cela il a été nécessaire d'entreprendre un lessivage préalable. Cette étape consiste à diminuer d'un Log environ la densité cellulaire dans le fermenteur. Le lessivage s'effectue par l'application d'une augmentation importante du débit d'entrée du milieu de culture (33 mL/min) et par l'augmentation du débit d'azote gazeux afin de pallier à l'arrivée importante d'oxygène via le milieu frais. Ainsi, la densité cellulaire diminue progressivement et, un suivi de croissance, après basculement du fermenteur en mode batch, a pu être effectué à partir d'une concentration à 1.10<sup>7</sup> cellules/mL comme décrit dans dans la partie II.2., ci-après.

### **II.2. Application de stress en cours de croissance**

Un "stress acide" de l'organisme *Pab* été effectué en milieu de phase de croissance exponentielle, après passage du mode continu au mode « batch » par un ajout contrôlé d'acide chlorhydrique au milieu de culture. Le pH du milieu a été abaissé précisément de deux unités pH (de 6,8 à 4,8). Le stress a été appliqué durant 120 minutes. Un comptage cellulaire et des prélèvements pour évaluer la survie des cellules ont été effectués toutes les 30 minutes.

### **II.3. Mesure de la viabilité cellulaire de *Pab* suite au stress pH**

Un dénombrement des cellules viables a été effectué par la technique statistique du « nombre le plus probable » à trois tubes (NPP). Au cours de l'application du stress acide, 1 mL (x 3) de culture a été prélevé en anaérobiose toutes les trente minutes pendant 120 minutes et réparti dans trois tubes scellés en verre. Dix dilutions successives (au 1/10<sup>ème</sup>) ont ensuite été réalisées dans différents tubes. La concentration en cellules viables de l'échantillon a été estimée statistiquement après 48 h de culture dans les mêmes conditions que décrites en I.1 à l'aide de la table de Mac Grady.

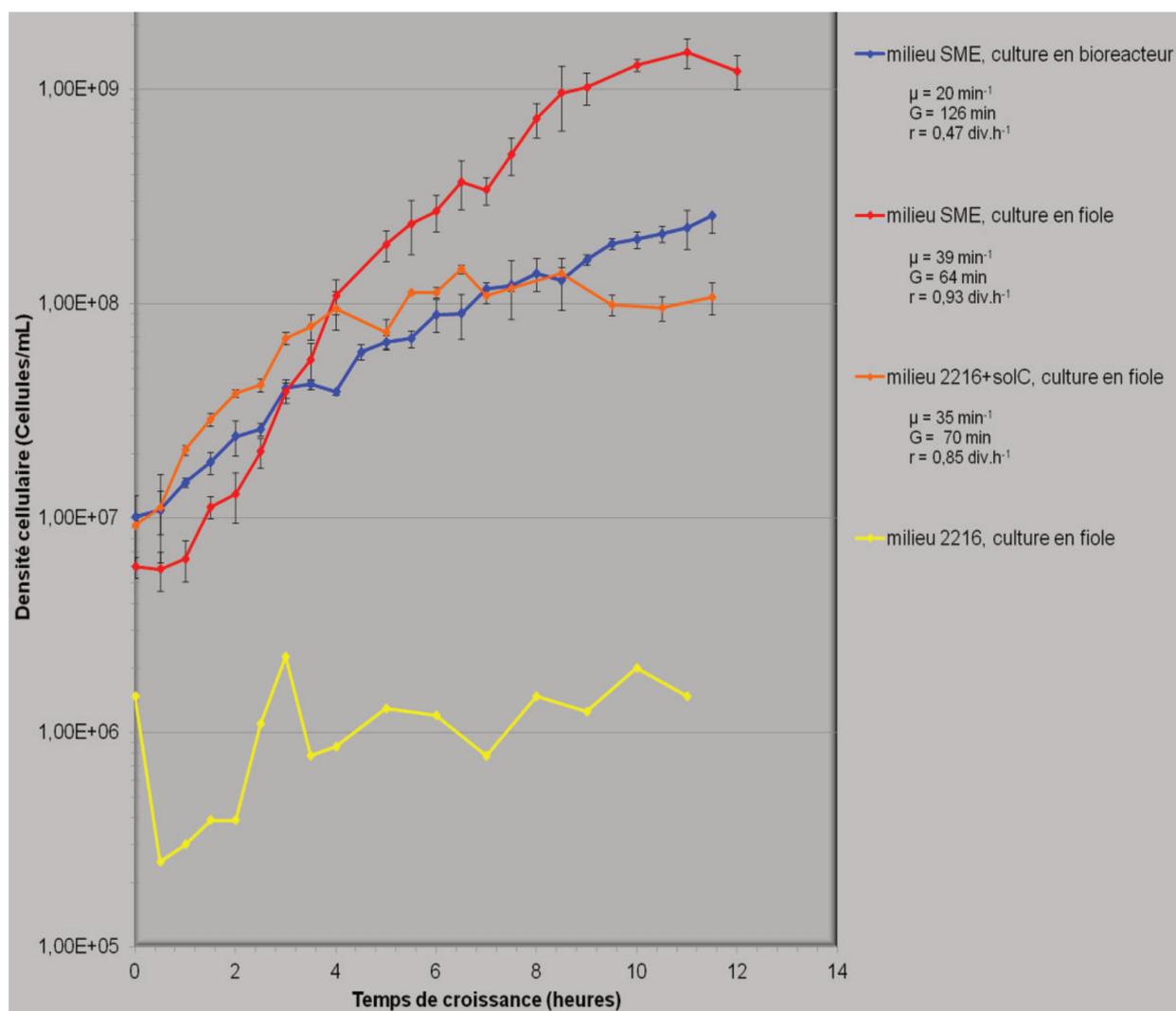
# Troisième partie : Résultats

---

## **I. Mesure du pool intracellulaire de dNTPs et NTPs chez *Pyrococcus abyssi***

### **I.1. Suivi de croissance de *Pyrococcus abyssi* en fiole anaérobie et en fermenteur**

Les milieux de culture habituellement utilisés pour la culture d'organismes hyperthermophiles anaérobies sont des milieux très riches en minéraux, et peuvent parfois biaiser les analyses chimiques qui suivent la culture. C'est pourquoi il a été nécessaire d'optimiser un milieu de culture ayant une composition minimale en minéraux mais assurant une densité cellulaire suffisante ( $> 1.10^8$  cellules/mL). Suite à différents essais expérimentaux de suivi de croissance dans différents milieux de culture en fiole pénicilline de 20 mL ou en bioréacteur (Figure 29), nous avons opté pour un milieu de culture minimum (milieu 2216) auquel nous avons ajouté les éléments chimiques reconnus indispensables à la multiplication de *Pab* à savoir, des sels de bore, de tungstène, de molybdène et de sélénium (milieu 2216 + solution C, annexe 3 page 191). Et nous avons décidé d'effectuer les cultures en fiole et non pas en bioréacteur.



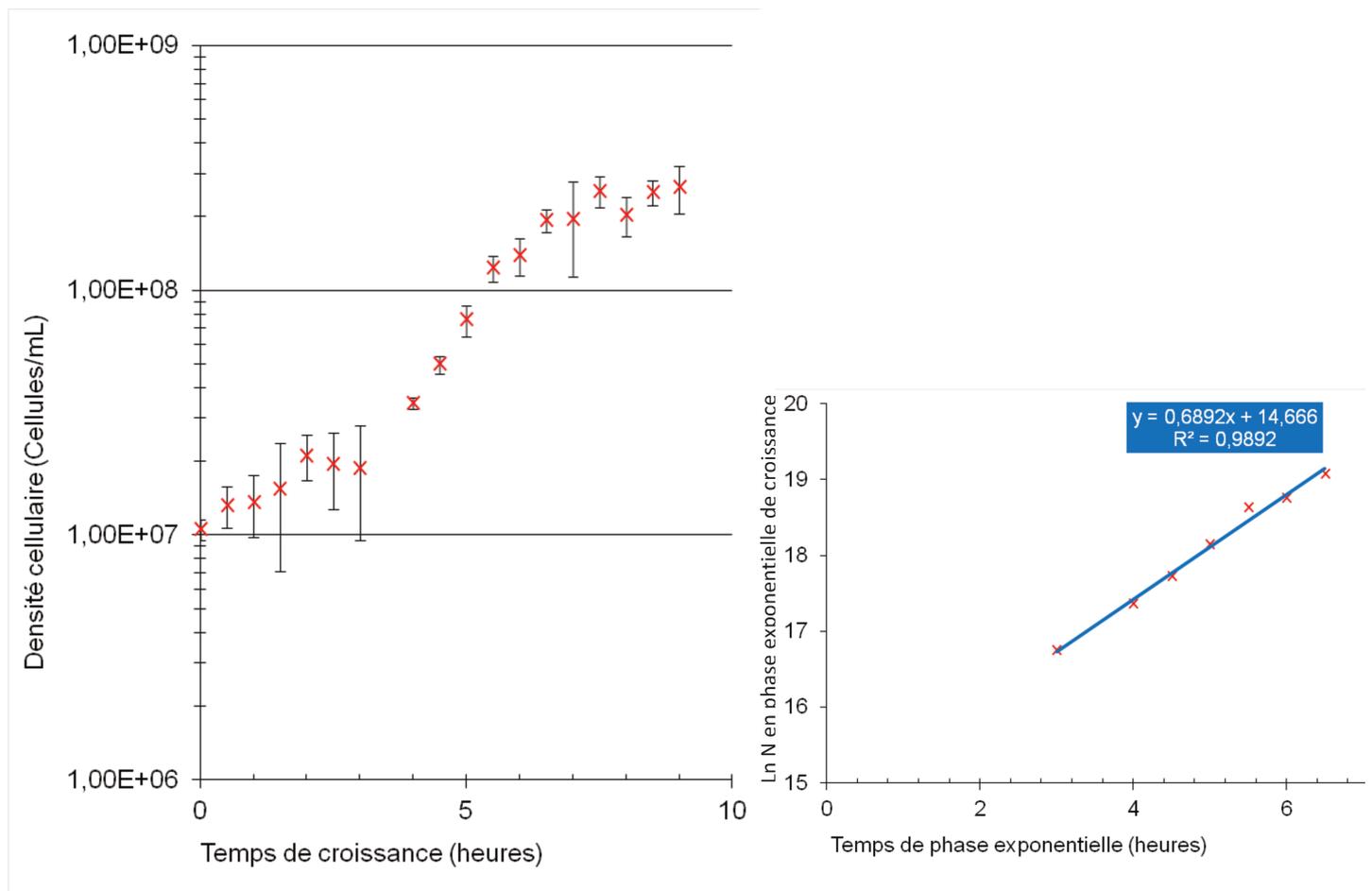
**Figure 29. Suivi de la croissance de *P. abyssi* en mode « batch » dans trois différents milieux de culture en fioles pénicilline de 20 mL ou en bioréacteur « gas-lift » de 2 L.**

Un dénombrement cellulaire a été effectué toutes les 30 min, en triplicat, par prélèvement puis observation des cellules placées en hématimètre de Thoma (profondeur 0,02 mm) au microscope photonique. Les paramètres de la croissance ( $\mu$ , G et r) ont été déterminés à partir de la linéarisation de chaque phase exponentielle et indiqués à droite du graphique.

Par la suite, afin d'obtenir un maximum de biomasse, nous avons travaillé dans des fioles pénicilline de grande capacité (1 L). Afin de connaître la répartition des phases de croissance dans le temps, nous avons effectué un suivi de croissance. D'après la Figure 30A, on peut remarquer que les cellules de *Pab* suivent une croissance classique qui se décompose en trois phases successives en fonction du temps et donc de l'appauvrissement du milieu de culture (croissance en mode "batch"). Tout d'abord, apparaît la phase de latence, qui s'étend sur environ 3 h, puis vient la phase exponentielle de croissance qui s'étend sur environ 5 h, à partir de laquelle les paramètres de la croissance ont été définis. Ainsi, d'après l'équation de la courbe Figure 30B,  $\ln(\text{nombre de cellules}) = ax + t$ , les paramètres de la croissance de *Pab*, sont les suivants :

- Taux spécifique de croissance  $\mu = 41 \text{ min}^{-1}$
- Temps de génération G = 60 min

Enfin, apparaît la phase stationnaire, pendant laquelle il y a autant de cellules en division que de cellules qui meurent ; elle survient après 7 h de culture et reste stable jusqu'à la phase de déclin qui n'est pas présentée sur ce graphique. Afin d'effectuer la mesure de la concentration cellulaire en dNTPs, il est nécessaire de travailler sur des cellules en division. Les cellules ont donc été prélevées en milieu de phase de croissance, c'est-à-dire après 5 h de culture. Afin d'apprécier l'évolution de cette concentration cellulaire en fonction de l'état de division des cellules, des prélèvements ont aussi été effectués en phase stationnaire de croissance, soit après 9 h de culture mais malheureusement la mesure du pool n'a pas été concluante.

**A****B**

**Figure 30. Suivi de croissance de *Pab* en mode « batch » en fiole de 800 mL dans le milieu 2216 + solution C.**

Des cultures de la souche GE5 de *Pab* ont été réalisées dans un volume de 800 mL de milieu de culture 2216+solution C, en mode batch (sans apport de milieu frais), en fiole en verre scellée pour assurer une anaérobiose stricte, à la température de 95 °C et à pH 6,8. La concentration cellulaire, notée N, a été mesurée toutes les 15-30 min en triplicat, pendant 9 h de culture, par observation des cellules grâce à un hématimètre de Thoma (profondeur 0,02 mm) au microscope photonique. **(A)**, La courbe de croissance est exprimée en log de la concentration cellulaire en fonction du temps de culture,  $\text{Log}(N)=f(t)$ . **(B)**, La partie de courbe exponentielle de **(A)** a été linéarisée en  $\text{Ln}(N)=f(t)$ , le coefficient directeur  $R^2$  d'une valeur de 0,9892 témoigne de la linéarité de la courbe. Les paramètres de la croissance ont pu être déterminés.

## I.2. Mesure du pool de dNTPs et NTPs chez *Pyrococcus abyssi*

Ces mesures ont été effectuées au sein de l'University of Umeå (Suède), Department of Medical Biochemistry and Biophysics, dans l'équipe du Pr. Andrei Chabes, avec l'aide technique de Lisette Marjavaara. Cette mission a été financée par un abondement Européen Mer.

### I.2.1. Adaptation du protocole initial

Sachant que les dNTPs sont faiblement concentrés par rapport au NTPs dans les cellules (Chapitre I, Tableau 3 page 45, cf. ratios NTPs/dNTPs de 12 à 180 environ tous types cellulaires confondus), il est nécessaire, après extraction chimique, d'effectuer une étape de concentration des échantillons de dNTPs. Plusieurs degrés de concentration ont été effectués par lyophilisation puis par rassemblement de plusieurs échantillons. Malgré la méthode de concentration, la [dCTP] mesurée chez *Pab* restait très faible (résultats non montrés). Cela pouvait s'expliquer par un problème méthodologique d'extraction et/ou de détection du dCTP, ou à une valeur physiologique faible chez *Pab*. La base pyrimidique cytosine a une absorbance maximale plus proche de 280 nm que de 260 nm contrairement aux bases puriques (G et A). L'expérience de détection en HPLC-UV a donc été réalisée à 260 nm puis reproduite à 280 nm. Cependant, aucune hausse de la quantité du dCTP n'a été observée. Nous avons donc décidé de modifier le protocole original et d'abandonner la collecte des cellules par la méthode de filtration. Ainsi, l'obtention de la biomasse s'est effectuée par centrifugation dont le but est de récolter un maximum de cellules.

### I.2.2. Résultats obtenus après mises au point expérimentales

La mesure des concentrations cellulaires en dCTP, dTTP, dATP et dGTP a été effectuée sur des cellules de *Pab* en cours de phase exponentielle de croissance. Chez *Pab*, parmi les quatre dNTPs, le dTTP semble être celui synthétisé en plus grande quantité (Figure 31A) comme déjà montré chez les cellules eucaryotes (Chapitre I, tableau 3). D'un point de vue quantitatif, l'abondance respective du pool de dNTPs chez *Pab* en phase exponentielle se décline de la façon suivante : dTTP > dATP > dCTP > dGTP (Figure 31 A). Ces valeurs vont dans le même sens que celle décrites chez les eucaryotes, à savoir dTTP > dATP > dCTP > dGTP pour *S. cerevisiae* et dTTP > dCTP > dATP > dGTP chez l'Homme et, de façon intéressante, sont plus éloignées de celles mesurées chez le procaryote *E. coli* dCTP > dTTP > dATP > dGTP (Tableau 3, page 45). Ainsi, la balance du « pool » de nucléotides de *Pab* se rapprocherait plus de celui de l'eucaryote que de la bactérie.

