



HAL
open science

**Maintenance génomique chez l'Archaea
hyperthermophile *Pyrococcus abyssi*: découverte de
nouvelles interactions physiques et caractérisation
fonctionnelle**

Julien Briffotiaux

► **To cite this version:**

Julien Briffotiaux. Maintenance génomique chez l'Archaea hyperthermophile *Pyrococcus abyssi*: découverte de nouvelles interactions physiques et caractérisation fonctionnelle. Biochimie [q-bio.BM]. Université de Bretagne occidentale - Brest, 2008. Français. NNT: . tel-00273692

HAL Id: tel-00273692

<https://theses.hal.science/tel-00273692>

Submitted on 15 Apr 2008

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.

Maintenance génomique chez l'Archaea hyperthermophile *Pyrococcus abyssi* : découverte de nouvelles interactions physiques et caractérisation fonctionnelle

THÈSE

présentée et soutenue publiquement le 31 janvier 2008

pour l'obtention du

Doctorat de l'Université de Bretagne Occidentale

Discipline : Microbiologie

par

Julien BRIFFOTAUX

Composition du jury

Rapporteurs : Docteur Christiane ELIE, CNRS, Université de Paris XI, UMR8621.
Docteur Jean ARMENGAUD, CEA, Bagnol sur Cèze.

Examineurs : Docteur Charles PINEAU, Inserm, Rennes.
Professeur Marie-Thérèse THEBAULT, Université de Bretagne Occidentale.
Professeur Mohamed JEBBAR, Université de Bretagne Occidentale, UMR 6197.
Docteur Jean-Paul RAFFIN, CNRS, UMR 6197, Brest.
Docteur Didier FLAMENT, Ifremer, UMR 6197, Brest.

Remerciements

A large, empty rectangular box with a thin black border, occupying most of the page below the title. It is intended for the user to write their acknowledgments.

Tout un chacun peut dans sa vie quotidienne, se poser des questions sur l'origine, le commencement des choses. Comment est né le temps, comment est arrivée la poule qui a pondu le tout premier œuf, comment un objet disparaît juste au moment où j'en ai le plus besoin ou comment un plan que je consulte en ville indique déjà « vous êtes ici » alors que je viens juste d'arriver...

Par la pensée, il est possible de voyager dans le temps. Nous sommes en septembre 2004, je viens juste d'arriver à Brest. Le plan me confirme que je suis bien ici, tout va bien ; Didier m'attend à la gare, je découvre cette ville à l'architecture bizarre, sa grande place de la liberté stalinienne, des goélands partout, l'absence de montagne, quelque soit l'endroit où l'on porte les yeux... Un choc culturel m'attendait en Bretagne. Provenant de ma Savoie d'origine, parlant un langage bien chargé en y, et habitué à la gastronomie bien grasse (essentiellement à base de patates, lards et reblochon), me voilà aux pays des sonneurs de biniou, mangeurs de crêpes et parlant une langue inintelligible. Le milieu semble hostile. On ne vient pas à Brest par hasard, ce n'est point une ville que l'on traverse, où l'on passe. Si on vient à Brest, c'est que l'on a quelque chose à y faire. Moi j'y ai effectué ma formation, en tant que jeune Padawan, aboutissant à la thèse.

Ma thèse s'est déroulée au laboratoire de Microbiologie des Environnements Extrêmes, à l'Ifremer de Brest. Je remercie le docteur Joël Querellou, responsable du laboratoire, pour m'avoir permis d'effectuer ma thèse dans de bonnes conditions. Je remercie aussi ici l'Océan que j'ai pu contempler durant ces trois années.

Au cours de cette aventure, j'ai eu le privilège de partager mon bureau avec mon Maître Jedi, Didier Flament. Grand maître de l'eau et du vent, arroseur de plantes, expert de la godille dans la brume et amateur de vieux gréements, sa maîtrise de la Force lui permet de concevoir, analyser et découvrir de nouveaux phénomènes scientifiques. Je suis très admiratif envers Didier pour sa simplicité, son calme et sa maîtrise de la conception et de la gestion de projet de recherche. Sa maîtrise de la grande voile et du gouvernail lui a permis de superviser mes travaux et je lui en suis très reconnaissant. Je le remercie pour m'avoir transmis ses compétences en matière de pêche aux protéines.

Jean-Paul Raffin, un barde d'aspect druidesque et maître de la Biochimie, venu d'Alsace, il est arrivé en Bretagne comme réfugié politique. Je remercie ce personnage singulier, auteur des raffinades, qui, entre une fureur contre l'organisation de la recherche et une remarque insolite sur la couleur des cornemuses, a été mon directeur de thèse.

Je remercie également Ghislaine Henneke, le troisième pilier du groupe « Maintenance génomique chez les Archées ». Son énergie et sa rigueur scientifique contribuent indéniablement au dynamisme et à la productivité de l'équipe.

Au bout de quelques mois, a débarqué au laboratoire un personnage tout droit issue de féerie et de l'imaginaire breton, Adeline Palud. C'était une korrigan de Camaret, une petite personne avec de grands yeux dans lesquels j'aurai bien pu me

noyer, parfois affublé d'une jupe verte, et de tongs elfiques. Je la remercie pour sa vitalité et son sourire qui rendent les journées au laboratoire agréables.

Adeline est souvent associée à ces compères de bavardage, commérage et découverte des nouveaux potins lors des pauses sur la terrasse, à l'heure du thé. Parmi ces pipelettes, Sophie Schmitt et sa bonne humeur quotidienne a fortement contribué à la bonne ambiance du laboratoire. Audrey le Goaziou (une fille de Prat !) n'était bien sûr jamais très loin. Je remercie particulièrement Audrey et Sophie pour avoir éclairé ma lanterne en m'expliquant la différence entre les jupes et les robes. Parmi les ex. du labo, Nathalie Benz et son illustre voiture a également rendu son passage mémorable !

Le laboratoire est également fréquenté par des Farfadets aux drôles de coutumes : Benoît Castrec et Sébastien Lelaz. Ces deux jeunes bretons, vivent en ermites à l'extérieur du laboratoire dans une grotte qu'ils appellent algéco. Ils apparaissent à leur paillasse sans crier gare afin d'élaborer quelques potions et mixtures dont ils gardent le secret. De ces expériences saugrenues, Benoît a été contaminé par une grave maladie, la polystyrénophobie aiguë. Cette exposition lui cause également quelques problèmes pour s'orienter au niveau de la place de la Liberté. Plus à l'est, c'est l'ouest.

C'était il y a un an. Les personnages de l'imaginaire me poursuivent. Une petite personne féérique à la voix douce, vêtue le plus souvent d'un gros pull pour se protéger des rigueurs du climat breton s'est jointe à notre équipe. Caroline Sotto, c'est toi « ma stagiaire préférée ». Merci pour ton efficacité pour la production des protéines, mais surtout pour ta gentillesse et ton amitié. Tu es une belle personne, simple, avec le cœur du bon côté. Par ailleurs, je ne vois plus le vortex de la même manière depuis ta rencontre !!

Il y a d'autres lutins que j'ai eu le plaisir de rencontrer. Claire Tessier (merci pour m'avoir ramener chez moi de nombreuses fois et nous avons une histoire de batterie en commun!), Henintsoa Randrianalivelo (une journée sympa sur la presqu'île de Crozon, les discussions sur Madagascar), Magalie La Bretonne qui m'a cédé son héritage (même si je n'ai jamais retrouvé le fameux portoir bleu !), Eloïse Delage (initiatrice du journal de potins du labo : Persiflage), Virginie Estébanes et son accent chantant, Nathalie Byrne (de bons souvenir grâce à l'association des doctorants), An Shu (zhù nǐ hǎo yùn 祝你好运.), Léna Rajouan, une belle personne joyeuse et positive (c'était cool le concert de Renan Luce au jeudi du port !), Lucile Durand (cavalière du Rohan, sa mère devait être une amazone et son père un centaure), Mathilde (une fille de Camaret !), Valérie, Cassandre Lazar (pour sa maîtrise de la conception de la pâte à pancake), Marie (pour ses gateaux au Chocolat), Laetitia (pour les poissons rouges), Anne Postec, Floriane, la fée Mélusine, Amandine.

Des gnomes et gobelins ont également fréquenté ces lieux, le majestueux Sébastien Lemoine (Spécialiste En Bêtise Accentuant sa Scrupuleuse Tendance à l'Ironie en Essayant d'être Naturel), l'artiste Christophe Rouillon (maintenant en

compagnie des petites anglaises !), Mathieu Gonnet (oui, j'ai appris à parler italien il y a quelques années !), Erwan Roussel (fais gaffe aux deux farfadets pas loin de ton nouveau bureau), Erwan (merci pour le pliage de la tente), Erwan (l'autre), Vincent Lecomte (gnome enchanteur en botanique et ornithologie), Anthony et Xavier du laboratoire limithrophe avec lequel nous entretenons néanmoins des rapports cordiaux.

Merci également à tout le personnel du LM2E qui a rendu ce séjour à l'Ifremer agréable.

J'ai eu la chance, au cours de ce doctorat, de me rendre en environnement extrême, dans un pays étrange et magnifique : l'Islande. En plus d'être parti à la recherche des Trolls et de la mythologie nordique, je remercie les deux chercheurs vikings de la société Prokaria qui m'ont accueilli et encadré lors de ce stage, les docteurs Viggó Marteinson et Ólafur Friðjónsson.

Je remercie les docteurs Raymund Wellinger, Michel Tremblay et Eric Cohen qui m'ont permis de présenter mes travaux au Canada. Je remercie également Elise, Charlotte et Jane pour leur sympathique accueil lors de ma première expérience de Couchsurfing !

Je n'oublie pas la fameuse crème aux œufs de la cantine, apportant un sourire extatique à tous ceux qui en mangent ; le Pic-vert qui est régulièrement venu me rendre visite lors de cette période de rédaction, juste en face de mon bureau.

Il y a quelques artistes qui, par leurs œuvres musicales, ont sans aucun doute influencer le quotidien. Les Têtes Raides, Debout sur le Zinc, Mano Solo, Thomas Fersen, Yann Tiersen, Ben Harper, Xavier Rudd et Elliott Smith sont parmi ceux qui ont comptés. J'ai bien sûr eu plaisir de pénétrer dans les univers épiques ou intimistes, merveilleux et lyriques de Joe Hisaishi et John Williams.

Après différentes lectures, j'ai aussi eu de l'admiration pour des chercheurs humanistes comme Axel Kahn, Hubert Reeves et Albert Jacquard. Grâce à eux, j'ai conscience des intrications rapprochant la science, l'économie, la société et la place de l'Homme et cela me permet d'avoir une réflexion critique et éthique sur les enjeux de demain.

J'ai évidemment une pensée pour Clermont-Ferrand. C'est là que tout est né ! Merci donc aux anciens copains de la fac et particulièrement à Fred Beckouet et à Lydie Pradel. Je vous souhaite bonne chance dans vos attentes de la vie et vos post-docs respectifs.

Merci à la Savoie et ses montagnes où je retournerai régulièrement me ressourcer dans ces paysages grandioses. Merci à Chambéry et à mes anciens camarades du Lycée.

C'est l'Ifremer et le réseau d'excellence européen Marine Genomics qui ont financé cette thèse. Je remercie Catherine Boyen et Michèle Barbier, ainsi que Marine Genomics pour m'avoir donné l'opportunité de présenter mes travaux en Allemagne, en Italie et en Crète et réaliser l'étude sur l'ADN polymérase de virus en Islande. Je remercie ici toutes les personnes avec lesquelles j'ai interagi lors de ces colloques.

C'est Christine Lamy qui a contribué à la bonne gestion administrative des travaux réalisés au cours du doctorat. C'est son rire spontané qui a égayé le couloir et les goûters à côté de l'imprimante.

Je souhaite naturellement remercier les chercheurs qui ont accepté de participer à mon jury de thèse : Christiane Elie, Jean Armengaud, Charles Pineau, Marie-Thérèse Thebault et Mohamed Jebbar. Avant la rédaction de ce manuscrit, ce travail a été évalué par mon comité de thèse, Michel Gurvan et Franck Zal que je remercie. C'est avec l'équipe de Hannu Myllykallio que nous avons interrogé et c'est la plateforme de protéomique Ouest Genopole qui a permis l'identification des nouveaux complexes.