De la même façon, les valeurs du « pool » de NTPs ont été obtenues chez *Pab*. Les valeurs sont supérieures aux valeurs des dNTPs (Figure 31B), comme déjà décrit chez plusieurs organismes (Kornberg et al., 1992 ; Traut, 1994 ; Nick McElhinny et al., 2010a) entre 5 à 20 fois supérieures pour le C et le A respectivement pour *Pab*. Parmi les quatre NTPs, l'ATP semble être celui synthétisé en plus grande quantité comme déjà montré chez les cellules eucaryotes (Tableau 3, page 45). Cela s'explique certainement en raison de son caractère ubiquitaire dans de nombreuses voies du métabolisme cellulaire. D'un point de vue quantitatif, l'abondance du « pool » de NTPs chez *Pab* en phase exponentielle se décline de la façon suivante : ATP > UTP > GTP > CTP (Figure 31B).

Ces différents niveau d'expression sont également retrouvés chez les eucaryotes et chez les procaryotes (Tableau 3, page 45) à savoir ATP > UTP > GTP > CTP.

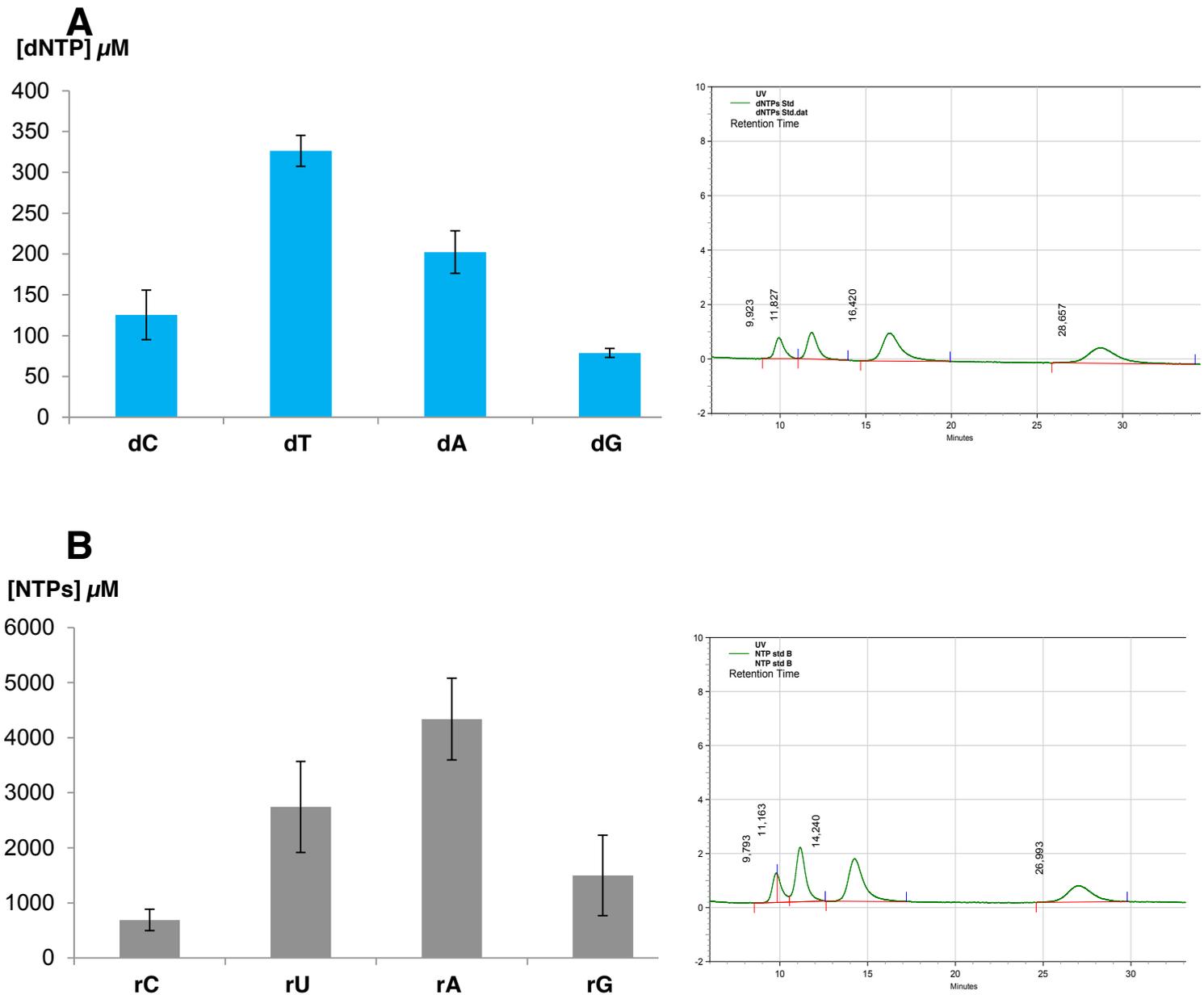
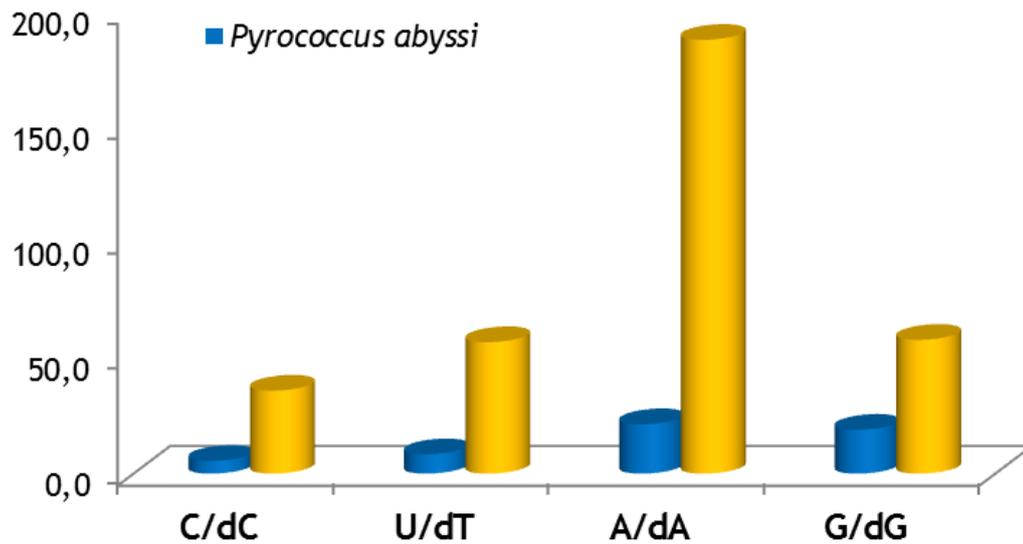


Figure 31. Concentrations des dNTPs et NTPs mesurées chez *Pyrococcus abyssi*.

La concentration des nucléotides a été calculée sur la base de la valeur estimée de la fraction soluble de la cellule de *Pab* à  $0,52 \mu\text{m}^3$ . Les valeurs représentent une moyenne pondérée de la mesure sur cinq échantillons de cellules. Les valeurs de chaque dNTPs (A) et NTPs (B) sont exprimés en  $\mu\text{M}$  d'après un « pool » standard, mesuré par HPLC-UV après extraction au fréon et chromatographie d'affinité. Les chromatogrammes des standards de dNTPs et de NTPs sont respectivement représentés en regard de chaque histogramme. L'élution fait apparaître les nucléotides au cours du temps de l'analyse dans le sens suivant : d/rCTP > dT/rUTP > d/rTTP > d/rGTP.

Les valeurs des ratios dNTPs/NTPs ont été calculées chez *Pab* et apparaissent relativement faible comparativement aux valeurs décrites chez l'eucaryote *S. cerevisiae* (Figure 32, histogrammes bleus).



<i>Pyrococcus abyssi</i>	dNTP	Concentration (μM)	rNTP	Concentration (μM)	Ratios rNTP/dNTP
	dC	125	rC	692	5,5
	dT	326	rU	2744	8,4
	dA	202	rA	4337	21,4
	dG	79	rG	1500	19,0

<i>Saccharomyces cerevisiae</i> (publié par Nick McElhinny et al. 2010a)	dNTP	Concentration (μM)	rNTP	Concentration (μM)	Ratios rNTP/dNTP
	dC	14	rC	500	35,7
	dT	30	rU	1700	56,7
	dA	16	rA	3000	187,5
	dG	12	rG	700	58,3

Figure 32. Calcul du ratio rNTPs/dNTPs pour *Pab* (histogrammes bleus) et comparaison avec ceux de l'eucaryote *S. cerevisiae* (histogrammes jaunes) (Nick McElhinny et al., 2010a)

## II. Mesure du pool intracellulaire de dNTPs et NTPs chez *Pyrococcus abyssi* ayant subi un stress physico-chimique

Initialement, le projet prévoyait d'étudier la variation du « pool » de dNTPs et NTPs intracellulaires suite à l'application de deux types de stress physico-chimiques au cours de la croissance de *Pyrococcus abyssi*. Un stress de type conventionnel, c'est-à-dire un événement environnemental fluctuant auquel *Pab* est soumis dans son écosystème naturel comme les variations de pH. Et un stress non conventionnel auquel *Pab* n'est sensée jamais subir comme le cas de la confrontation aux UV. L'objectif était de comprendre la réponse du « pool » de dNTPs aux stress physico-chimiques environnementaux qui reflètent les mécanismes d'accommodation ou d'adaptation de *Pab*, à l'image de ceux identifiés chez la levure *S. cerevisiae* (Chabes et al., 2003). Ainsi, ces deux stress (UV et pH) ont été appliqués *in vivo* au cours de la croissance de *Pab* de façon contrôlée. Malheureusement, par manque de temps du fait d'une longue mise au point du protocole d'application de stress aux cellules de *Pab*, le projet n'a pas pu être terminé. Cependant, l'effet d'une modification du pH du milieu de culture sur la viabilité cellulaire a pu être estimé. Egalement, une mesure de l'effet de l'exposition d'un UV mimétique sur la variation du pool de NTPs a pu être calculée. Ces résultats, toutefois préliminaires, seront exposés dans la partie qui suit.

### II.1. Culture de *Pyrococcus abyssi*, en bioréacteur « gas-lift »

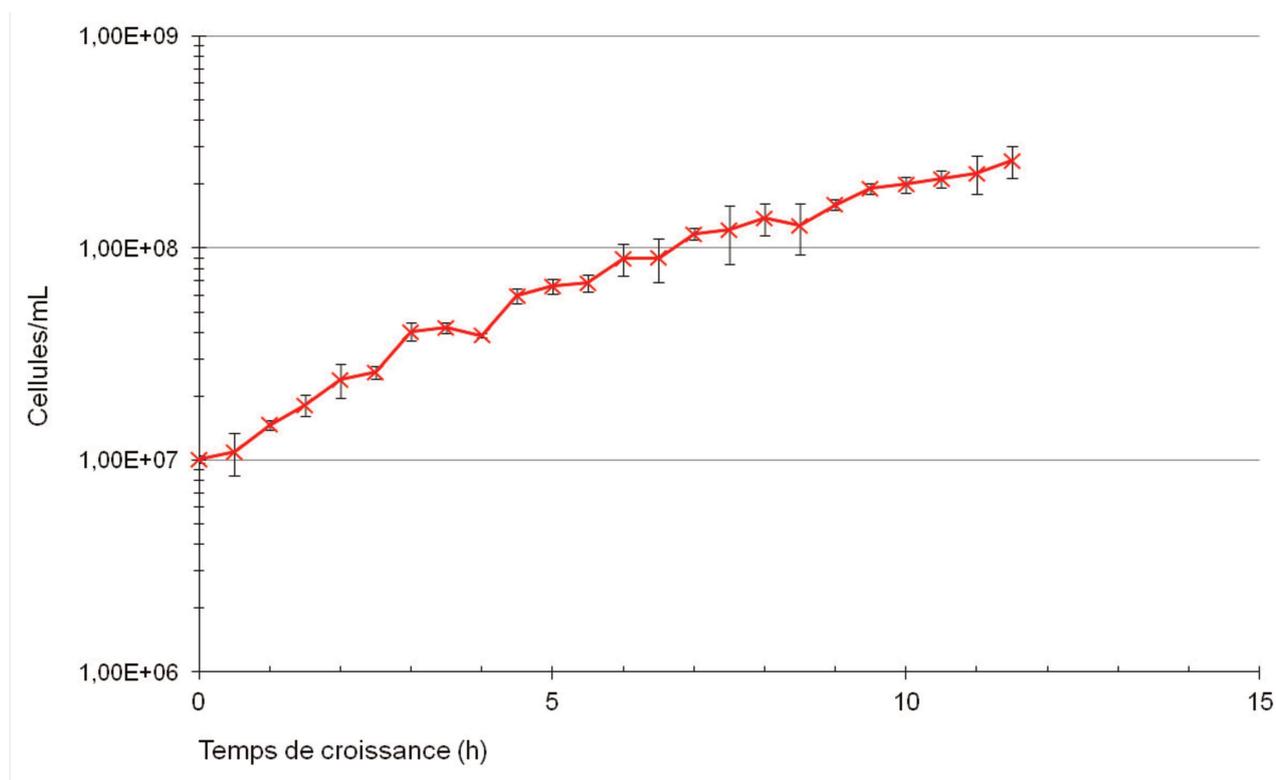
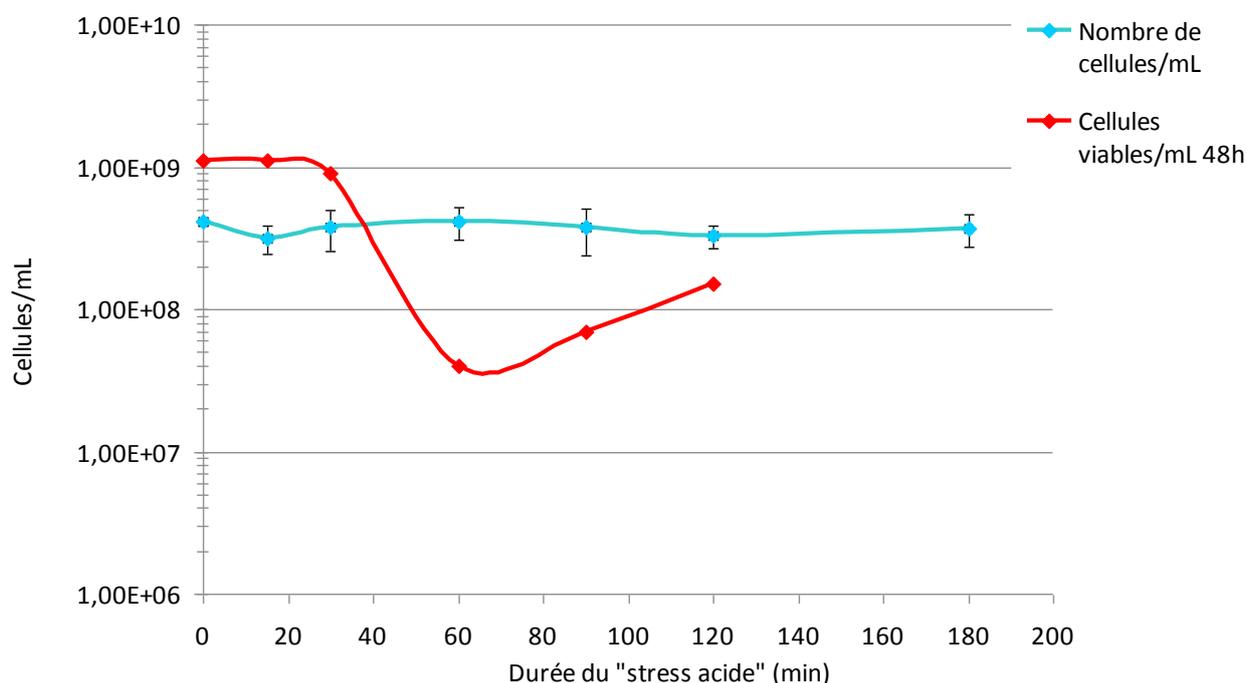


Figure 33. Courbe de suivi de croissance de *Pab* en mode « batch » en bioréacteur « gas-lift ».

Le temps de génération est de 126 min.