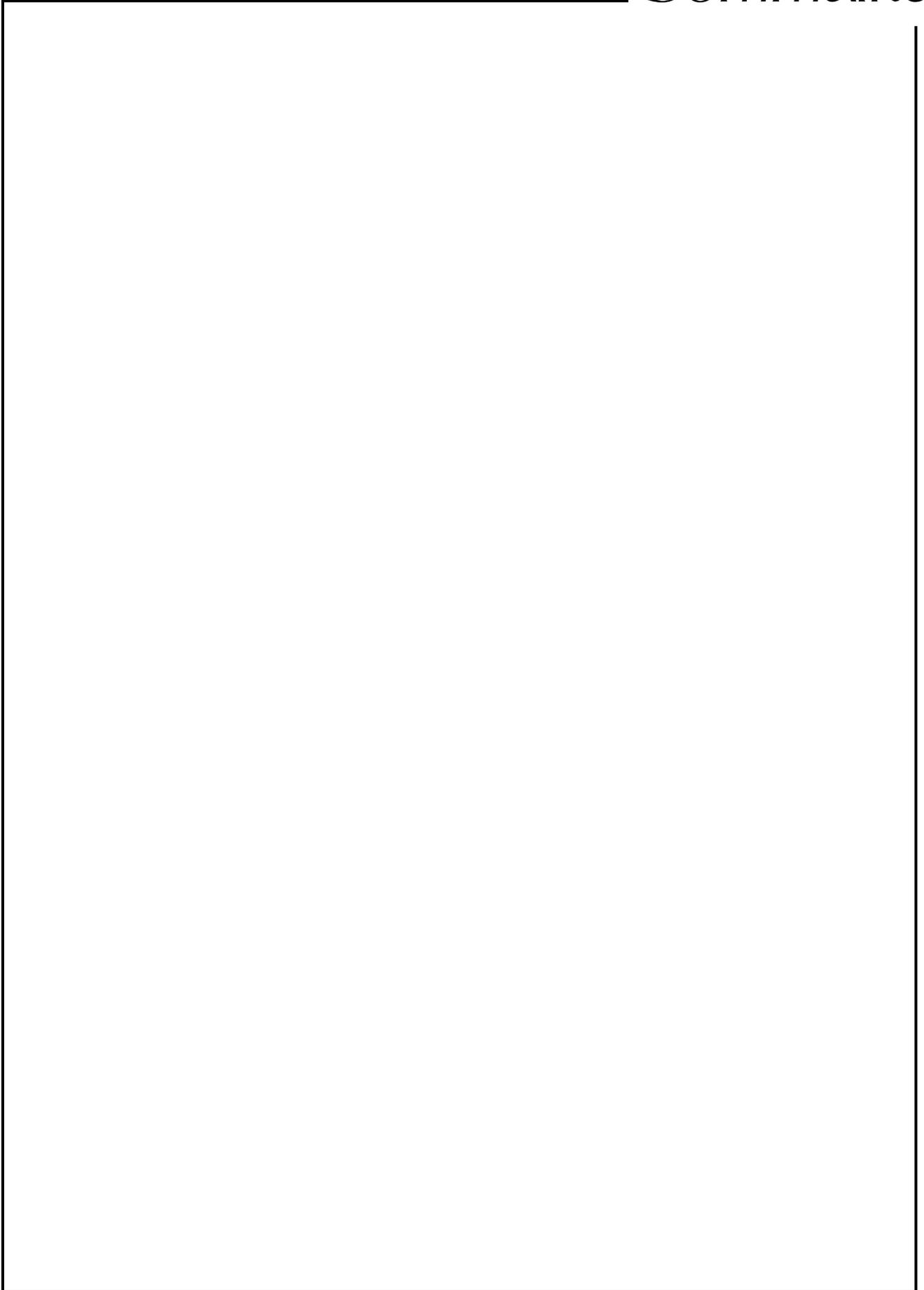
Je voudrais enfin dire toute mon affection à ma famille qui m'a soutenue depuis le début dans mon parcours scientifique et parfois chaotique. Je remercie mes parents, Isabelle et Jean-François, mon frère Aurélien (même s'il n'a pas pris conscience que le cidre de Normandie ne valait pas le cidre de Bretagne) et ma sœur Mélanie (Mel In Taiwan 祝你成功.).

C'est Didier Flament qui a tout relit.



Deux partenaires sur une plage islandaise, métaphore de l'interaction entre protéines.

Sommaire



Sommaire

Abréviations	6
INTRODUCTION	9
Introduction générale	10
I. Jusqu'où s'étend la biosphère ? Les écosystèmes hydrothermaux profonds	13
II. Les Archaea : troisième domaine du vivant	15
2.1. Découverte et caractéristiques des <i>Archaea</i>	15
2.2. Les principaux groupes d' <i>Archaea</i>	18
2.3. <i>Archaea</i> et origine de la vie.....	19
2.4. Ecologie des <i>Archaea</i>	21
2.5. Adaptation à la thermophilie	23
2.6. Potentialités biotechnologiques des Archées	25
2.7. <i>Pyrococcus abyssi</i>	27
III. Maintenance génomique chez les Archaea	29
3.1. La réplication de l'ADN.....	29
3.1.1. L'initiation de la réplication	31
3.1.2. Le replisome : l'élongation de l'ADN.....	32
3.1.3. Les protéines de liaison à l'ADN simple brin	33
3.1.4. L'ADN Primase.....	33
3.1.5. Les ADN Polymérase	35
3.1.5.1. Topologie générale des ADN polymérase	35
3.1.5.2. Différentes familles d'ADN polymérase	36
3.1.5.1. Répartition dans les trois domaines du vivant.....	36
3.1.6. Les facteurs accessoires	37
Le PCNA et le RFC : « the loader of the rings »	37
Le RF-C.....	38
Le PCNA et ses partenaires : la communauté de l'anneau.....	39
3.1.7. Les protéines interagissant avec le PCNA et intervenant dans la maturation des fragments d'Okazaki.	41
3.1.8. Autres protéines : les GINS.....	42
3.2. Processus de réparation de l'ADN chez les <i>Archaea</i>	48
3.2.1. La réparation par excision de nucléotide (<i>nucleotide excision repair</i> <i>NER</i>).....	49
3.2.2. La réparation par excision de bases (<i>Base excision repair</i> <i>BER</i>).....	50
3.2.3. La réparation des mésappariements (<i>Mismatch Repair</i> <i>MMR</i>).....	51
3.2.4. Réparation des cassures double brin (<i>Double strand break repair</i> <i>DSBR</i>).....	52
IV. Dis-moi avec qui tu interagis, je te dirais qui tu es...	55
4.1. Interactome : De l'interaction aux réseaux	55