## II.2. Influence d'un stress acide sur la viabilité cellulaire de *Pyrococcus abyssi*

Dans le but d'étudier la réponse physiologique de *Pab* à un stress de pH acide, les cellules en cours de phase exponentielle de croissance ont subi un abaissement de 2 unités pH du milieu de culture. Après abaissement du pH de 6,8 à 4,8, grâce au système analogique de contrôle du frementeur, des prélèvements de suspension cellulaire ont été effectués toutes les 20 min durant 2 h pour le dénombrement des cellules viables, c'est-à-dire ayant la capacité de se diviser, et durant 3h pour le comptage de cellules. D'après le dénombrement de cellules viables sous stress, une réponse de mort cellulaire est observée environ 35 min après application du stress (Figure 34, courbe rouge). Ainsi, un abaissement de 2 unités pH du milieu de culture de *Pab* serait toxique. Le nombre de cellules viables décroît progressivement d'un Log et demi pendant 1 h après application du stress. Ainsi, en 40 min un déclin cellulaire s'est opéré puisque seulement 3 % des cellules de *Pab* sont encore viables (Figure 34, courbe rouge). Puis, une lente reprise de la croissance est observée (Figure 34, à  $t = 60$  min, courbe rouge). Il semblerait que *Pab* soit capable de mettre en place des systèmes de protection et/ou de réparation génomique en réponse à un stress acide, dans les conditions testées. Le phénomène de résurgence donne une augmentation du nombre de cellules viables d'environ un Log en 60 min. Ce phénomène est probablement dû au fait qu'une partie de la population de *Pab* a été insensible au stress acide, et a conservé sa capacité à se diviser.



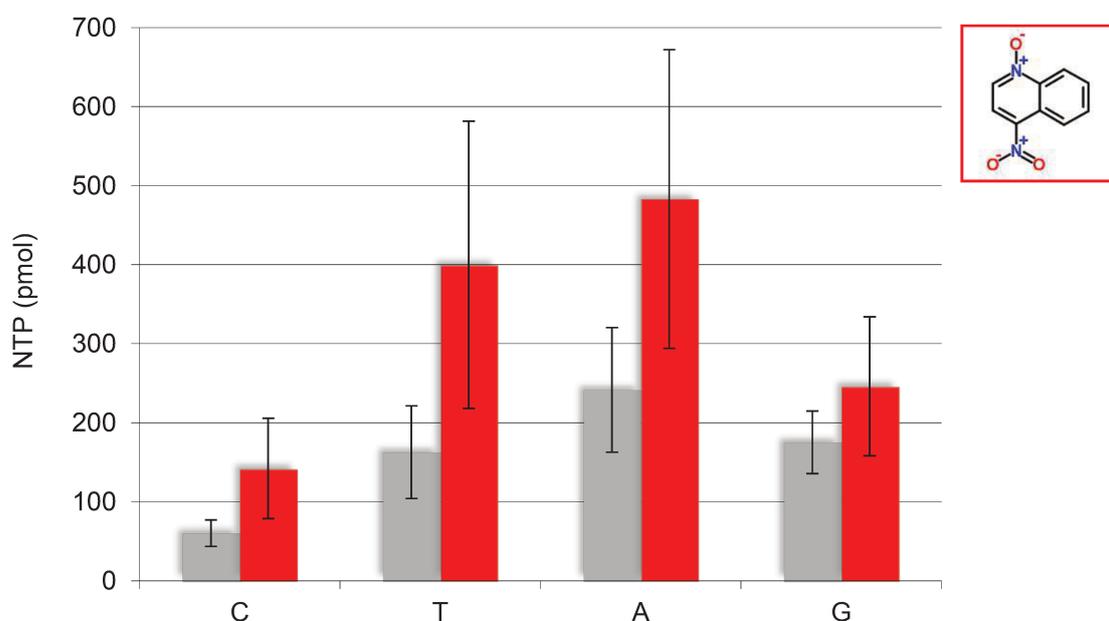
**Figure 34. Suivi de la viabilité cellulaire de *Pab* pendant un stress acide.**

Courbe rouge, à chaque prélèvement de suspension de *Pab* en culture, un dénombrement des cellules viables a été effectué par la technique statistique du « nombre le plus probable » à trois tubes (NPP). La concentration en cellules viables de l'échantillon a été estimée statistiquement à l'aide de la table de Mac Grady. Courbe bleue, à chaque prélèvement de suspension de *Pab* en culture, un comptage cellulaire direct a été effectué par observation au microscope et comptage en hématimètre de Thoma.

Par la suite, il serait intéressant de mesurer la concentration en dNTPs à chaque point de mesure de la viabilité afin de voir si un lien existe entre l'application d'un stress acide et la modulation du pool de dNTPs.

### II.3. Dosage préliminaire du pool de dNTPs et de NTPs après un stress UV

Seuls les valeurs des NTPs seront montrées dans cette partie raison du manque de reproductibilité obtenu pour les valeurs de dNTPs. La variation de la valeur du pool de NTP chez *Pab* sous l'influence de l'agent génotoxique N-Oxyde de 4-NitroQuinoléine (4-NQO), agent alkylant de l'ADN, (Figure 35) pendant sa phase de croissance exponentielle a été mesurée. De façon unanime, les valeurs de chaque NTPs sont augmentées (Figure 35, histogrammes rouges). Une augmentation de 2,5 fois est observée pour le CTP et le TTP, de 2 fois pour l'ATP et de 1,5 fois pour le GTP. Cependant, au vue de la taille des écart-types, les propos doivent être modérés. Cela suggère qu'une modification métabolique s'opère sous l'influence de l'agent alkylant.



**Figure 35. Quantité de chaque NTPs dans les cellules de *Pab* après exposition à l'agent génotoxique 4-NQO.**

Les cellules de *Pab* en cours de croissance exponentielle ont été exposées à 0,2 mg/mL d'agent alkylant 4-NQO pendant 5 h. Les obtenues sont préliminaires et ont été comparées aux valeurs des NTPs de cellules témoins non exposées à la drogue (histogrammes gris). La structure chimique du composé 4-NQO est schématisée à droite du graphique.

### **III. Conclusion-Perspectives**

L'ensemble de ces résultats est encore préliminaire mais encourageant. Il est prévu d'utiliser une autre méthode de séparation/détection des dNTPs par une technique de séparation en chromatographie couplée à une double détection en spectrométrie de masse (HPLC-MS/MS) au CEA de Grenoble (Laboratoire des lésions des acides nucléiques, Dr Jean-Luc Ravanat). En raison de ses conditions de vie à haute température, des désaminations de la base cytosine en uracile dans l'ADN de l'Archaea hyperthermophile *Pab* peuvent s'opérer. C'est pourquoi nous souhaitons mesurer la concentration intracellulaire de dUTP dans les cellules de *Pab* ; également, nous désirons estimer sa fréquence d'apparition au sein du génome de *Pab*. D'autre part, la valeur du pool de dNTPs restant à confirmer, il serait intéressant d'utiliser cette concentration dans nos expériences d'extension d'amorce d'ADN *in vitro*, afin de se rapprocher des conditions de répllication *in vivo* et ainsi de simuler au mieux les conditions physiologiques de *Pyrococcus abyssi*. Enfin, le projet prévoit d'effectuer la mesure du pool de nucléotides chez une Archaea mésophile, *Haloferax volcanii*, afin de comprendre si le facteur température a une part de responsabilité dans les fortes valeurs de dNTPs et de NTPs mesurés chez *Pyrococcus abyssi* au cours de cette thèse.

## CONCLUSION GENERALE ET PERSPECTIVES

---

Les Archaea sont un domaine du Vivant dont les mécanismes biologiques, à la frontière entre eucaryote et procaryote, commencent peu à peu à être compris. La science de la phylogénie et l'émergence de nouvelles techniques de biologie moléculaire ont, en autres, permis d'initier l'étude de ces organismes fascinants. Cependant, un nombre encore restreint d'étude est entrepris sur ces organismes. Les Archaea hyperthermophiles suscitent de nombreuses questions quant à la résistance des molécules biologiques (protéines et acides nucléiques en particulier) à la dénaturation par la chaleur et au maintien de l'information génétique dans de telles conditions physico-chimiques.

La question centrale posée au cours de ce projet de thèse était « quelles sont les particularités des métabolismes et des mécanismes de la réplication de l'ADN spécifiques aux Archaea hyperthermophiles qui concourent à assurer le bon développement de l'organisme dans son environnement de haute température ? »

Le modèle d'étude utilisé était l'Euryarchaea hyperthermophile *Pyrococcus abyssi*. Nous nous sommes intéressés particulièrement à l'enzyme qui a un rôle central dans la réplication de l'ADN : l'ADN polymérase B. *Pab* possède deux pôles répliquatives déjà bien caractérisées biochimiquement. Les hautes températures participent à la création de lésions sur la molécule d'ADN. La présence d'un domaine spécifique de reconnaissance des bases désaminées de l'ADN au sein de la famille B des ADN pol archéennes, largement décrit dans la littérature, témoigne de la capacité de ces organismes à faire face aux dommages. Les expériences biochimiques menées au cours de ce projet ont montré l'importance des ions catalytiques dans les activités enzymatiques intrinsèques de l'enzyme. La propriété qu'ils donnent à l'ADN pol pourrait avoir un rôle dans la relation de la pol avec les dommages de l'ADN. Afin de préserver un taux de mutation à l'ADN relativement bas, toutes cellules du Vivant possèdent des mécanismes de réparation très spécifiques. En outre, le maintien du pool de nucléotides, qui sont les briques de l'ADN, participe aussi à l'établissement de cet équilibre. Les expériences de quantification de la concentration intracellulaire des dNTPs et NTPs au cours de ce sujet de thèse laisseraient penser que *Pab* présente un taux constitutif de nucléotides relativement élevé et aurait un rôle dans la maintenance génomique.

Trois projets principaux ont donc été menés en parallèle au cours de cette thèse. Des collaborations étroites avec d'autres équipes de recherche ont permis de compléter l'approche biochimique effectuée au laboratoire avec notamment l'étude de PabPolB en cristallographie en collaboration avec Jérôme Gouge et Marc Delarue de l'Institut Pasteur à Paris. Cette collaboration a abouti à la compréhension très fine de la relation structure/fonction de PabPolB avec l'ADN sain ou endommagé. L'approche écophysiologique reposant sur la mesure du « pool » de nucléotides (dNTPs et NTPs) chez *Pab* a fait appel à des compétences microbiologiques de culture des hyperthermophiles (Bioréacteur « gas-lift »). La maintenance au quotidien des cultures en bioréacteur et/ou en flacon a fait appel à la contribution d'Anne Godfroy, directrice du laboratoire. D'autre part, la collecte des

cellules, l'extraction et de la quantification des nucléotides effectuée à partir d'un protocole bien décrit chez la levure, ont été optimisées chez *Pab* en partenariat avec Andrei Chabes et Lisette Marjavaara de l'université d'Umeå en Suède.

Les cofacteurs métalliques divalents, logés dans le cœur catalytique des pols, sont indispensables à la réaction catalytique de synthèse d'ADN. Les ions utilisés par les pols sont le magnésium et le manganèse pour l'activité exonucléase, et principalement le magnésium pour l'activité polymérase. De façon surprenante, nous avons mis en évidence que PabPolB est capable d'utiliser le calcium pour mener à bien la synthèse d'ADN ; cette synthèse présente des caractéristiques bien différentes de celle induite par le magnésium comme, par exemple, une vitesse de synthèse très lente. Cette propriété pourrait s'expliquer par une très grande affinité pour la matrice et également par un changement conformationnel inhabituel de la pol. Des études très précises de caractérisation de l'affinité de PabPolB à une matrice ADN amorcée en présence de calcium sont actuellement en cours de réalisation par la technique de résonance plasmonique de surface (SPR). Les mesures de processivité de la pol le long de la matrice sont en train de s'achever. L'ensemble de ces études pourrait conforter un mécanisme de translocation unique de PabPolB en présence de calcium.

D'autre part, des résultats préliminaires montrent que le calcium confère un comportement différent de PabPolB vis-à-vis des nucléotides endommagés. Notamment, la synthèse erronée serait défavorisée et la sélectivité des nucléotides augmentée. Ainsi, la modulation par le calcium de ces propriétés intrinsèques compenserait l'inhibition de correction des erreurs par PabPolB. Un fait particulièrement intéressant est celui du comportement de PabPolD, autre pol répliquative des Euryarchaea. Celle-ci est capable d'utiliser le calcium comme cofacteur de l'activité exonucléase, mais ne peut l'utiliser pour mener à bien la synthèse d'ADN. Le comportement des deux pols de *Pab* est donc opposé vis à vis du cofacteur calcium. Il serait très captivant de poursuivre cette étude du comportement de PabPolB face aux dommages à l'ADN en présence de  $Ca^{2+}$  et d'explorer son rôle potentiel dans la réparation de l'ADN. L'étude d'un panel d'ADN polymérase de famille et d'origine diverses laisse penser que cette capacité à utiliser l'ion calcium dans la synthèse d'ADN serait exclusive aux pols de la famille B.

Pour la première fois, la structure de PabPolB a été résolue par diffraction aux rayons X sur le cristal (Gouge et al., 2012). La structure à très haute résolution de PabPolB en conformation édition en présence d'une construction amorce/matrice d'ADN saine a été obtenue. Egalement, la structure de la forme apo a été déterminée mais à plus faible résolution. L'organisation en domaines (« pouce », « doigt », « paume » et exonucléase) de la protéine est similaire à celle des ADN polymérase de la famille B répliquatives déjà résolues chez différents organismes. La poche de reconnaissance de l'uracile formée par trois acides aminés décrite chez les Pols B d'archées apparaît rigide chez PabPolB. Des structures avec les bases désaminées (uracile et hypoxantine) et canoniques (A, T, C et G) logées dans la poche de reconnaissance ont été obtenues à très haute résolution. Il semblerait que le temps de résidence des bases canoniques soit moins long que celui des bases désaminées. D'autre part, un ion métallique divalent a été découvert et son rôle dans la

réplication de matrice endommagé a été examiné. L'ensemble des résultats obtenus suggère que cet ion augmenterait la discrimination entre les bases canoniques et les bases désaminées par PabPolB, et faciliterait le glissement de la pol en présence de bases désaminées. Ces résultats ne sont que préliminaires et une étude du rôle de l'ion dans l'activité de correction des erreurs de PabPolB serait intéressante. Dans ce cas, la modularité des deux fonctions catalytiques de PabPolB serait à analyser.

Pour la première fois, des valeurs de la concentration intracellulaire du pool de nucléotide libre (dNTPs, NTPs) ont été obtenues chez l'Archaea. Les valeurs des dNTPs mesurés sont relativement plus élevées que celles obtenues chez l'eucaryote *S. cerevisiae*. Mais leur balance est exactement identique avec une prédominance du dTTP. Les valeurs des NTPs mesurés sont également relativement plus importantes que celles obtenues chez *S. cerevisiae* et leur balance est exactement identique avec une prédominance de l'ATP. Ces résultats soulèvent de nombreuses questions et des expériences pourraient être mises en place afin de comprendre comment les valeurs élevées du pool de nucléotides participent aux processus de maintenance du génome de *Pab*. Il serait utile de voir le comportement *in vitro* de PabPolB en cours de réplication de l'ADN endommagé lorsqu'une grande quantité de substrat dNTP est introduite. De plus, il serait intéressant d'explorer l'occurrence des ribonucléotides dans le génome de *Pab* compte tenu de la concentration élevée en NTP mesuré au niveau intracellulaire.



## REFERENCES BIBLIOGRAPHIQUES

---



- Abdulovic, A.L., and Jinks-Robertson, S. (2006). The *in vivo* characterization of translesion synthesis across UV-induced lesions in *Saccharomyces cerevisiae*: insights into Pol zeta- and Pol eta-dependent frameshift mutagenesis. *Genetics* **172**, 1487–1498.
- Ajon, M., Fröls, S., van Wolferen, M., Stoecker, K., Teichmann, D., Driessen, A.J.M., Grogan, D.W., Albers, S.-V., and Schleper, C. (2011). UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. *Mol. Microbiol.* **82**, 807–817.
- Albers, S.-V., and Meyer, B.H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426.
- Albers, S.-V., and Pohlschröder, M. (2009). Diversity of archaeal type IV pilin-like structures. *Extremophiles* **13**, 403–410.
- Allinson, S.L., Dianova, I.I., and Dianov, G.L. (2001). DNA polymerase beta is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts. *EMBO J.* **20**, 6919–6926.
- Andrade, M.A., Ouzounis, C., Sander, C., Tamames, J., and Valencia, A. (1999). Functional classes in the three domains of life. *J. Mol. Evol.* **49**, 551–557.
- Asami, Y., Murakami, M., Shimizu, M., Pisani, F.M., Hayata, I., and Nohmi, T. (2006). Visualization of the interaction between archaeal DNA polymerase and uracil-containing DNA by atomic force microscopy. *Genes Cells* **11**, 3–11.
- Augustin, M.A., Huber, R., and Kaiser, J.T. (2001). Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nat. Struct. Biol.* **8**, 57–61.
- Auzende, J.-M., Urabe, T., Bendel, V., Deplus, C., Eissen, J.-P., Grimaud, D., Huchon, P., Ishibashi, J., Joshima, M., Lagabrielle, Y., et al. (1991). In situ geological and geochemical study of an active hydrothermal site on the North Fiji Basin ridge. *Marine Geology* **98**, 259–269.
- Baliga, N.S., Bonneau, R., Facciotti, M.T., Pan, M., Glusman, G., Deutsch, E.W., Shannon, P., Chiu, Y., Weng, R.S., Gan, R.R., et al. (2004). Genome sequence of *Haloarcula marismortui*: a halophilic archaeon from the Dead Sea. *Genome Res.* **14**, 2221–2234.
- Baños, B., Lázaro, J.M., Villar, L., Salas, M., and de Vega, M. (2008). Characterization of a *Bacillus subtilis* 64-kDa DNA polymerase X potentially involved in DNA repair. *J. Mol. Biol.* **384**, 1019–1028.
- Baños, B., Villar, L., Salas, M., and de Vega, M. (2010). Intrinsic apurinic/aprimidinic (AP) endonuclease activity enables *Bacillus subtilis* DNA polymerase X to recognize, incise, and further repair abasic sites. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19219–19224.
- Barns, S.M., Delwiche, C.F., Palmer, J.D., and Pace, N.R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9188–9193.
- Barry, E.R., and Bell, S.D. (2006). DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.* **70**, 876–887.
- Bartlett, M.S. (2005). Determinants of transcription initiation by archaeal RNA polymerase. *Curr. Opin. Microbiol.* **8**, 677–684.
- Batra, V.K., Shock, D.D., Beard, W.A., McKenna, C.E., and Wilson, S.H. (2012). Binary complex crystal structure of DNA polymerase  $\beta$  reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 113–118.
- Bauer, R.J., Wolff, I.D., Zuo, X., Lin, H.-K., and Trakselis, M.A. (2013). Assembly and Distributive Action of an Archaeal DNA Polymerase Holoenzyme. *J. Mol. Biol.*
- Bavoux, C., Hoffmann, J.S., and Cazaux, C. (2005). Adaptation to DNA damage and stimulation of

- genetic instability: the double-edged sword mammalian DNA polymerase kappa. *Biochimie* 87, 637–646.
- Beard, W.A., Prasad, R., and Wilson, S.H. (2006). Activities and mechanism of DNA polymerase beta. *Meth. Enzymol.* 408, 91–107.
- Beese, L.S., and Steitz, T.A. (1991). Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* 10, 25–33.
- Belkin, S., and Jannasch, H.W. (1985). A new extremely thermophilic, sulfur-reducing heterotrophic, marine bacterium. *Arch. Microbiol.* 141, 181–186.
- Bell, S.D. (2005). Archaeal transcriptional regulation--variation on a bacterial theme? *Trends Microbiol.* 13, 262–265.
- Bell, S.D., and Jackson, S.P. (2001). Mechanism and regulation of transcription in archaea. *Curr. Opin. Microbiol.* 4, 208–213.
- Berquist, B.R., DasSarma, P., and DasSarma, S. (2007). Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1. *BMC Genet.* 8, 31.
- Bertocci, B., De Smet, A., Weill, J.-C., and Reynaud, C.-A. (2006). Nonoverlapping functions of DNA polymerases mu, lambda, and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination *in vivo*. *Immunity* 25, 31–41.
- Bétous, R., Pillaire, M.-J., Pierini, L., van der Laan, S., Recolin, B., Ohi-Séguy, E., Guo, C., Niimi, N., Grúz, P., Nohmi, T., et al. (2013). DNA polymerase kappa-dependent DNA synthesis at stalled replication forks is important for CHK1 activation. *EMBO J.* 32, 2172–2185.
- Bickley, J., Short, J.K., McDowell, D.G., and Parkes, H.C. (1996). Polymerase chain reaction (PCR) detection of *Listeria monocytogenes* in diluted milk and reversal of PCR inhibition caused by calcium ions. *Lett. Appl. Microbiol.* 22, 153–158.
- Biles, B.D., and Connolly, B.A. (2004). Low-fidelity *Pyrococcus furiosus* DNA polymerase mutants useful in error-prone PCR. *Nucleic Acids Res.* 32, e176.
- Blasius, M., Shevelev, I., Jolivet, E., Sommer, S., and Hübscher, U. (2006). DNA polymerase X from *Deinococcus radiodurans* possesses a structure-modulated 3'-->5' exonuclease activity involved in radioresistance. *Mol. Microbiol.* 60, 165–176.
- Blöchl, E., Rachel, R., Burggraf, S., Hafenbradl, D., Jannasch, H.W., and Stetter, K.O. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C. *Extremophiles* 1, 14–21.
- Bolhuis, H., Poole, E.M.T., and Rodriguez-Valera, F. (2004). Isolation and cultivation of Walsby's square archaeon. *Environ. Microbiol.* 6, 1287–1291.
- Bott, T.L., and Brock, T.D. (1969). Bacterial growth rates above 90 degrees C in Yellowstone hot springs. *Science* 164, 1411–1412.
- Boucher, Y. (2007). Lipids : Biosynthesis, Function and Evolution. In *Archaea: molecular and cellular biology*, R. Cavicchioli, (ed.) (Washington, DC, ASM Press), p. 341-353.
- Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001). *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol eta. *Nucleic Acids Res.* 29, 4607–4616.
- Boudsocq, F., Kokoska, R.J., Plosky, B.S., Vaisman, A., Ling, H., Kunkel, T.A., Yang, W., and Woodgate, R. (2004). Investigating the role of the little finger domain of Y-family DNA polymerases in low fidelity synthesis and translesion replication. *J. Biol. Chem.* 279, 32932–32940.

- Bouthier de la Tour, C., Portemer, C., Nadal, M., Stetter, K.O., Forterre, P., and Duguet, M. (1990). Reverse gyrase, a hallmark of the hyperthermophilic archaeobacteria. *J. Bacteriol.* *172*, 6803–6808.
- Bouthier de la Tour, C., Portemer, C., Huber, R., Forterre, P., and Duguet, M. (1991). Reverse gyrase in thermophilic eubacteria. *J. Bacteriol.* *173*, 3921–3923.
- Boyd, E.S., Hamilton, T.L., Wang, J., He, L., and Zhang, C.L. (2013). The role of tetraether lipid composition in the adaptation of thermophilic archaea to acidity. *Front Microbiol* *4*, 62.
- Brasier, M.D., Green, O.R., Jephcoat, A.P., Kleppe, A.K., Van Kranendonk, M.J., Lindsay, J.F., Steele, A., and Grassineau, N.V. (2002). Questioning the evidence for Earth's oldest fossils. *Nature* *416*, 76–81.
- Brautigam, C.A., and Steitz, T.A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* *8*, 54–63.
- Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005). Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* *6*, R42.
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* *6*, 245–252.
- Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* *14*, 274–281.
- Brock, T.D., Brock, K.M., Belly, R.T., and Weiss, R.L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol* *84*, 54–68.
- Brown, J.A., and Suo, Z. (2011). Unlocking the sugar “steric gate” of DNA polymerases. *Biochemistry* *50*, 1135–1142.
- Bunting, K.A., Roe, S.M., and Pearl, L.H. (2003). Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the beta-clamp. *EMBO J.* *22*, 5883–5892.
- Burgers, P.M.J. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. *J. Biol. Chem.* *284*, 4041–4045.
- Cambon-Bonavita, M.A., Schmitt, P., Zieger, M., Flaman, J.M., Lesongeur, F., Raguénès, G., Bindel, D., Frisch, N., Lakkis, Z., Dupret, D., et al. (2000). Cloning, expression, and characterization of DNA polymerase I from the hyperthermophilic archaea *Thermococcus fumicolans*. *Extremophiles* *4*, 215–225.
- Cann, I.K.O. (2008). Cell sorting protein homologs reveal an unusual diversity in archaeal cell division. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 18653–18654.
- Cann, I.K., Komori, K., Toh, H., Kanai, S., and Ishino, Y. (1998). A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* *95*, 14250–14255.
- Cann, I.K., Ishino, S., Hayashi, I., Komori, K., Toh, H., Morikawa, K., and Ishino, Y. (1999a). Functional interactions of a homolog of proliferating cell nuclear antigen with DNA polymerases in Archaea. *J. Bacteriol.* *181*, 6591–6599.
- Cann, I.K., Ishino, S., Nomura, N., Sako, Y., and Ishino, Y. (1999b). Two family B DNA polymerases from *Aeropyrum pernix*, an aerobic hyperthermophilic crenarchaeote. *J. Bacteriol.* *181*, 5984–5992.
- Cann, I.K., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H., and Ishino, Y. (2001). Biochemical analysis of replication factor C from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* *183*, 2614–2623.

- Castrec, B., Rouillon, C., Henneke, G., Flament, D., Querellou, J., and Raffin, J.-P. (2009). Binding to PCNA in Euryarchaeal DNA Replication requires two PIP motifs for DNA polymerase D and one PIP motif for DNA polymerase B. *J. Mol. Biol.* *394*, 209–218.
- Cavicchioli, R. (2007). *Archaea: molecular and cellular biology*. xii + 523 pp.
- Cavicchioli, R., Curmi, P.M.G., Saunders, N., and Thomas, T. (2003). Pathogenic archaea: do they exist? *Bioessays* *25*, 1119–1128.
- Chaban, B., Ng, S.Y.M., and Jarrell, K.F. (2006). Archaeal habitats from the extreme to the ordinary. *Can. J. Microbiol.* *52*, 73–116.
- Chabes, A., and Stillman, B. (2007). Constitutively high dNTP concentration inhibits cell cycle progression and the DNA damage checkpoint in yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* *104*, 1183–1188.
- Chabes, A., Georgieva, B., Domkin, V., Zhao, X., Rothstein, R., and Thelander, L. (2003). Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell* *112*, 391–401.
- Charbonnier, F., and Forterre, P. (1994). Comparison of plasmid DNA topology among mesophilic and thermophilic eubacteria and archaeobacteria. *J. Bacteriol.* *176*, 1251–1259.
- Charbonnier, F., Erauso, G., Barbeyron, T., Prieur, D., and Forterre, P. (1992). Evidence that a plasmid from a hyperthermophilic archaeobacterium is relaxed at physiological temperatures. *J. Bacteriol.* *174*, 6103–6108.
- Chia, N., Cann, I., and Olsen, G.J. (2010). Evolution of DNA replication protein complexes in eukaryotes and Archaea. *PLoS ONE* *5*, e10866.
- Cohen, G.N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., Poch, O., Prieur, D., Quéréllou, J., Ripp, R., et al. (2003). An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol. Microbiol.* *47*, 1495–1512.
- Cozens, C., Pinheiro, V.B., Vaisman, A., Woodgate, R., and Holliger, P. (2012). A short adaptive path from DNA to RNA polymerases. *Proc. Natl. Acad. Sci. U.S.A.* *109*, 8067–8072.
- Cubonová, L., Richardson, T., Burkhart, B.W., Kelman, Z., Connolly, B.A., Reeve, J.N., and Santangelo, T.J. (2013). Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in *Thermococcus kodakarensis*. *J. Bacteriol.* *195*, 2322–2328.
- Darwin, C. (1859). *On the Origin of the Species by Natural Selection* (Murray).
- DeLong, E.F. (2000). Resolving a methane mystery. *Nature* *407*, 577, 579.
- DeLong, E.F., and Pace, N.R. (2001). Environmental diversity of bacteria and archaea. *Syst. Biol.* *50*, 470–478.
- DeLong, E.F., Wu, K.Y., Prézelin, B.B., and Jovine, R.V. (1994). High abundance of Archaea in Antarctic marine picoplankton. *Nature* *371*, 695–697.
- Dietrich, J., Schmitt, P., Zieger, M., Preve, B., Rolland, J.-L., Chaabihi, H., and Gueguen, Y. (2002). PCR performance of the highly thermostable proof-reading B-type DNA polymerase from *Pyrococcus abyssi*. *FEMS Microbiol. Lett.* *217*, 89–94.
- Dirmeier, R., Keller, M., Hafenbradl, D., Braun, F.J., Rachel, R., Burggraf, S., and Stetter, K.O. (1998). *Thermococcus acidaminovorans* sp. nov., a new hyperthermophilic alkaliphilic archaeon growing on amino acids. *Extremophiles* *2*, 109–114.
- Domínguez, O., Ruiz, J.F., Laín de Lera, T., García-Díaz, M., González, M.A., Kirchhoff, T., Martínez-

- A, C., Bernad, A., and Blanco, L. (2000). DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J.* *19*, 1731–1742.
- Doolittle, W.F., and Bapteste, E. (2007). Pattern pluralism and the Tree of Life hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* *104*, 2043–2049.
- Dopson, M., Baker-Austin, C., Hind, A., Bowman, J.P., and Bond, P.L. (2004). Characterization of *Ferroplasma* isolates and *Ferroplasma acidarmanus* sp. nov., extreme acidophiles from acid mine drainage and industrial bioleaching environments. *Appl. Environ. Microbiol.* *70*, 2079–2088.
- Doublié, S., and Ellenberger, T. (1998). The mechanism of action of T7 DNA polymerase. *Curr. Opin. Struct. Biol.* *8*, 704–712.
- Drake, J.W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. U.S.A.* *88*, 7160–7164.
- Dridi, B., Raoult, D., and Drancourt, M. (2011). Archaea as emerging organisms in complex human microbiomes. *Anaerobe* *17*, 56–63.
- Dumstorf, C.A., Clark, A.B., Lin, Q., Kissling, G.E., Yuan, T., Kucherlapati, R., McGregor, W.G., and Kunkel, T.A. (2006). Participation of mouse DNA polymerase iota in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 18083–18088.
- Edgell, D.R., Klenk, H.P., and Doolittle, W.F. (1997). Gene duplications in evolution of archaeal family B DNA polymerases. *J. Bacteriol.* *179*, 2632–2640.
- Elkins, J.G., Podar, M., Graham, D.E., Makarova, K.S., Wolf, Y., Randau, L., Hedlund, B.P., Brochier-Armanet, C., Kunin, V., Anderson, I., et al. (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 8102–8107.
- Emptage, K., O'Neill, R., Solovyova, A., and Connolly, B.A. (2008). Interplay between DNA polymerase and proliferating cell nuclear antigen switches off base excision repair of uracil and hypoxanthine during replication in archaea. *J. Mol. Biol.* *383*, 762–771.
- Erauso, G., Reysenbach, A.-L., Godfroy, A., Meunier, J.-R., Crump, B., Partensky, F., Baross, J.A., Marteinson, V., Barbier, G., Pace, N.R., et al. (1993). *Pyrococcus abyssi* sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* *160*, 338–349.
- Erauso, G., Marsin, S., Benbouzid-Rollet, N., Baucher, M.F., Barbeyron, T., Zivanovic, Y., Prieur, D., and Forterre, P. (1996). Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. *J. Bacteriol.* *178*, 3232–3237.
- Evans, S.J., Fogg, M.J., Mamone, A., Davis, M., Pearl, L.H., and Connolly, B.A. (2000). Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* *28*, 1059–1066.
- De Felice, M., Sensen, C.W., Charlebois, R.L., Rossi, M., and Pisani, F.M. (1999). Two DNA polymerase sliding clamps from the thermophilic archaeon *Sulfolobus solfataricus*. *J. Mol. Biol.* *291*, 47–57.
- De Felice, M., Medagli, B., Esposito, L., De Falco, M., Pucci, B., Rossi, M., Grüz, P., Nohmi, T., and Pisani, F.M. (2007). Biochemical evidence of a physical interaction between *Sulfolobus solfataricus* B-family and Y-family DNA polymerases. *Extremophiles* *11*, 277–282.
- Ferraro, P., Franzolin, E., Pontarin, G., Reichard, P., and Bianchi, V. (2010). Quantitation of cellular deoxynucleoside triphosphates. *Nucleic Acids Res.* *38*, e85.
- Ferry, J.G., and Kestead, K.A. (2007). Methanogenesis. In *Archaea: molecular and cellular biology*, R. Cavicchioli, (ed.) (Washington, DC, ASM Press),