4.2. Réseau d'interaction au niveau de la réplication et la maintenance de l'ADN.....	56
IV. Présentation de l'étude	62
RESULTATS ET DISCUSSION	65
I. Le PCNA acteur central de la maintenance génomique	67
1.1. Fen1, RNase HII et l'ADN Ligase 1 interagissent au niveau d'un site unique de fixation sur le PCNA	70
Article 1	71
1.2. <i>Pab2263</i> est un nouveau partenaire du PCNA.....	83
1.3. Le complexe Mre11 et Rad50 est physiquement associé au PCNA	86
1.3.1. Analyse des séquences primaires de Mre11 et Rad50	88
1.3.2. Expression et purification des protéines <i>PabMre11</i> et <i>PabRad50</i>	89
1.3.3. Validation de l'interaction <i>in vitro</i> entre les protéines recombinantes PCNA et Mre11/Rad50.....	90
1.3.4. Recherche de partenaires du complexe Mre11/Rad50.....	91
1.3.4.1. Le PCNA est associé au complexe Mre11/Rad50	93
1.3.4.2. Les protéines NurA et HerA, partenaires du complexe Mre11/Rad50	93
1.3.4.3. La protéine <i>Pab1633</i> de la famille PilT	93
1.3.4.4. La protéine <i>Pab1110</i>	94
II. Les protéines Fen1 et l'ADN primase interviennent dans des mécanismes de réparation de l'ADN.....	96
2.1. Association physique entre <i>PabFen1</i> et l'ADN primase de l'Archaea hyperthermophile <i>Pyrococcus abyssi</i> – implication pour des mécanismes de réparation.....	97
Article 2	98
2.2. Résultat complémentaire de l'étude : recherche de partenaires de p41	121
2.2.1. Les protéines RPA, partenaires de l'ADN primase.....	122
2.2.2. La reverse gyrase est-elle un nouveau partenaire de l'ADN primase ?.....	122
2.2.3. Les ARN polymérases, associées au RPA ou à l'ADN primase ?.....	123
2.2.4. La protéine <i>Pab1478</i>	124
2.3. Fen1 est associé à la topoisomérase VI.....	124
III. Analyse du protéome partiel.....	127
IV. Etude d'une ADN polymérase issue de virus thermophile	129
4.1. Introduction	129
4.2. Clonage des deux sous-unités de l'ADN polymérase indépendamment.....	129
4.3. Clonage des deux sous-unités de l'ADN polymérase dans un même vecteur et sur-expression.....	131
CONCLUSION ET PERSPECTIVES	135

MATERIELS ET METHODES	143
I. Techniques de Microbiologie	145
1.1. Culture de <i>Pyrococcus abyssi</i>	145
1.1.1. Culture de <i>P. abyssi</i> en batch	145
1.1.2. Culture de <i>P. abyssi</i> en bioréacteur	145
1.2. Culture d'Escherichia coli	147
II. Techniques de Biologie Moléculaire	148
2.1. Extraction de l'ADN génomique de <i>P. abyssi</i>	148
2.2. Amplification par PCR	148
2.3. Séparation et purification des fragments d'ADN	149
2.4. Techniques de Clonage	150
2.4.1. Vecteur	150
2.4.2. Clonage de deux gènes dans un même vecteur	152
2.4.3. Traitement des inserts par l'ADN polymérase T4	152
2.4.4. Réaction d' <i>annealing</i> du vecteur, de l'adaptateur et des inserts	152
2.4.5. Purification de l'ADN plasmidique	153
2.4.6. Séquençage de l'ADN	153
2.4.7. Transformation des cellules compétentes	153
2.4.7.1. Transformation en cellules de stockage	153
2.4.7.2. Transformation en cellules d'expression	154
III. Technique de Biochimie	155
3.1. Technique de production des protéines recombinantes	155
3.1.1. Expression des protéines en petit volume	155
3.1.2. Expression des protéines en grand volume	156
3.2. Purification des protéines	156
3.2.1. Précipitation des protéines non thermostables	156
3.2.2. Purification par chromatographie	156
3.2.3. Dialyse	157
3.3. Dosage des protéines	158
3.4. Séparation et visualisation des protéines par électrophorèse sur gel de polyacrylamide	159
3.4.1. Préparation des échantillons	159
3.4.2. Migration	160
3.4.3. Coloration et séchage des gels	160
3.5. Technique de Western Blot	160
3.5.1. Préparation des anticorps	160
3.5.2. Electro-transfert des protéines sur membrane	161
3.5.3. Immunodétection des protéines	161
3.6. Préparation de l'extrait cellulaire total de <i>Pyrococcus abyssi</i>	162
3.7. Capture de partenaires par <i>pull-down</i>	162
3.8. Capture de partenaires par Co-immunoprécipitation	163
3.9. Identification des protéines en spectrométrie de masse	164
3.10. Analyse du protéome partiel de <i>P. abyssi</i>	166

3.11. Etude des interactions protéine-protéine par résonance plasmonique de surface (SPR)	166
3.11.1. Principe de la résonance plasmonique de surface	166
3.11.2. Principe de fonctionnement de l'appareil utilisé	167
3.11.3. Expériences réalisées au cours de cette étude	168
3.12. Préparation des substrats	169
3.12.1. Marquage des oligonucléotides en 5'	169
3.12.2. Préparation de la matrice ADN amorcé	170
3.13. Protocole d'élongation de la matrice amorcée	171
3.14. Révélation des produits de réaction obtenus en électrophorèse sur gel de polyacrylamide dénaturant	171
2.15. Test d'activité de l'ADN polymérase	172
REFERENCES BIBLIOGRAPHIQUES	175
ANNEXES	195