- Fidalgo da Silva, E., and Reha-Krantz, L.J. (2007). DNA polymerase proofreading: active site switching catalyzed by the bacteriophage T4 DNA polymerase. *Nucleic Acids Res.* *35*, 5452–5463.
- Filée, J., Forterre, P., Sen-Lin, T., and Laurent, J. (2002). Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *J. Mol. Evol.* *54*, 763–773.
- Firbank, S.J., Wardle, J., Heslop, P., Lewis, R.J., and Connolly, B.A. (2008). Uracil recognition in archaeal DNA polymerases captured by X-ray crystallography. *J. Mol. Biol.* *381*, 529–539.
- Fogg, M.J., Pearl, L.H., and Connolly, B.A. (2002). Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat. Struct. Biol.* *9*, 922–927.
- Forterre, P., Charbonnier, F., Marguet, E., Harper, F., and Henckes, G. (1992). Chromosome structure and DNA topology in extremely thermophilic archaeobacteria. *Biochem. Soc. Symp.* *58*, 99–112.
- Franklin, M.C., Wang, J., and Steitz, T.A. (2001). Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* *105*, 657–667.
- Franzmann, P.D., Springer, N., Ludwig, W., Conway De Macario, E., and Rohde, M. (1992). A Methanogenic Archaeon from Ace Lake, Antarctica: *Methanococcoides burtonii* sp. nov. *Systematic and Applied Microbiology* *15*, 573–581.
- Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R., and Steitz, T.A. (1988). Cocystal structure of an editing complex of Klenow fragment with DNA. *Proc. Natl. Acad. Sci. U.S.A.* *85*, 8924–8928.
- Freudenthal, B.D., Beard, W.A., Shock, D.D., and Wilson, S.H. (2013). Observing a DNA polymerase choose right from wrong. *Cell* *154*, 157–168.
- Fröls, S. (2013). Archaeal biofilms: widespread and complex. *Biochem. Soc. Trans.* *41*, 393–398.
- Fuchs, R.P., Fujii, S., and Wagner, J. (2004). Properties and functions of *Escherichia coli*: Pol IV and Pol V. *Adv. Protein Chem.* *69*, 229–264.
- Gao, J., and Wang, J. (2012). Re-annotation of two hyperthermophilic archaea *Pyrococcus abyssi* GE5 and *Pyrococcus furiosus* DSM 3638. *Curr. Microbiol.* *64*, 118–129.
- Garcia-Diaz, M., Bebenek, K., Gao, G., Pedersen, L.C., London, R.E., and Kunkel, T.A. (2005). Structure-function studies of DNA polymerase lambda. *DNA Repair (Amst.)* *4*, 1358–1367.
- García-Díaz, M., Domínguez, O., López-Fernández, L.A., de Lera, L.T., Saníger, M.L., Ruiz, J.F., Párraga, M., García-Ortiz, M.J., Kirchhoff, T., del Mazo, J., et al. (2000). DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis. *J. Mol. Biol.* *301*, 851–867.
- Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003). PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, “*Pyrococcus abyssi*.” *J. Bacteriol.* *185*, 3888–3894.
- Geslin, C., Gaillard, M., Flament, D., Rouault, K., Le Romancer, M., Prieur, D., and Erauso, G. (2007). Analysis of the first genome of a hyperthermophilic marine virus-like particle, PAV1, isolated from *Pyrococcus abyssi*. *J. Bacteriol.* *189*, 4510–4519.
- Gill, S., O’Neill, R., Lewis, R.J., and Connolly, B.A. (2007). Interaction of the family-B DNA polymerase from the archaeon *Pyrococcus furiosus* with deaminated bases. *J. Mol. Biol.* *372*, 855–863.
- Godfroy, A., Meunier, J.R., Guezennec, J., Lesongeur, F., Raguénès, G., Rimbault, A., and Barbier, G. (1996). *Thermococcus fumicolans* sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent in the north Fiji Basin. *Int. J. Syst. Bacteriol.* *46*, 1113–1119.

- Godfroy, A., Raven, N.D., and Sharp, R.J. (2000). Physiology and continuous culture of the hyperthermophilic deep-sea vent archaeon *Pyrococcus abyssi* ST549. *FEMS Microbiol. Lett.* *186*, 127–132.
- Göksenin, A.Y., Zahurancik, W., LeCompte, K.G., Taggart, D.J., Suo, Z., and Pursell, Z.F. (2012). Human DNA polymerase  $\epsilon$  is able to efficiently extend from multiple consecutive ribonucleotides. *J. Biol. Chem.* *287*, 42675–42684.
- Gon, S., Napolitano, R., Rocha, W., Coulon, S., and Fuchs, R.P. (2011). Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 19311–19316.
- Goodman, M.F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* *71*, 17–50.
- Gouge, J., Ralec, C., Henneke, G., and Delarue, M. (2012). Molecular recognition of canonical and deaminated bases by *P. abyssi* family B DNA polymerase. *J. Mol. Biol.* *423*, 315–336.
- Grabowski, B., and Kelman, Z. (2003). Archeal DNA replication: eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.* *57*, 487–516.
- Greagg, M.A., Fogg, M.J., Panayotou, G., Evans, S.J., Connolly, B.A., and Pearl, L.H. (1999). A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc. Natl. Acad. Sci. U.S.A.* *96*, 9045–9050.
- Grice, E.A., and Segre, J.A. (2011). The skin microbiome. *Nat. Rev. Microbiol.* *9*, 244–253.
- Grice, E.A., and Segre, J.A. (2012). The human microbiome: our second genome. *Annu Rev Genomics Hum Genet* *13*, 151–170.
- Grúz, P., Shimizu, M., Pisani, F.M., De Felice, M., Kanke, Y., and Nohmi, T. (2003). Processing of DNA lesions by archaeal DNA polymerases from *Sulfolobus solfataricus*. *Nucleic Acids Res.* *31*, 4024–4030.
- Gueguen, Y., Rolland, J.L., Lecompte, O., Azam, P., Le Romancer, G., Flament, D., Raffin, J.P., and Dietrich, J. (2001). Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi*. *Eur. J. Biochem.* *268*, 5961–5969.
- Guo, D., Wu, X., Rajpal, D.K., Taylor, J.S., and Wang, Z. (2001). Translesion synthesis by yeast DNA polymerase zeta from templates containing lesions of ultraviolet radiation and acetylaminofluorene. *Nucleic Acids Res.* *29*, 2875–2883.
- Guy, L., and Ettema, T.J.G. (2011). The archaeal “TACK” superphylum and the origin of eukaryotes. *Trends Microbiol.* *19*, 580–587.
- Håkansson, P., Hofer, A., and Thelander, L. (2006). Regulation of mammalian ribonucleotide reduction and dNTP pools after DNA damage and in resting cells. *J. Biol. Chem.* *281*, 7834–7841.
- Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., Inoue, T., and Kai, Y. (2001). Crystal structure of DNA polymerase from hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. *J. Mol. Biol.* *306*, 469–477.
- Hayashi, I., Morikawa, K., and Ishino, Y. (1999). Specific interaction between DNA polymerase II (PolD) and RadB, a Rad51/Dmc1 homolog, in *Pyrococcus furiosus*. *Nucleic Acids Res.* *27*, 4695–4702.
- Henneke, G., Gueguen, Y., Flament, D., Azam, P., Querellou, J., Dietrich, J., Hübscher, U., and Raffin, J.-P. (2002). Replication factor C from the hyperthermophilic archaeon *Pyrococcus abyssi* does not need ATP hydrolysis for clamp-loading and contains a functionally conserved RFC PCNA-binding domain. *J. Mol. Biol.* *323*, 795–810.

- Henneke, G., Flament, D., Hübscher, U., Querellou, J., and Raffin, J.-P. (2005). The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. *J. Mol. Biol.* *350*, 53–64.
- Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J.N. (1995). 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* *3*, 1295–1306.
- Henzler-Wildman, K., and Kern, D. (2007). Dynamic personalities of proteins. *Nature* *450*, 964–972.
- Hingorani, M.M., and O'Donnell, M. (2000). Sliding clamps: a (tail)ored fit. *Curr. Biol.* *10*, R25–29.
- Hofer, A., Ekanem, J.T., and Thelander, L. (1998). Allosteric regulation of *Trypanosoma brucei* ribonucleotide reductase studied *in vitro* and *in vivo*. *J. Biol. Chem.* *273*, 34098–34104.
- Hofer, A., Crona, M., Logan, D.T., and Sjöberg, B.-M. (2012). DNA building blocks: keeping control of manufacture. *Crit. Rev. Biochem. Mol. Biol.* *47*, 50–63.
- Hogg, M., Wallace, S.S., and Doublé, S. (2004). Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J.* *23*, 1483–1493.
- Hopfner, K.P., Eichinger, A., Engh, R.A., Laue, F., Ankenbauer, W., Huber, R., and Angerer, B. (1999). Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl. Acad. Sci. U.S.A.* *96*, 3600–3605.
- Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C., and Stetter, K.O. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* *417*, 63–67.
- Ishibashi, J., Grimaud, D., Nojiri, Y., Auzende, J.-M., and Urabe, T. (1994). Fluctuation of chemical compositions of the phase-separated hydrothermal fluid from the North Fiji Basin Ridge. *Marine Geology* *116*, 215–226.
- Ishino, Y., and Cann, I.K. (1998). The euryarchaeotes, a subdomain of Archaea, survive on a single DNA polymerase: fact or farce? *Genes Genet. Syst.* *73*, 323–336.
- Ishino, Y., and Ishino, S. (2012). Rapid progress of DNA replication studies in Archaea, the third domain of life. *Sci China Life Sci* *55*, 386–403.
- Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* *202*, 715–728.
- Jaenicke, R. (1996). Stability and folding of ultrastable proteins: eye lens crystallins and enzymes from thermophiles. *FASEB J.* *10*, 84–92.
- Jaenicke, R., and Böhm, G. (2001). Thermostability of proteins from *Thermotoga maritima*. *Meth. Enzymol.* *334*, 438–469.
- Jain, R., Rivera, M.C., Moore, J.E., and Lake, J.A. (2002). Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol* *61*, 489–495.
- Jarrell, K.F., and Albers, S.-V. (2012). The archaellum: an old motility structure with a new name. *Trends Microbiol.* *20*, 307–312.
- Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases  $\iota$  and  $\zeta$  act sequentially to bypass DNA lesions. *Nature* *406*, 1015–1019.
- Jokela, M., Eskelinen, A., Pospiech, H., Rouvinen, J., and Syväoja, J.E. (2004). Characterization of the 3' exonuclease subunit DP1 of *Methanococcus jannaschii* replicative DNA polymerase D. *Nucleic Acids Res.* *32*, 2430–2440.

- Jolivet, E., Matsunaga, F., Ishino, Y., Forterre, P., Prieur, D., and Myllykallio, H. (2003). Physiological responses of the hyperthermophilic archaeon "Pyrococcus abyssi" to DNA damage caused by ionizing radiation. *J. Bacteriol.* *185*, 3958–3961.
- Jolivet, E., Corre, E., L'Haridon, S., Forterre, P., and Prieur, D. (2004). *Thermococcus marinus* sp. nov. and *Thermococcus radiotolerans* sp. nov., two hyperthermophilic archaea from deep-sea hydrothermal vents that resist ionizing radiation. *Extremophiles* *8*, 219–227.
- Jones, M.L. (1981). *Riftia pachyptila* Jones: Observations on the Vestimentiferan Worm from the Galapagos Rift. *Science* *213*, 333–336.
- Jones, D.S., Williams, D.F., and Romanek, C.S. (1986). Life history of symbiont-bearing giant clams from stable isotope profiles. *Science* *231*, 46–48.
- Joyce, C.M. (1997). Choosing the right sugar: how polymerases select a nucleotide substrate. *Proc. Natl. Acad. Sci. U.S.A.* *94*, 1619–1622.
- Joyce, G.F. (2012). Bit by bit: the Darwinian basis of life. *PLoS Biol.* *10*, e1001323.
- Juez, G., Rodriguez-Valera, F., Ventosa, A., and Kushner, D.J. (1986). *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec. nov., Two New Species of Extremely Halophilic Archaeobacteria. *Systematic and Applied Microbiology* *8*, 75–79.
- Kamtekar, S., Berman, A.J., Wang, J., Lázaro, J.M., de Vega, M., Blanco, L., Salas, M., and Steitz, T.A. (2004). Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage phi29. *Mol. Cell* *16*, 609–618.
- Kandler, O., and König, H. (1998). Cell wall polymers in Archaea (Archaeobacteria). *Cell. Mol. Life Sci.* *54*, 305–308.
- Karlin, S., Brocchieri, L., Campbell, A., Cyert, M., and Mrázek, J. (2005). Genomic and proteomic comparisons between bacterial and archaeal genomes and related comparisons with the yeast and fly genomes. *Proc. Natl. Acad. Sci. U.S.A.* *102*, 7309–7314.
- Kashefi, K., and Lovley, D.R. (2003). Extending the upper temperature limit for life. *Science* *301*, 934.
- Kawai, A., Hashimoto, H., Higuchi, S., Tsunoda, M., Sato, M., Nakamura, K.T., and Miyamoto, S. (2011). A novel heterotetrameric structure of the crenarchaeal PCNA2-PCNA3 complex. *J. Struct. Biol.* *174*, 443–450.
- Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., et al. (1999). Complete genome sequence of an aerobic hyperthermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res.* *6*, 83–101, 145–152.
- Kelman, Z. (2000). DNA replication in the third domain (of life). *Curr. Protein Pept. Sci.* *1*, 139–154.
- Kelman, Z., and O'Donnell, M. (1995). DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.* *64*, 171–200.
- Kelman, Z., and White, M.F. (2005). Archaeal DNA replication and repair. *Curr. Opin. Microbiol.* *8*, 669–676.
- Kelman, Z., Pietrokovski, S., and Hurwitz, J. (1999). Isolation and characterization of a split B-type DNA polymerase from the archaeon *Methanobacterium thermoautotrophicum* deltaH. *J. Biol. Chem.* *274*, 28751–28761.
- Khare, V., and Eckert, K.A. (2002). The proofreading 3'→5' exonuclease activity of DNA polymerases: a kinetic barrier to translesion DNA synthesis. *Mutat. Res.* *510*, 45–54.
- Khym, J.X. (1975). An Analytical System for Rapid Separation of Tissue Nucleotides at Low Pressures

on Conventional Anion Exchangers. *Clinical Chemistry* 21, 1245–1252.

Killelea, T., and Connolly, B.A. (2011). Role of disulfide bridges in archaeal family-B DNA polymerases. *Chembiochem* 12, 1330–1336.

Killelea, T., Ghosh, S., Tan, S.S., Heslop, P., Firbank, S.J., Kool, E.T., and Connolly, B.A. (2010). Probing the interaction of archaeal DNA polymerases with deaminated bases using X-ray crystallography and non-hydrogen bonding isosteric base analogues. *Biochemistry* 49, 5772–5781.

Kim, S.W., Kim, D.-U., Kim, J.K., Kang, L.-W., and Cho, H.-S. (2008). Crystal structure of *Pfu*, the high fidelity DNA polymerase from *Pyrococcus furiosus*. *Int. J. Biol. Macromol.* 42, 356–361.

Kirouac, K.N., and Ling, H. (2009). Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase  $\iota$ . *EMBO J.* 28, 1644–1654.

Komori, K., and Ishino, Y. (2000). Functional interdependence of DNA polymerizing and 3'→5' exonucleolytic activities in *Pyrococcus furiosus* DNA polymerase I. *Protein Eng.* 13, 41–47.

Kornberg, A., KORNBERG, S.R., and SIMMS, E.S. (1956). Metaphosphate synthesis by an enzyme from *Escherichia coli*. *Biochim. Biophys. Acta* 20, 215–227.

Krupovic, M., Gonnet, M., Hania, W.B., Forterre, P., and Erauso, G. (2013). Insights into dynamics of mobile genetic elements in hyperthermophilic environments from five new *Thermococcus* plasmids. *PLoS ONE* 8, e49044.

Kulaeva, O.I., Koonin, E.V., McDonald, J.P., Randall, S.K., Rabinovich, N., Connaughton, J.F., Levine, A.S., and Woodgate, R. (1996). Identification of a DinB/UmuC homolog in the archeon *Sulfolobus solfataricus*. *Mutat. Res.* 357, 245–253.

Kulczyk, A.W., Akabayov, B., Lee, S.-J., Bostina, M., Berkowitz, S.A., and Richardson, C.C. (2012). An interaction between DNA polymerase and helicase is essential for the high processivity of the bacteriophage T7 replisome. *J. Biol. Chem.* 287, 39050–39060.

Kumar, S., Tsai, C.J., and Nussinov, R. (2000). Factors enhancing protein thermostability. *Protein Eng.* 13, 179–191.

Kunkel, T.A. (2004). DNA replication fidelity. *J. Biol. Chem.* 279, 16895–16898.

Kunkel, T.A. (2011). Balancing eukaryotic replication asymmetry with replication fidelity. *Curr Opin Chem Biol* 15, 620–626.