Abréviations

ADN : Acide DésoxyRiboNucléique

ARN : Acide Ribonucléique

ARNr : ARN ribosomal

ATP : Adénosine Triphosphate

BER : Base Excision Repair

BET : bromure d'éthyidium

BLAST : Basic Local Alignment Search Tool

BSA : Bovine Serum Albumin

Co-IP : Co-Immunoprécipitation

Da : dalton

dNTP : desoxyribo Nucleotide Tri Phosphate

DSBR : Double Strand Break Repair

DTT : dithiothreitol

ECL : Enhanced Chemiluminescence

EDTA : acide

éthylènediaminotétraacétique sel disodique

Fen-1 : Flap Endonuclease

FPLC : Fast Protein Liquid Chromatography

HRP : horseradish peroxydase

HSP : heat shock proteins

IPTG : Isopropyl- β -D-galactoside

kpb : kilo paires de bases

LB : Luria-Bertani (milieu liquide ou gélosé)

LUCA : Last Universal Common Ancestor

MALDI-TOF : Matrix Assisted Laser Desorption Ionisation - Time of Flight

MCM : Mini Chromosome Maintenance

MMR : MisMatch Repair

MOPS : 3-(N-Morpholino)-propanesulfonic acid

MS : Spectrométrie de Masse

NER : Nucleic Excision Repair

ORC : Origin Recognition Complex

ORF : Open Reading Frame (cadre de lecture ouvert)

pb : paires de bases

PBS : Phosphate Buffered Saline

PCNA : Proliferating Cell Nuclear Antigen

PCR : Polymerase Chain Reaction

PIP : PCNA Interacting Protein

PVDF : PolyVinylidene DiFluoride

RF-C : Replication Factor C

RP-A : Replication Protein A

SDS-PAGE : Sodium Dodecyl Sulfate – PolyAcrylamide Gel Electrophoresis

SPR : Résonance Plasmonique de Surface

SSB : Single Strand Binding Protein

TAE : Tris/Ammonium/EDTA (tampon)

TBE : Tris/Borate/EDTA (tampon)

TBS : Tris Buffered Saline

TEMED : N,N,N',N' Tétréthylènediamine

UV : Ultraviolet

Introduction

Introduction générale

Qu'est-ce que la vie ?

Si chacun de nous à une idée intuitive de ce qu'est la vie, elle est en réalité difficile à définir et à expliquer. On peut consentir que la vie possède trois propriétés fondamentales qui sont l'autoconservation (la capacité de se maintenir en vie, le métabolisme), l'autoreproduction (se multiplier et transmettre ses propriétés à sa descendance) et l'autorégulation (le contrôle du métabolisme, la réaction aux stimuli). Les scientifiques et philosophes ont, depuis l'origine de l'Histoire, cherché à élucider la nature de la vie. Il est possible de distinguer le vivant et l'inerte (les objets inanimés), mais on sait aujourd'hui, avec l'exemple des virus, que la frontière entre ces deux domaines peut être difficile à placer. Les Hommes ont longtemps pensé que la matière vivante différait de la matière inerte. Le vitalisme présentait les organismes vivants doués d'une force vitale à l'origine de tous les mécanismes de la vie. Les découvertes du XX^eS ont depuis mis en évidence une unité de structure et de fonction du monde vivant : la cellule. On sait aujourd'hui grâce à la physique, à la chimie organique et à la biochimie que le vivant est composé de matière, c'est-à-dire d'atomes, de molécules, identiques à ceux composant l'inerte. François Jacob propose qu'« en réalité, la vie est un processus, une organisation de la matière »¹. La vie est donc une des structures de la matière. A la lueur des avancés scientifiques, Hubert Reeves définit la vie comme « cette tendance mystérieuse et universelle de la matière à s'associer, à s'organiser, à se complexifier »². Une autre propriété du vivant serait donc cette capacité à s'auto-organiser. C'est à partir de l'agencement et des interactions entre les molécules qu'émergent des propriétés nouvelles. « Mettre ensemble, ce n'est pas ajouter, c'est faire apparaître du nouveau »³.

Au sein des cellules, si les gènes permettent de stocker et de coder de l'information, ce sont les protéines qui assurent une fonction, ce sont des acteurs majeurs de la vie cellulaire. De part leurs structures tridimensionnelles, les protéines sont capables de catalyser des réactions et d'interagir avec d'autres protéines. La notion d'auto-organisation du vivant se retrouve ici, dans les relations et les interactions entre protéines, leur capacité à former des complexes multiprotéiques résultant de leur affinité et de leur complémentarité. Les interactions protéine-protéine peuvent être schématisées sous forme de réseau d'interaction. Du protéome (l'ensemble des protéines de la cellule, dans une condition donnée), nous passons à l'ère de l'interactome (l'ensemble des interactions cellulaires). C'est probablement dans la richesse de ces interactions que réside la complexité du fonctionnement des cellules vivantes.

¹ François Jacob, *Qu'est-ce que la vie ? in La Vie*, Université de tous les savoirs, Editions Odile Jacob, 2002.

² Hubert Reeves, *Poussières d'étoiles*, Edition du Seuil, 1984.

³ Albert Jacquard, *La légende de la Vie*, Flammarion, 1999.

Une des propriétés fondamentales du vivant est de pouvoir se reproduire et se multiplier. Cette propriété requiert avant tout la duplication du génome, afin de transmettre l'information essentielle à la survie de la nouvelle cellule formée. C'est la réplication de l'ADN qui assure le maintien et la transmission de l'information génétique ainsi que la propagation des êtres vivants. Un nombre important de travaux permettant de mieux comprendre le mécanisme moléculaire de la réplication de l'ADN ont été effectués dans les trois domaines du vivant : *Eukaryota*, *Bacteria* et *Archaea*. De nombreuses protéines impliquées dans ce processus interagissent de façon stable ou transitoire en formant un complexe appelé replisome. Le mécanisme de la réplication de l'ADN doit être assez robuste afin d'éviter les erreurs dans la transmission de l'information génétique. Le taux d'erreur dans la synthèse de l'ADN est d'environ 10^{-9} *in vivo*, mais malgré la fiabilité du processus, ce taux n'est pas négligeable. L'ADN subit de plus diverses attaques ou réactions chimiques pouvant entraîner des lésions. Il existe donc d'autres mécanismes, mettant en jeu un grand nombre de protéines et permettant d'assurer la maintenance de l'ADN, en participant à la détection et la réparation des erreurs et des lésions éventuelles introduites dans le génome.

Il est temps maintenant d'introduire notre modèle d'étude, *Pyrococcus abyssi*, un micro-organisme procaryote, une archée, rencontré dans un environnement qui peut nous paraître insolite, les sources hydrothermales sous-marines. Cet organisme est qualifié d'extrêmophile car il présente la capacité de se multiplier à des températures proches de 100°C. En plus de cette particularité, *Pyrococcus abyssi* présente comme toutes les archées un intérêt tout particulier au niveau évolutif : si elle ressemble en certain point aux procaryotes bactériens, l'analyse du génome montre que les gènes intervenant dans la maintenance et l'expression du génome sont homologues à ceux des eucaryotes. Le dessein de ce projet de thèse était donc de contribuer à la découverte de nouvelles protéines, de nouvelles interactions intervenant dans ces mécanismes de maintenance génomique que sont la réplication et la réparation de l'ADN. En effet, les travaux menés au sein de l'équipe Maintenance Génomique chez les *Archaea*, du laboratoire de Microbiologie des Environnements Extrêmes de l'Ifremer concernent l'étude des mécanismes de réplication et de réparation des *Archaea* hyperthermophiles. Un certain nombre de protéines intervenant dans la machinerie moléculaire de la réplication de l'ADN de *Pyrococcus abyssi* ont été caractérisées et sont utilisables en guise d'appâts pour la découverte de nouveaux partenaires.

Au cours de l'analyse bibliographique de ce manuscrit, après un bref descriptif de l'environnement extrême que sont les cheminées hydrothermales, nous décrirons plus en détail les caractéristiques des *Archaea*, le troisième domaine du vivant, puis *Pyrococcus abyssi* notre modèle d'étude. Enfin, les notions de réplication et de réparation de l'ADN seront expliquées, les protéines intervenant dans ces machineries moléculaires seront décrites, ceci afin d'avoir une vision globale de ces mécanismes au sein des trois domaines du vivant. La partie expérimentale détaillera ensuite l'approche utilisée dans ce travail, pour mettre à jour un réseau d'interactions protéiques qui a permis d'envisager de nouvelles perspectives

dans le domaine de la maintenance génomique de *P. abyssi*. En plus de l'interaction physique entre protéines, j'ai tenté d'explicitier la fonction de ces interactions. En confrontant mes travaux à d'autres données bibliographiques, je leur propose ici des rôles potentiels dans la maintenance du génome.

* *
*



Nuwa, personnage de la mythologie chinoise, déesse créatrice, forme un couple avec Fuxi. Les corps entrelacés rappellent la structure de l'ADN. Le réseau d'interaction protéique est schématisé en arrière plan.

I. Jusqu'où s'étend la biosphère ? Les écosystèmes hydrothermaux profonds

Les progrès technologiques au cours des années 1970 ont permis la mise au point de petits submersibles permettant l'exploration des fonds océaniques. En 1977, c'est grâce à l'un d'eux, l'Alvin, qu'une équipe de géologues, dirigée par John Corliss, observait des gisements minéraux sur la dorsale des Galápagos. Ils ne se doutaient pas qu'ils étaient sur le point de faire une découverte qui allait révolutionner la géologie et la biologie. Ils observèrent à 2500 mètres sous la surface, dans l'obscurité totale, de grandes cheminées crachant des fumées noires, avec aux alentours une faune luxuriante. Les sources hydrothermales venaient d'être découvertes (Corliss and Ballard, 1977). Les campagnes de plongées ultérieures ont permis la mise en évidence et l'exploration d'autres sites hydrothermaux au niveau des dorsales océaniques de l'Atlantique, du Pacifique et de l'Océan Indien.

Les phénomènes hydrothermaux sont une conséquence indirecte des phénomènes d'extension et d'accrétion des plaques tectoniques. La croûte océanique au niveau des dorsales océaniques, est parsemée d'anfractuosités où l'eau de mer peut s'infiltrer. C'est en traversant la croûte terrestre que l'eau peut se charger en divers composés solubles et en métaux (lithium, potassium, calcium, fer, cuivre, manganèse, zinc...). L'eau qui peut s'infiltrer jusqu'à plusieurs kilomètres de profondeur se réchauffe à proximité des chambres magmatiques présentes sous les dorsales. Sous l'effet de la pression, le fluide chaud peut ensuite remonter vers le plancher océanique, puis jaillit en des points localisés, constituant les sources hydrothermales. Les minéraux présents dans le fluide chaud précipitent ensuite à proximité de la zone d'émission pour former des édifices hydrothermaux (figure 1 et 2).

La présence d'une densité biologique très importante autour de ces sources hydrothermales, qualifiées d'oasis de vie, a intrigué les chercheurs pendant de nombreuses années. En effet, tous les écosystèmes sur Terre dépendent de la présence d'organismes autotrophes, c'est à dire capable de synthétiser de la matière organique à partir du dioxyde de carbone. Or, l'absence de lumière à cette profondeur rend la photosynthèse impossible. Avant 1977, l'environnement abyssal était d'ailleurs considéré comme dépourvu de vie car dénué de production primaire photosynthétique. Les travaux suivants ont démontré l'existence de bactéries autotrophes dans l'écosystème hydrothermal. Ces bactéries sont capables de transformer le CO₂ en molécules organiques en utilisant non pas l'énergie solaire, mais l'énergie dégagée par l'oxydation

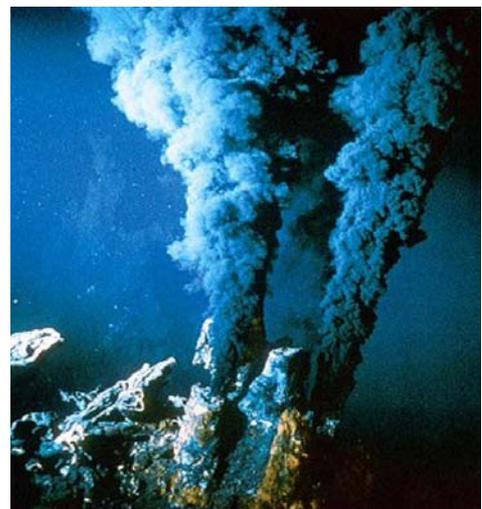


Figure 1 : fumeur noir (Ifremer)

de divers composés chimiques : c'est la chimiosynthèse. Ces micro-organismes sont donc à la base de la chaîne alimentaire de cet écosystème. Les animaux vivant aux alentours des sources hydrothermales peuvent se nourrir directement des bactéries alors que d'autres, comme les vestimentifères, vivent en association symbiotique avec ces bactéries.