Kunkel, T.A., and Bebenek, K. (1988). Recent studies of the fidelity of DNA synthesis. *Biochim. Biophys. Acta* 951, 1–15.

Kuroita, T., Matsumura, H., Yokota, N., Kitabayashi, M., Hashimoto, H., Inoue, T., Imanaka, T., and Kai, Y. (2005). Structural mechanism for coordination of proofreading and polymerase activities in archaeal DNA polymerases. *J. Mol. Biol.* 351, 291–298.

Lamers, M.H., Georgescu, R.E., Lee, S.-G., O'Donnell, M., and Kuriyan, J. (2006). Crystal structure of the catalytic alpha subunit of *E. coli* replicative DNA polymerase III. *Cell* 126, 881–892.

Le Breton, M., Henneke, G., Norais, C., Flament, D., Myllykallio, H., Querellou, J., and Raffin, J.-P. (2007). The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair? *J. Mol. Biol.* 374, 1172–1185.

Lecoite, F., Shevelev, I.V., Bailone, A., Sommer, S., and Hübscher, U. (2004). Involvement of an X family DNA polymerase in double-stranded break repair in the radioresistant organism *Deinococcus radiodurans*. *Mol. Microbiol.* 53, 1721–1730.

Lee, J.W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T.A., Wang, Z., and Povirk, L.F.

- (2004). Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J. Biol. Chem.* *279*, 805–811.
- Li, Y., Kong, Y., Korolev, S., and Waksman, G. (1998). Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci.* *7*, 1116–1123.
- Li, Z., Kelman, L.M., and Kelman, Z. (2013). *Thermococcus kodakarensis* DNA replication. *Biochem. Soc. Trans.* *41*, 332–338.
- Li-bo L. I. N., Yu-fen, L.I.U., Xi-peng, L.I.U., Jian-hua, L.I.U., (2012). Biochemical Characterization of Uracil-DNA Glycosylase from *Pyrococcus furiosus*. *CHEMICAL RESEARCH IN CHINESE UNIVERSITIES* *28*, 477–482.
- Lin, L.-J., Yoshinaga, A., Lin, Y., Guzman, C., Chen, Y.-H., Mei, S., Lagunas, A.M., Koike, S., Iwai, S., Spies, M.A., et al. (2010). Molecular analyses of an unusual translesion DNA polymerase from *Methanosarcina acetivorans* C2A. *J. Mol. Biol.* *397*, 13–30.
- Lindahl, T. (1979). DNA glycosylases, endonucleases for apurinic/aprimidinic sites, and base excision-repair. *Prog. Nucleic Acid Res. Mol. Biol.* *22*, 135–192.
- Lindahl, T., and Karlström, O. (1973). Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution. *Biochemistry* *12*, 5151–5154.
- Lindahl, T., and Nyberg, B. (1974). Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* *13*, 3405–3410.
- Lindås, A.-C., and Bernander, R. (2013). The cell cycle of archaea. *Nat Rev Micro* *11*, 627–638.
- Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001). Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell* *107*, 91–102.
- Lombard, J., López-García, P., and Moreira, D. (2012). The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* *10*, 507–515.
- López-García, P., and Forterre, P. (1997). DNA topology in hyperthermophilic archaea: reference states and their variation with growth phase, growth temperature, and temperature stresses. *Mol. Microbiol.* *23*, 1267–1279.
- Lucas, S., Toffin, L., Zivanovic, Y., Charlier, D., Moussard, H., Forterre, P., Prieur, D., and Erauso, G. (2002). Construction of a shuttle vector for, and spheroplast transformation of, the hyperthermophilic archaeon *Pyrococcus abyssi*. *Appl. Environ. Microbiol.* *68*, 5528–5536.
- Lundberg, K.S., Shoemaker, D.D., Adams, M.W., Short, J.M., Sorge, J.A., and Mathur, E.J. (1991). High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* *108*, 1–6.
- Lundgren, M., Andersson, A., Chen, L., Nilsson, P., and Bernander, R. (2004). Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc. Natl. Acad. Sci. U.S.A.* *101*, 7046–7051.
- Lundin, D., Gribaldo, S., Torrents, E., Sjöberg, B.-M., and Poole, A.M. (2010). Ribonucleotide reduction - horizontal transfer of a required function spans all three domains. *BMC Evol. Biol.* *10*, 383.
- MacNeill, S.A. (2001). Understanding the enzymology of archaeal DNA replication: progress in form and function. *Mol. Microbiol.* *40*, 520–529.
- Madsen, E.L. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. *Curr. Opin. Biotechnol.* *22*, 456–464.

- Makarova, K.S., Yutin, N., Bell, S.D., and Koonin, E.V. (2010). Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* **8**, 731–741.
- Mäkinieniemi, M., Pospiech, H., Kilpeläinen, S., Jokela, M., Vihinen, M., and Syväoja, J.E. (1999). A novel family of DNA-polymerase-associated B subunits. *Trends Biochem. Sci.* **24**, 14–16.
- Marguet, E., and Forterre, P. (1994). DNA stability at temperatures typical for hyperthermophiles. *Nucleic Acids Res.* **22**, 1681–1686.
- Markkanen, E., van Loon, B., Ferrari, E., Parsons, J.L., Dianov, G.L., and Hübscher, U. (2012). Regulation of oxidative DNA damage repair by DNA polymerase  $\lambda$  and MutYH by cross-talk of phosphorylation and ubiquitination. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 437–442.
- Marteinsson, V.T., Moulin, P., Birrien, J., Gambacorta, A., Vernet, M., and Prieur, D. (1997). Physiological Responses to Stress Conditions and Barophilic Behavior of the Hyperthermophilic Vent Archaeon *Pyrococcus abyssi*. *Appl. Environ. Microbiol.* **63**, 1230–1236.
- Matray, T.J., and Kool, E.T. (1999). A specific partner for abasic damage in DNA. *Nature* **399**, 704–708.
- Matsui, I., Urushibata, Y., Shen, Y., Matsui, E., and Yokoyama, H. (2011). Novel structure of an N-terminal domain that is crucial for the dimeric assembly and DNA-binding of an archaeal DNA polymerase D large subunit from *Pyrococcus horikoshii*. *FEBS Lett.* **585**, 452–458.
- Matsunaga, F., Norais, C., Forterre, P., and Myllykallio, H. (2003). Identification of short “eukaryotic” Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Rep.* **4**, 154–158.
- Maxwell, B.A., and Suo, Z. (2013). Single-molecule investigation of substrate binding kinetics and protein conformational dynamics of a B-family replicative DNA polymerase. *J. Biol. Chem.* **288**, 11590–11600.
- Mayanagi, K., Kiyonari, S., Nishida, H., Saito, M., Kohda, D., Ishino, Y., Shirai, T., and Morikawa, K. (2011). Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1845–1849.
- McCulloch, S.D., Kokoska, R.J., Masutani, C., Iwai, S., Hanaoka, F., and Kunkel, T.A. (2004). Preferential cis-syn thymine dimer bypass by DNA polymerase eta occurs with biased fidelity. *Nature* **428**, 97–100.
- McHenry, C.S. (2011). DNA replicases from a bacterial perspective. *Annu. Rev. Biochem.* **80**, 403–436.
- Meng, X., Zhou, Y., Zhang, S., Lee, E.Y.C., Frick, D.N., and Lee, M.Y.W.T. (2009). DNA damage alters DNA polymerase delta to a form that exhibits increased discrimination against modified template bases and mismatched primers. *Nucleic Acids Res.* **37**, 647–657.
- Mikheikin, A.L., Lin, H.-K., Mehta, P., Jen-Jacobson, L., and Trakselis, M.A. (2009). A trimeric DNA polymerase complex increases the native replication processivity. *Nucleic Acids Res.* **37**, 7194–7205.
- Miyabe, I., Kunkel, T.A., and Carr, A.M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet.* **7**, e1002407.
- Mojica, F.J., Charbonnier, F., Juez, G., Rodríguez-Valera, F., and Forterre, P. (1994). Effects of salt and temperature on plasmid topology in the halophilic archaeon *Haloferax volcanii*. *J. Bacteriol.* **176**, 4966–4973.
- Morita, R., Nakane, S., Shimada, A., Inoue, M., Iino, H., Wakamatsu, T., Fukui, K., Nakagawa, N., Masui, R., and Kuramitsu, S. (2010). Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. *J Nucleic Acids* **2010**, 179594.

- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* *51 Pt 1*, 263–273.
- Myllykallio, H., Lopez, P., López-García, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H., and Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* *288*, 2212–2215.
- Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y., and Yang, W. (2012). Watching DNA polymerase  $\eta$  make a phosphodiester bond. *Nature* *487*, 196–201.
- Nakane, S., Nakagawa, N., Kuramitsu, S., and Masui, R. (2009). Characterization of DNA polymerase X from *Thermus thermophilus* HB8 reveals the POLXc and PHP domains are both required for 3'-5' exonuclease activity. *Nucleic Acids Res.* *37*, 2037–2052.
- Nakane, S., Ishikawa, H., Nakagawa, N., Kuramitsu, S., and Masui, R. (2012). The structural basis of the kinetic mechanism of a gap-filling X-family DNA polymerase that binds Mg(2+)-dNTP before binding to DNA. *J. Mol. Biol.* *417*, 179–196.
- Nealson, K. (2008). A Korarchaeote yields to genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 8805–8806.
- Nelson, J.R., Lawrence, C.W., and Hinkle, D.C. (1996). Thymine-thymine dimer bypass by yeast DNA polymerase zeta. *Science* *272*, 1646–1649.
- Nick McElhinny, S.A., and Ramsden, D.A. (2004). Sibling rivalry: competition between Pol X family members in V(D)J recombination and general double strand break repair. *Immunol. Rev.* *200*, 156–164.
- Nick McElhinny, S.A., Watts, B.E., Kumar, D., Watt, D.L., Lundström, E.-B., Burgers, P.M.J., Johansson, E., Chabes, A., and Kunkel, T.A. (2010a). Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 4949–4954.
- Nick McElhinny, S.A., Kumar, D., Clark, A.B., Watt, D.L., Watts, B.E., Lundström, E.-B., Johansson, E., Chabes, A., and Kunkel, T.A. (2010b). Genome instability due to ribonucleotide incorporation into DNA. *Nat. Chem. Biol.* *6*, 774–781.
- Nishida, H., Mayanagi, K., Kiyonari, S., Sato, Y., Oyama, T., Ishino, Y., and Morikawa, K. (2009). Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. *Proc. Natl. Acad. Sci. U.S.A.* *106*, 20693–20698.
- Norais, C., Hawkins, M., Hartman, A.L., Eisen, J.A., Myllykallio, H., and Allers, T. (2007). Genetic and Physical Mapping of DNA Replication Origins in *Haloferax volcanii*. *PLoS Genet* *3*.
- Northam, M.R., Garg, P., Baitin, D.M., Burgers, P.M., and Shcherbakova, P.V. (2006). A novel function of DNA polymerase  $\zeta$  regulated by PCNA. *EMBO J* *25*, 4316–4325.
- Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., Chee, G.-J., Hattori, M., Kanai, A., Atomi, H., et al. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* *39*, 3204–3223.
- Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Welte, W., and Marx, A. (2010). Replication through an abasic DNA lesion: structural basis for adenine selectivity. *EMBO J.* *29*, 1738–1747.
- Offre, P., Spang, A., and Schleper, C. (2013). Archaea in Biogeochemical Cycles. *Annu. Rev. Microbiol.*
- Ohmori, H., Friedberg, E.C., Fuchs, R.P., Goodman, M.F., Hanaoka, F., Hinkle, D., Kunkel, T.A.,

- Lawrence, C.W., Livneh, Z., Nohmi, T., et al. (2001). The Y-family of DNA polymerases. *Mol. Cell* **8**, 7–8.
- Olsen, G.J., and Woese, C.R. (1996). Lessons from an Archaeal genome: what are we learning from *Methanococcus jannaschii*? *Trends Genet.* **12**, 377–379.
- Pace, N.R. (1997). A molecular view of microbial diversity and the biosphere. *Science* **276**, 734–740.
- Palud, A., Villani, G., L'Haridon, S., Querellou, J., Raffin, J.-P., and Henneke, G. (2008). Intrinsic properties of the two replicative DNA polymerases of *Pyrococcus abyssi* in replicating abasic sites: possible role in DNA damage tolerance? *Mol. Microbiol.* **70**, 746–761.
- Paper, W., Jahn, U., Hohn, M.J., Kronner, M., Näther, D.J., Burghardt, T., Rachel, R., Stetter, K.O., and Huber, H. (2007). *Ignicoccus hospitalis* sp. nov., the host of "Nanoarchaeum equitans." *Int. J. Syst. Evol. Microbiol.* **57**, 803–808.
- Patel, P.H., and Loeb, L.A. (2001). Getting a grip on how DNA polymerases function. *Nat. Struct. Biol.* **8**, 656–659.
- Pavlov, Y.I., Shcherbakova, P.V., and Kunkel, T.A. (2001). In vivo consequences of putative active site mutations in yeast DNA polymerases alpha, epsilon, delta, and zeta. *Genetics* **159**, 47–64.
- Pearl, L.H. (2000). Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.* **460**, 165–181.
- Pelletier, H., Sawaya, M.R., Wolfle, W., Wilson, S.H., and Kraut, J. (1996). Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* **35**, 12742–12761.
- Pelve, E.A., Lindås, A.-C., Knöppel, A., Mira, A., and Bernander, R. (2012). Four chromosome replication origins in the archaeon *Pyrobaculum calidifontis*. *Mol. Microbiol.* **85**, 986–995.
- Perler, F.B., Kumar, S., and Kong, H. (1996). Thermostable DNA polymerases. *Adv. Protein Chem.* **48**, 377–435.
- Perrino, F.W., Harvey, S., Blans, P., Gelhaus, S., Lacourse, W.R., and Fishbein, J.C. (2005). Polymerization past the N2-isopropylguanine and the N6-isopropyladenine DNA lesions with the translesion synthesis DNA polymerases eta and iota and the replicative DNA polymerase alpha. *Chem. Res. Toxicol.* **18**, 1451–1461.
- Pisani, F.M., and Rossi, M. (1994). Evidence that an archaeal alpha-like DNA polymerase has a modular organization of its associated catalytic activities. *J. Biol. Chem.* **269**, 7887–7892.
- Pisani, F.M., De Martino, C., and Rossi, M. (1992). A DNA polymerase from the archaeon *Sulfolobus solfataricus* shows sequence similarity to family B DNA polymerases. *Nucleic Acids Res.* **20**, 2711–2716.
- Pochart, P., Doré, J., Lémann, F., Goderel, I., and Rambaud, J.C. (1992). Interrelations between populations of methanogenic archaea and sulfate-reducing bacteria in the human colon. *FEMS Microbiol. Lett.* **77**, 225–228.
- Potapova, O., Grindley, N.D.F., and Joyce, C.M. (2002). The mutational specificity of the Dbh lesion bypass polymerase and its implications. *J. Biol. Chem.* **277**, 28157–28166.
- Prakash, S., Johnson, R.E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem.* **74**, 317–353.
- Probst, A.J., Auerbach, A.K., and Moissl-Eichinger, C. (2013). Archaea on human skin. *PLoS ONE* **8**, e65388.

- Ramadan, K., Shevelev, I., and Hübscher, U. (2004). The DNA-polymerase-X family: controllers of DNA quality? *Nat. Rev. Mol. Cell Biol.* *5*, 1038–1043.
- Raven, N., Ladwa, N., Cossar, D., and Sharp, R. (1992). Continuous culture of the hyperthermophilic archaeum *Pyrococcus furiosus*. *Appl Microbiol Biotechnol* *38*, 263–267.
- Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA synthesis. *Annu. Rev. Biochem.* *57*, 349–374.
- Reiter, W.D., Hüdepohl, U., and Zillig, W. (1990). Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. *Proc. Natl. Acad. Sci. U.S.A.* *87*, 9509–9513.
- Richardson, T.T., Gilroy, L., Ishino, Y., Connolly, B.A., and Henneke, G. (2013a). Novel inhibition of archaeal family-D DNA polymerase by uracil. *Nucleic Acids Res.* *41*, 4207–4218.
- Richardson, T.T., Wu, X., Keith, B.J., Heslop, P., Jones, A.C., and Connolly, B.A. (2013b). Unwinding of primer-templates by archaeal family-B DNA polymerases in response to template-strand uracil. *Nucleic Acids Res.* *41*, 2466–2478.
- Riera, J., Robb, F.T., Weiss, R., and Fontecave, M. (1997). Ribonucleotide reductase in the archaeon *Pyrococcus furiosus*: a critical enzyme in the evolution of DNA genomes? *Proc. Natl. Acad. Sci. U.S.A.* *94*, 475–478.
- Rodriguez, A.C., Park, H.W., Mao, C., and Beese, L.S. (2000). Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon *Thermococcus* sp. 9 degrees N-7. *J. Mol. Biol.* *299*, 447–462.
- Rothwell, P.J., Mitaksov, V., and Waksman, G. (2005). Motions of the fingers subdomain of klenTaq1 are fast and not rate limiting: implications for the molecular basis of fidelity in DNA polymerases. *Mol. Cell* *19*, 345–355.
- Rothwell, P.J., Allen, W.J., Sisamakias, E., Kalinin, S., Felekyan, S., Widengren, J., Waksman, G., and Seidel, C.A.M. (2013). dNTP-dependent conformational transitions in the fingers subdomain of KlenTaq1 DNA polymerase: insights into the role of the “nucleotide-binding” state. *J. Biol. Chem.* *288*, 13575–13591.
- Rouillon, C., Henneke, G., Flament, D., Querellou, J., and Raffin, J.-P. (2007). DNA polymerase switching on homotrimeric PCNA at the replication fork of the euryarchaea *Pyrococcus abyssi*. *J. Mol. Biol.* *369*, 343–355.
- Ruiz, J.F., Juárez, R., García-Díaz, M., Terrados, G., Picher, A.J., González-Barrera, S., Fernández de Henestrosa, A.R., and Blanco, L. (2003). Lack of sugar discrimination by human Pol mu requires a single glycine residue. *Nucleic Acids Res.* *31*, 4441–4449.
- Russell, H.J., Richardson, T.T., Emptage, K., and Connolly, B.A. (2009). The 3'-5' proofreading exonuclease of archaeal family-B DNA polymerase hinders the copying of template strand deaminated bases. *Nucleic Acids Res.* *37*, 7603–7611.
- Sabouri, N., Viberg, J., Goyal, D.K., Johansson, E., and Chabes, A. (2008). Evidence for lesion bypass by yeast replicative DNA polymerases during DNA damage. *Nucleic Acids Res.* *36*, 5660–5667.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* *239*, 487–491.
- Sakofsky, C.J., Foster, P.L., and Grogan, D.W. (2012). Roles of the Y-family DNA polymerase Dbh in accurate replication of the *Sulfolobus* genome at high temperature. *DNA Repair (Amst.)* *11*, 391–400.

- Salguero, I., Guarino, E., Shepherd, M.E.A., Deegan, T.D., Havens, C.G., MacNeill, S.A., Walter, J.C., and Kearsley, S.E. (2012). Ribonucleotide reductase activity is coupled to DNA synthesis via proliferating cell nuclear antigen. *Curr. Biol.* *22*, 720–726.
- Sandigursky, M., and Franklin, W.A. (1999). Thermostable uracil-DNA glycosylase from *Thermotoga maritima* a member of a novel class of DNA repair enzymes. *Curr. Biol.* *9*, 531–534.
- Savino, C., Federici, L., Johnson, K.A., Vallone, B., Nastopoulos, V., Rossi, M., Pisani, F.M., and Tsernoglou, D. (2004). Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus solfataricus*. *Structure* *12*, 2001–2008.
- Schleper, C., Puehler, G., Holz, I., Gambacorta, A., Janekovic, D., Santarius, U., Klenk, H.P., and Zillig, W. (1995). *Picrophilus* gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. *J. Bacteriol.* *177*, 7050–7059.
- Schopf, J.W., and Packer, B.M. (1987). Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. *Science* *237*, 70–73.
- Seki, M., Masutani, C., Yang, L.W., Schuffert, A., Iwai, S., Bahar, I., and Wood, R.D. (2004). High-efficiency bypass of DNA damage by human DNA polymerase  $\eta$ . *EMBO J.* *23*, 4484–4494.
- Seki, M., Gearhart, P.J., and Wood, R.D. (2005). DNA polymerases and somatic hypermutation of immunoglobulin genes. *EMBO Rep.* *6*, 1143–1148.
- Servant, L., Cazaux, C., Bieth, A., Iwai, S., Hanaoka, F., and Hoffmann, J.-S. (2002). A role for DNA polymerase  $\beta$  in mutagenic UV lesion bypass. *J. Biol. Chem.* *277*, 50046–50053.
- Shamoo, Y., and Steitz, T.A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* *99*, 155–166.
- Shcherbakova, P.V., and Fijalkowska, I.J. (2006). Translesion synthesis DNA polymerases and control of genome stability. *Front. Biosci.* *11*, 2496–2517.
- Shen, J.C., Rideout, W.M., 3rd, and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.* *22*, 972–976.
- Sherman, P.A., and Fyfe, J.A. (1989). Enzymatic assay for deoxyribonucleoside triphosphates using synthetic oligonucleotides as template primers. *Anal. Biochem.* *180*, 222–226.
- Shevelev, I.V., and Hübscher, U. (2002). The 3' 5' exonucleases. *Nat. Rev. Mol. Cell Biol.* *3*, 364–376.
- Shewach, D.S. (1992). Quantitation of deoxyribonucleoside 5'-triphosphates by a sequential boronate and anion-exchange high-pressure liquid chromatographic procedure. *Analytical Biochemistry* *206*, 178–182.
- Shimizu, M., Gruz, P., Kamiya, H., Kim, S.-R., Pisani, F.M., Masutani, C., Kanke, Y., Harashima, H., Hanaoka, F., and Nohmi, T. (2003). Erroneous incorporation of oxidized DNA precursors by Y-family DNA polymerases. *EMBO Rep.* *4*, 269–273.
- Shuttleworth, G., Fogg, M.J., Kurpiewski, M.R., Jen-Jacobson, L., and Connolly, B.A. (2004). Recognition of the pro-mutagenic base uracil by family B DNA polymerases from archaea. *J. Mol. Biol.* *337*, 621–634.
- Silvian, L.F., Toth, E.A., Pham, P., Goodman, M.F., and Ellenberger, T. (2001). Crystal structure of a DinB family error-prone DNA polymerase from *Sulfolobus solfataricus*. *Nat. Struct. Biol.* *8*, 984–989.
- Slobodkin, Jeanthon, L'Haridon, Nazina, Miroshnichenko, and Bonch-Osmolovskaya (1999). Dissimilatory reduction of Fe(III) by thermophilic bacteria and archaea in deep subsurface petroleum reservoirs of western siberia. *Curr. Microbiol.* *39*, 99–102.

- Soler, N., Marguet, E., Cortez, D., Desnoves, N., Keller, J., van Tilbeurgh, H., Sezonov, G., and Forterre, P. (2010). Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. *Nucleic Acids Res.* *38*, 5088–5104.
- Spang, A., Hatzepichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., Streit, W., Stahl, D.A., Wagner, M., and Schleper, C. (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* *18*, 331–340.
- Starkuviene, V., and Fritz, H.-J. (2002). A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res.* *30*, 2097–2102.
- Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U.S.A.* *90*, 6498–6502.
- Stetter, K.O. (1982). Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105 °C. *Nature* *300*, 258–260.
- Stetter, K.O., Thomm, M., Winter, J., Wildgruber, G., Huber, H., Zillig, W., Jané-Covic, D., König, H., Palm, P., and Wunderl, S. (1981). *Methanothermus fervidus*, sp. nov., a novel extremely thermophilic methanogen isolated from an Icelandic hot spring. *Zentralblatt Für Bakteriologie Mikrobiologie Und Hygiene: I. Abt. Originale C: Allgemeine, Angewandte Und Ökologische Mikrobiologie* *2*, 166–178.
- Stetter, K.O., König, H., and Stackebrandt, E. (1983). *Pyrodictium* gen. nov., a New Genus of Submarine Disc-Shaped Sulphur Reducing Archaeobacteria Growing Optimally at 105°C. *Syst. Appl. Microbiol.* *4*, 535–551.
- Subuddhi, U., Hogg, M., and Reha-Krantz, L.J. (2008). Use of 2-aminopurine fluorescence to study the role of the beta hairpin in the proofreading pathway catalyzed by the phage T4 and RB69 DNA polymerases. *Biochemistry* *47*, 6130–6137.
- Swan, M.K., Johnson, R.E., Prakash, L., Prakash, S., and Aggarwal, A.K. (2009). Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase delta. *Nat. Struct. Mol. Biol.* *16*, 979–986.
- Tahirov, T.H., Makarova, K.S., Rogozin, I.B., Pavlov, Y.I., and Koonin, E.V. (2009). Evolution of DNA polymerases: an inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. *Biol. Direct* *4*, 11.
- Tang, M., Pham, P., Shen, X., Taylor, J.S., O'Donnell, M., Woodgate, R., and Goodman, M.F. (2000). Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* *404*, 1014–1018.
- Tang, X.-F., Shen, Y., Matsui, E., and Matsui, I. (2004). Domain topology of the DNA polymerase D complex from a hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry* *43*, 11818–11827.
- Thauer, R.K., and Shima, S. (2006). Biogeochemistry: methane and microbes. *Nature* *440*, 878–879.
- Tissier, A., Frank, E.G., McDonald, J.P., Iwai, S., Hanaoka, F., and Woodgate, R. (2000). Misinsertion and bypass of thymine-thymine dimers by human DNA polymerase iota. *EMBO J.* *19*, 5259–5266.
- Torreblanca, M., Rodriguez-Valera, F., Juez, G., Ventosa, A., Kamekura, M., and Kates, M. (1986). Classification of Non-alkaliphilic Halobacteria Based on Numerical Taxonomy and Polar Lipid Composition, and Description of Haloarcula gen. nov. and Haloferax gen. nov. *Systematic and Applied Microbiology* *8*, 89–99.
- Traut, T.W. (1994). Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* *140*, 1–22.
- Trincao, J., Johnson, R.E., Escalante, C.R., Prakash, S., Prakash, L., and Aggarwal, A.K. (2001). Structure of the catalytic core of *S. cerevisiae* DNA polymerase eta: implications for translesion DNA

synthesis. *Mol. Cell* **8**, 417–426.

Tsunoda, H., Kudo, T., Masaki, Y., Ohkubo, A., Seio, K., and Sekine, M. (2011). Biochemical behavior of N-oxidized cytosine and adenine bases in DNA polymerase-mediated primer extension reactions. *Nucleic Acids Res.* **39**, 2995–3004.

Uemori, T., Sato, Y., Kato, I., Doi, H., and Ishino, Y. (1997). A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes to Cells* **2**, 499–512.

Vértessy, B.G., and Tóth, J. (2009). Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc. Chem. Res.* **42**, 97–106.

Wang, J., Sattar, A.K., Wang, C.C., Karam, J.D., Konigsberg, W.H., and Steitz, T.A. (1997). Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087–1099.

Wang, Z., Castaño, I.B., De Las Peñas, A., Adams, C., and Christman, M.F. (2000). Pol kappa: A DNA polymerase required for sister chromatid cohesion. *Science* **289**, 774–779.

Wardle, J., Burgers, P.M.J., Cann, I.K.O., Darley, K., Heslop, P., Johansson, E., Lin, L.-J., McGlynn, P., Sanvoisin, J., Stith, C.M., et al. (2008). Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res.* **36**, 705–711.

Watt, D.L., Johansson, E., Burgers, P.M., and Kunkel, T.A. (2011). Replication of ribonucleotide-containing DNA templates by yeast replicative polymerases. *DNA Repair (Amst.)* **10**, 897–902.

Wilson, R.C., Jackson, M.A., and Pata, J.D. (2013). Y-family polymerase conformation is a major determinant of fidelity and translesion specificity. *Structure* **21**, 20–31.

Wilson, S.H., Sobol, R.W., Beard, W.A., Horton, J.K., Prasad, R., and Vande Berg, B.J. (2000). DNA polymerase beta and mammalian base excision repair. *Cold Spring Harb. Symp. Quant. Biol.* **65**, 143–155.

Woese, C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5088–5090.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576–4579.

Van Wolferen, M., Ajon, M., Driessen, A.J.M., and Albers, S.-V. (2013). How hyperthermophiles adapt to change their lives: DNA exchange in extreme conditions. *Extremophiles* **17**, 545–563.

Woffle, W.T., Johnson, R.E., Minko, I.G., Lloyd, R.S., Prakash, S., and Prakash, L. (2005). Human DNA polymerase iota promotes replication through a ring-closed minor-groove adduct that adopts a syn conformation in DNA. *Mol. Cell. Biol.* **25**, 8748–8754.

Wu, Y., Wilson, R.C., and Pata, J.D. (2011). The Y-family DNA polymerase Dpo4 uses a template slippage mechanism to create single-base deletions. *J. Bacteriol.* **193**, 2630–2636.

Wuchter, C., Abbas, B., Coolen, M.J.L., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., et al. (2006). Archaeal nitrification in the ocean. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12317–12322.

Xia, S., Wang, J., and Konigsberg, W.H. (2013). DNA mismatch synthesis complexes provide insights into base selectivity of a B family DNA polymerase. *J. Am. Chem. Soc.* **135**, 193–202.

Yamasaki, K., Urushibata, Y., Yamasaki, T., Arisaka, F., and Matsui, I. (2010). Solution structure of the N-terminal domain of the archaeal D-family DNA polymerase small subunit reveals evolutionary

relationship to eukaryotic B-family polymerases. *FEBS Lett.* *584*, 3370–3375.

Yamtich, J., and Sweasy, J.B. (2010). DNA polymerase Family X: Function, structure, and cellular roles. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* *1804*, 1136–1150.

Yang, G., Lin, T., Karam, J., and Konigsberg, W.H. (1999). Steady-state kinetic characterization of RB69 DNA polymerase mutants that affect dNTP incorporation. *Biochemistry* *38*, 8094–8101.

Yang, S., Gong, H., Zhang, L., Liu, Y., and He, Z.-G. (2010). Characterization of physical and functional interactions between eukaryote-like Orc1/Cdc6 proteins and Y-family DNA polymerase in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochem. Biophys. Res. Commun.* *396*, 755–762.

Yao, N.Y., Schroeder, J.W., Yurieva, O., Simmons, L.A., and O'Donnell, M.E. (2013). Cost of rNTP/dNTP pool imbalance at the replication fork. *Proc. Natl. Acad. Sci. U.S.A.* *110*, 12942–12947.

Zahn, K.E., Belrhali, H., Wallace, S.S., and Doublé, S. (2007). Caught bending the A-rule: crystal structures of translesion DNA synthesis with a non-natural nucleotide. *Biochemistry* *46*, 10551–10561.

Zbinden, M., and Cambon-Bonavita, M.-A. (2003). Occurrence of Deferribacterales and Entomoplasmatales in the deep-sea Alvinocarid shrimp *Rimicaris exoculata* gut. *FEMS Microbiol. Ecol.* *46*, 23–30.

Zeng, X., Birrien, J.-L., Fouquet, Y., Cherkashov, G., Jebbar, M., Querellou, J., Oger, P., Cambon-Bonavita, M.-A., Xiao, X., and Prieur, D. (2009). *Pyrococcus CH1*, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life. *ISME J* *3*, 873–876.

Zhao, Y., Jeruzalmi, D., Moarefi, I., Leighton, L., Lasken, R., and Kuriyan, J. (1999). Crystal structure of an archaeobacterial DNA polymerase. *Structure* *7*, 1189–1199.

Zillig, W., Tu, J., and Holz, I. (1981). Thermoproteales: a third order of thermoacidophilic archaeobacteria. *Nature* *293*, 85–86.

Zobell, C.E. (1941). Studies on marine bacteria. The cultural requirement of heterotrophic aerobes. *J. Marine Research*, *4*, 42-76.







Séquences peptidiques et poids moléculaires de PabPolB et de ses mutants, avec leur Tag.

En bleu : séquence extra His-Tag.

En vert : mutation D215A → activité exonucléasique abolie.

En rouge : mutation pour le T-cleft fixant l'ion métallique divalent.