La découverte étonnante des sources hydrothermales sous-marines a bouleversé nos préjugés en biologie, mais elle a aussi relancé le débat sur l'origine de la vie sur Terre. Elle nous a également permis de caractériser de nouveaux micro-organismes, qualifiés d'extrémophiles, capables de vivre et de se multiplier à des températures très élevées au plus proche des fluides chauds émergents des cheminées hydrothermales. L'étude de ces communautés microbiennes, leurs adaptations physiologiques aux conditions de vie très particulière de cet écosystème, leurs valorisations et leurs potentiels biotechnologiques sont des champs de recherche exaltant pour les scientifiques. Parmi ces micro-organismes, les thermophiles les plus extrêmes sont les *Archaea*, autrefois appelés archéobactéries. Nous commencerons donc ici par introduire ce domaine du vivant, fascinant à plus d'un titre. Après avoir décrit, l'archée *Pyrococcus abyssi*, notre modèle d'étude, nous recentrerons nos propos sur la maintenance génomique chez les archées : de la réplication à la réparation de l'ADN. Quelles sont leurs particularités, quelles sont leurs similarités avec les autres domaines du vivant ?

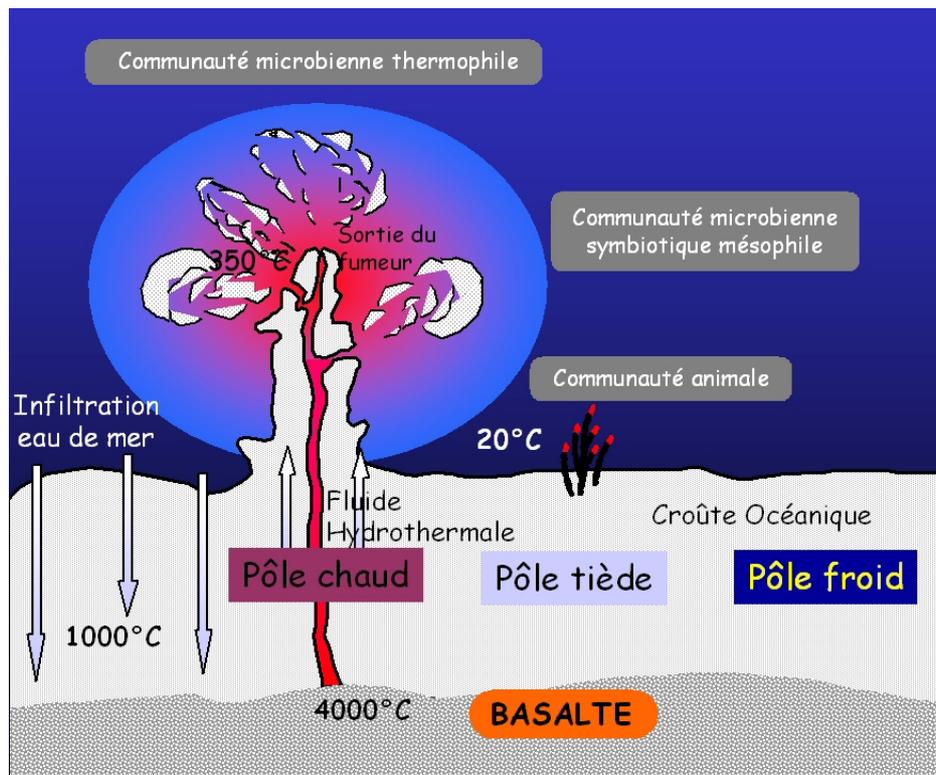


Figure 2 : Représentation schématique des sources hydrothermales océaniques

* *
*

II. Les *Archaea* : troisième domaine du vivant

2.1. Découverte et caractéristiques des *Archaea*

Depuis les fondements de la microbiologie, initiée par les travaux de Louis Pasteur et Robert Koch, la classification taxonomique des procaryotes était basée sur des caractères phénotypiques comme la morphologie des cellules, la physiologie et la pathogénicité. Cette classification permettait l'identification des micro-organismes. Cependant, cette classification ne tenait pas compte des relations évolutives entre les différents groupes taxonomiques. Pour remédier au problème de taxonomie chez les procaryotes, Carl Woese choisit la séquence de la petite sous-unité de l'ARN ribosomique (ARNr) comme horloge moléculaire universelle (Woese and Fox, 1977). Les travaux de Woese en phylogénie moléculaire l'ont donc conduit à diviser le vivant en trois grands domaines : les *Eukarya*, les *Bacteria* et les *Archaea* (Woese *et al.*, 1990) (figure 3). Cette découverte a permis d'initier de nouvelles hypothèses sur la nature de LUCA (*Last Universal Cellular Ancestor*) et sur les relations phylogénétiques entre les trois domaines du vivant (Woese, 2000) (Forterre *et al.*, 2002) (Cavalier-Smith, 2006).

Les procaryotes sont apparus sur Terre il y a 3,5 à 4 milliards d'années et ont constamment évolué depuis. En raison de cette longue période d'évolution, il existe une grande diversité de procaryotes, en particulier au niveau du métabolisme et de la physiologie. La diversité métabolique des Eucaryotes est au contraire plus limitée. Les domaines procaryotes *Archaea* et *Bacteria* possèdent des caractéristiques communes mais aussi des différences très spécifiques (tableau 1).

Les *Archaea* peuvent se distinguer des deux autres domaines par de nombreux points (tableau 1), notamment par la structure et la chimie de la membrane cellulaire. Les lipides membranaires des *Archaea* sont particuliers et consistent en des éthers de glycérol et d'isoprénoïdes. La paroi constituée de peptidoglycane chez les bactéries (hormis des Mycoplasmes) est absente chez les *Archaea*. A la place de la paroi, on peut trouver une couche régulière, constituée de sous-unités protéiques, appelée couche de surface (ou S-layer). Le chromosome circulaire est de type bactérien, mais il comporte des gènes mosaïques similaires à ceux des eucaryotes. De plus, le chromosome est associé à des protéines homologues aux histones eucaryotes. Les ARN polymérases ADN-dépendantes des *Archaea*, copiant l'ADN pour donner l'ARN messager, sont plus complexes que celles des bactéries, et similaires à celles des eucaryotes. Les protéines impliquées dans la réplication de l'ADN sont également plus proches de celles des eucaryotes que de celles des bactéries (Kelman, 2000).