- **PabPolB Exo +**

**MGSSHHHHHHSSGLVPRGSHAG**MIIDADYITEDGKPIIRIFKKEKGEFKVEYDRTFRPYIYALLKDDSAI  
 DEVKKITAERHGKIVRITEVEKVQKKFLGRPIEVWKLYLEHPQDVPAIREKIREHPAVVDIFEYDIPFAKR  
 YLIDKGLTPMEGNEELTFLAVDIETLYHEGEEFGKGPIIMISYADEEGAKVITWKSIDLPHYVEVSSERE  
 MIKRLVKVIREKDPDVIITYNGDNF**D**FPYLLKRAEKLGIKPLGRDNSEPKMQRMGDSLAVEIKGRIHFD  
 LFPVIRRTINLPTYTLEAVYEAIFGKSKEKVYAHEIAEAWETGKGLERVAKYSMEDAKVTFELGKEFFP  
 MEAQLARLVGQPVWDVSRSSSTGNLVEWFLLRKAYERNELAPNKPDEREYERRLRESYEGGYVKEPE  
 KGLWEGIVSLDFRSLYPSIIITHNVSPDTLNRENCKEYDVAPQVGHRFCKDFPGFIPSLGNLLEERQKI  
 KCRMKESKDPVEKLLDYRQRAIKILANSYGGYGYAKARWYCKECAESVTAWGRQYIDLVRRELES  
 RGFKVLYIDTDGLYATIPGAKHEEIKEKALKFVEYINSKLPGLLELEYEGFYARGFFVTKKKYALIDEEGK  
 IVTRGLEIVRRDWSEIAKETQAKVLEAILKHGNVDEAVKIVKEVTEKLSKYEIPPEKLVIEYEQITRPLSEYK  
 AIGPHVAVAKRLAAKGVKVPGMVIGYIVLRGDGPISKRAIAIEEFDPKKHKYDAEYIENQVLPVERIL  
 RAFGYRKEDLKYQTKQVGLGAWLKF

**PM= 91 784 Da**

- **PabPolB Exo –**

**MGSSHHHHHHSSGLVPRGSHAG**MIIDADYITEDGKPIIRIFKKEKGEFKVEYDRTFRPYIYALLKDDSAI  
 DEVKKITAERHGKIVRITEVEKVQKKFLGRPIEVWKLYLEHPQDVPAIREKIREHPAVVDIFEYDIPFAKR  
 YLIDKGLTPMEGNEELTFLAVDIETLYHEGEEFGKGPIIMISYADEEGAKVITWKSIDLPHYVEVSSERE  
 MIKRLVKVIREKDPDVIITYNGDNF**A**FPYLLKRAEKLGIKPLGRDNSEPKMQRMGDSLAVEIKGRIHFD  
 LFPVIRRTINLPTYTLEAVYEAIFGKSKEKVYAHEIAEAWETGKGLERVAKYSMEDAKVTFELGKEFFP  
 MEAQLARLVGQPVWDVSRSSSTGNLVEWFLLRKAYERNELAPNKPDEREYERRLRESYEGGYVKEPE  
 KGLWEGIVSLDFRSLYPSIIITHNVSPDTLNRENCKEYDVAPQVGHRFCKDFPGFIPSLGNLLEERQKI  
 KCRMKESKDPVEKLLDYRQRAIKILANSYGGYGYAKARWYCKECAESVTAWGRQYIDLVRRELES  
 RGFKVLYIDTDGLYATIPGAKHEEIKEKALKFVEYINSKLPGLLELEYEGFYARGFFVTKKKYALIDEEGK  
 IVTRGLEIVRRDWSEIAKETQAKVLEAILKHGNVDEAVKIVKEVTEKLSKYEIPPEKLVIEYEQITRPLSEYK  
 AIGPHVAVAKRLAAKGVKVPGMVIGYIVLRGDGPISKRAIAIEEFDPKKHKYDAEYIENQVLPVERIL  
 RAFGYRKEDLKYQTKQVGLGAWLKF

**PM= 91 740 Da**

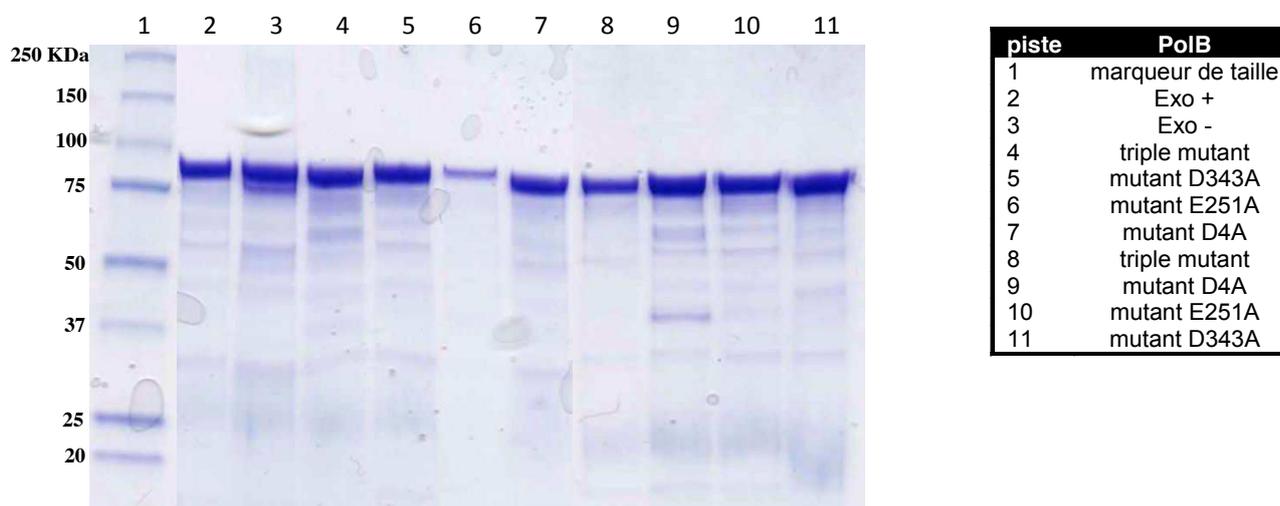
▪ **PabPolB Triple mutant**

**MGSSHHHHHSSGLVPRGSHAG**MII**A**ADYITEDGKPIIRIFKKEKGFEKVEYDRTFRPYIYALLKDDSAI  
 DEVKKITAERHGKIVRITEVEKVVQKKFLGRPIEVWKLYLEHPQDVPAIREKIREHPAVVDIFEYDIPFAKR  
 YLIDKGLTPMEGNEELTFLAVDIETLYHEGEEFGKGPIIMISYADEEGAKVITWKSIDLPHYVEVSSERE  
 MIKRLVKVIREKDPDVIITYNGDNF**A**FPYLLKRAEKLGIKPLGRDNSEPKMQRMGDSLAV**A**IKGRIHFD  
 LFPVIRRTINLPTYTLEAVYEAIFGKSKEKVYAHEIAEAWETGKGLERVAKYSMEDAKVTFELGKEFFP  
 MEAQLARLVGQP**V**AVSR**S**STGNLVEWFLLRKAYERNELAPNKPDEREYERRLRESYEGGYVKEPE  
 KGLWEGIVSLDFRSLYPSIIITHNVSPDTLNRENCKEYDVAPQVGHFRFCKDFPGFIPSLGNNLEERQKI  
 KKRMKESKDPVEKKLLDYRQRAIKILANSYYGYGYAKARWYCKECAESVTAWGRQYIDLVRRELES  
 RGFKVLYIDTDGLYATIPGAKHEEIKEKALKFVEYINSKLPLGLELEYEGFYARGFFVTKKKYALIDEEGK  
 IVTRGLEIVRRDWSEIAKETQAKVLEAILKHGNVDEAVKIVKEVTEKLSKYEIPPEKLVIEQITRPLSEYK  
 AIGPHVAVAKRLAAKGVKVPGMVIGYIVLRGDGPISKRAIAIEEFDPKKHKYDAEYIENQVLPVERIL  
 RAFGYRKEDLKYQTKQVGLGAWLKF

**PM= 91 594 Da**

**ANNEXE 2**

Contrôle du poids moléculaire et de la pureté de l'ensemble des PabPolBs produites.



Profil électrophorétique en gel de polyacrylamide (4 - 20 %) dénaturant (SDS) des différentes productions de protéines PabPolB et purifiées en tampon B (pistes 4 à 7) et en tampon C (pistes 8 à 11). 3,2 µg de protéines dénaturées ont été déposés par puits. La migration électrophorétique a été effectuée en présence d'un marqueur de poids moléculaire (piste 1).

Composition du milieu de culture 2216 + solution C.

Fonction	Composant	Quantité pour 1L
Source de carbone et/ou d'énergie	<b>Extrait de levure</b>	0.5 g
	<b>Peptone</b>	2 g
Sels et minéraux	<b>Sels de mer</b>	30 g
	<b>K<sub>2</sub>HPO<sub>4</sub></b>	0.05 g
	<b>H<sub>3</sub>BO<sub>3</sub></b>	7.5 mg
	<b>Na<sub>2</sub>WO<sub>4</sub>, 2 H<sub>2</sub>O</b>	3.3 mg
	<b>Na<sub>2</sub>MoO<sub>4</sub>, 2 H<sub>2</sub>O</b>	0.15 mg
	<b>Na<sub>2</sub>SeO<sub>3</sub></b>	5 µg
Indicateur coloré	<b>Solution de rézaurine (1g/L)</b>	0.5 mL
Tampon pH	<b>Sel dissodium PIPES</b>	3.46 g
Accepteur d'électrons	<b>Soufre élémentaire</b>	10 g
Agent réducteur	<b>Na<sub>2</sub>S 5%</b>	10 mL

Composition du milieu de culture SME modifié.

Fonction	Composants	Quantité pour 1L
Source de carbone et/ou d'énergie	<b>Extrait de levure</b>	1 g
	<b>Peptone</b>	2 g
Sels et minéraux	<b>NaCl</b>	28 g
	<b>Solution de sels de magnésium</b>	10 mL
	<b>Solution A</b>	1 mL
	<b>Solution B</b>	1 mL
	<b>Solution C</b>	1 mL
Vitamines	<b>Solution de vitamines</b>	0.5 mL
	<b>Solution de rézazurine (1g/L)</b>	0.5 mL
Tampon pH	<b>Sel dissodium PIPES</b>	6.05 g
Accepteur d'électrons	<b>L-cystéine</b>	0.5 g
	(pH ajusté à 7)	

Solution A	
Composant	Quantité pour 1L
MnSO <sub>4</sub> , 4 H <sub>2</sub> O	9 g
ZnSO <sub>4</sub> , 7 H <sub>2</sub> O	2.5 g
NiCl <sub>2</sub> , 6 H <sub>2</sub> O	2.5 g
AlK(SO <sub>4</sub> ) <sub>2</sub> , 12 H <sub>2</sub> O	0.3 g
CoCl <sub>2</sub> , 6 H <sub>2</sub> O	0.3 g
CuSO <sub>4</sub> , 5 H <sub>2</sub> O	0.15 g
Eau milliQ qsp 1 L	

Solution B	
Composant	Quantité pour 1L
CaCl <sub>2</sub> , 2 H <sub>2</sub> O	56 g
NaBr	25 g
KCl	16 g
KI	10 g
SrCl <sub>2</sub> , 6 H <sub>2</sub> O	4 g
Eau milliQ qsp 1 L	

Solution C	
Composant	Quantité pour 1L
K <sub>2</sub> HPO <sub>4</sub>	50 g
H <sub>3</sub> BO <sub>3</sub>	7.5 g
Na <sub>2</sub> WO <sub>4</sub> , 2 H <sub>2</sub> O	3.3 g
Na <sub>2</sub> MoO <sub>4</sub> , 2 H <sub>2</sub> O	0.15 g
Na <sub>2</sub> SeO <sub>3</sub>	0.005 g
Eau milliQ qsp 1 L	

Solution D	
Composant	Quantité pour 1L
FeCl <sub>2</sub> , 4 H <sub>2</sub> O	10 g
Eau milliQ qsp 1 L	

<b>Solution de Vitamines</b>	
<b>Composant</b>	<b>Quantité pour 1L</b>
Biotine	40 mg
Acide Folique	40 mg
Pyridoxine-HCl	200 mg
Thiamine HCl	100 mg
Riboflavine	100 mg
Acide nicotinique	100 mg
DL. Calcium pantothenate	100 mg
Vitamine B12	2 mg
Acide lipoïque	100 mg
Ethanol dans eau distillée 50 % V/V	qsp 1 L

<b>Solution de sels de magnesium</b>	
<b>Composant</b>	<b>Quantité pour 1L</b>
FeCl <sub>2</sub> , 4H <sub>2</sub> O	10 g
Eau milliQ qsp 1 L	





## ADN polymérase et acides nucléiques endommagés chez les Archaea: maintenance génomique et biotechnologies

**Résumé :** Les Archaea hyperthermophiles sont des micro-organismes particulièrement adaptés à la vie à très haute température. Les températures élevées sont responsables de la création de diverses lésions à l'ADN. Pourtant, le taux de mutations spontanées chez l'Archaea *Sulfolobus solfataricus* est proche de celui mesuré chez des organismes mésophiles. Ceci laisse suggérer que les Archaea doivent être équipées de mécanismes spécialisés efficaces dans la reconnaissance et réparation des dommages à l'ADN. La réplication de l'ADN est un mécanisme fonctionnellement conservé dans les trois domaines du Vivant assuré par des enzymes clés, les ADN polymérase. A ce jour, chez *Pyrococcus abyssi* (*Pab*), Euryarchaea hyperthermophile, deux familles de pol ont été identifiées, une pol répliquative de la famille B, présente dans les trois règnes du Vivant et une pol répliquative la famille D, spécifique de l'embranchement des Euryarchaea. Tous les membres de la famille B des pols archéennes possèdent une signature structurale unique impliquée dans la reconnaissance spécifique des bases endommagées (uracile et hypoxanthine). Ce doctorat a permis de démontrer que cette poche pouvait également accommoder les bases canoniques (A, T, C, G) contribuant ainsi à la haute fidélité de ces enzymes. Des structures cristallographiques à très haute résolution de l'ADN polymérase B (*PabPolB*) ont permis d'identifier la présence d'un ion métallique divalent situé à proximité de cette poche de reconnaissance. Il a été suggéré que cet ion modulerait la reconnaissance des bases désaminées et le glissement de l'ADN polymérase au cours de la synthèse d'ADN. De plus, nous avons montré que les ions catalytiques divalents ont un rôle central dans la modulation des propriétés intrinsèques de *PabPolB*. L'ion calcium est utilisé par *PabPolB* pour mener à bien la synthèse d'ADN mais engendre des caractéristiques cinétiques différentes de celles conférées par l'ion magnésium universel. D'autre part, nous avons déterminé pour la première fois chez un organisme archéen la concentration intracellulaire de dNTPs et NTPs. Leurs concentrations sont relativement plus élevées que celles mesurées par exemple chez l'eucaryote *Saccharomyces cerevisiae*, suggérant un mécanisme constitutif de réaction de protection à des agressions de l'ADN. L'ensemble de ces résultats participe à la compréhension des mécanismes moléculaires qui régissent la maintenance de l'intégrité du génome et, donc, de la vie des organismes hyperthermophiles.

**Mots clés :** ADN, Archéa, polymérase, *Pyrococcus*.

### DNA polymerases and damaged nucleic acids in archaea : Genome maintenance and Biotechnology

**Abstract :** Hyperthermophilic Archaea that thrive under harsh environments (elevated temperature, pH shifts, and ionizing radiations) are supposed to be exposed to massive DNA damages. However, the mutations frequencies in hyperthermophilic Archaea are comparable with those of other microorganisms (1) indicating they are equipped with unique and efficient molecular mechanisms to ensure their genome integrity. DNA replication is an essential and conserved process among the three domains of life. DNA polymerases are central enzymes involved in the joining of deoxyribonucleoside 5'-triphosphates (dNTPs) to form the growing DNA chain. In *Pyrococcus abyssi* (*Pab*), two families DNA polymerases have been described as replicases, one family B (*PabPol B*) with structural diversity and common mechanisms among all organisms in the three domains of life, and one family D found exclusively in Euryarchaea. All members of archaeal family B DNA polymerases possess a unique structural feature involved in the recognition of deaminated bases (uracil and hypoxanthine) called uracil-pocket. During my PhD, it has been shown that not only deaminated but also canonical bases (A, T, C, G) could enter this pocket in *PabPolB*. We therefore renamed it as the base-checking pocket and proposed a role in the high fidelity of *PabPolB*. High-resolution crystal structures of *PabPolB* highlighted the presence of divalent metal ion to the proximity of the base-checking pocket. It has been suggested that this metal ion may modulate the recognition of deaminated bases and the translocation of *PabPolB* on DNA during DNA synthesis. Moreover, the crucial role of metal ion on DNA synthesis and exonuclease activity by *PabPolB* has been evaluated. DNA polymerisation in the presence of calcium was as effective as the universal magnesium ion while showing different time-course kinetics. For the first time, intracellular concentrations of nucleotides (dNTPs and NTPs) have been determined in an archaeal microorganism. dNTPs and NTPs concentrations seem rather elevated for *Pab* than in the eukaryotic *Saccharomyces cerevisiae* cells, suggesting a constitutive mechanism for protecting the genome from DNA damage. Overall, these results contributed to unravel specific molecular mechanisms involved in the maintenance of the genome integrity with a particular emphasis on molecules (DNA polymerases, dNTPs and NTPs) crucial for cellular life.

**Keywords :** DNA, Archéa, polymerase, *Pyrococcus*.