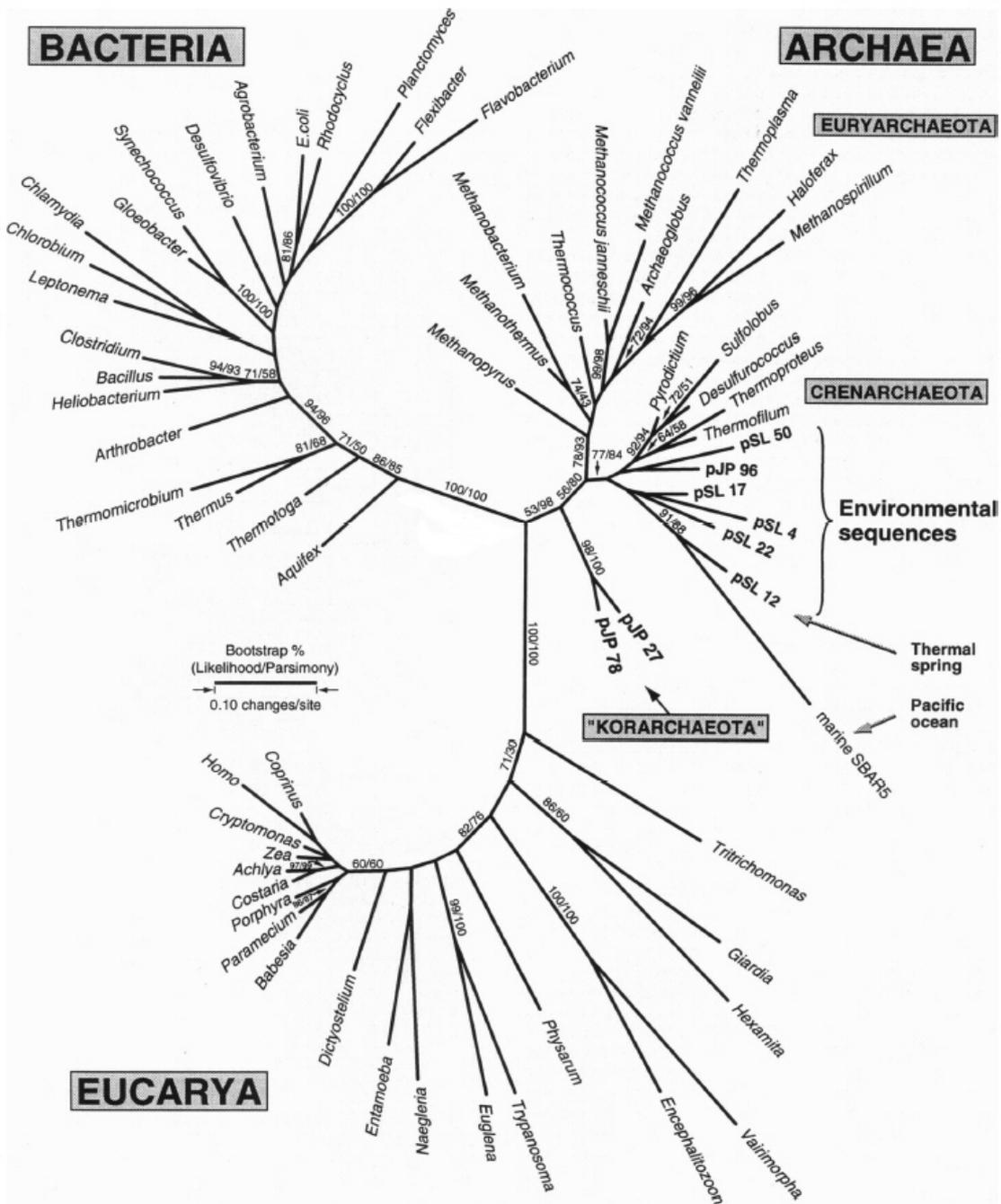


Figure 3 : Arbre phylogénétique non enraciné, issu d’une analyse comparative de séquences de gènes ribosomiques et montrant la séparation des trois grands domaines du vivant : Archaea, Bacteria et Eucaryote. D’après (Barns *et al.*, 1996).

	<i>Archaea</i>	<i>Bacteria</i>	Eucaryotes
Taille typique	1 à 4 µm	1 à 4 µm	Supérieure à 5 µm
Caractéristiques physiologiques	Aérobie et anaérobie	Aérobie et anaérobie, chlorophyllienne	Très largement aérobie, chlorophyllienne
Matériel génétique	Petit chromosome circulaire, plasmides et virus ; génome associé avec des histones	Petit chromosome circulaire, plasmides et virus ; pas d'histones	Noyau complexe avec plus d'un chromosome linéaire, virus ; génome associé avec des histones
Différenciation	Généralement unicellulaire, différenciation cellulaire rare	Généralement unicellulaire, différenciation cellulaire rare	Unicellulaire et pluricellulaire, différenciation cellulaire fréquente
Paroi cellulaire	Protéine, glycoprotéine, pseudomuréine, pas de paroi	Peptidoglycane (muréine) et LPS, rarement formée de protéines, formes sans parois rares	Grande variété, absence de peptidoglycane
Membrane cytoplasmique	Ether de glycérol et d'isoprénoïdes, site de la biosynthèse d'énergie	Ester du glycérol et des acides gras, site de la biosynthèse de l'énergie	Esters du glycérol et des acides gras, cholestérol fréquent
Membranes intracytoplasmiques	Généralement absentes	Généralement absentes ; quand présentes, contiennent de grandes quantités de protéines	Fréquentes dans les organites comme les mitochondries et les chloroplastes, le noyau, l'appareil de Golgi, le réticulum endoplasmique et les vacuoles ; site de la biosynthèse de l'énergie
Locomotion	Flagelle simple	Flagelle simple, glissement	Flagelles complexes, cils, pattes, nageoires, ailes
Traduction			
Synthèse protéique	Ribosome 70S, insensible au chloramphénicol et au cycloheximide, diphtamide présente dans le facteur d'élongation	Ribosome 70S, sensible au chloramphénicol, insensible au cycloheximide, diphtamide absente dans le facteur d'élongation	Ribosome 80S et 70S (organites), insensible au chloramphénicol (80S), sensible au cycloheximide (80S), diphtamide présente dans le facteur d'élongation
ARNr	5S, 16S, 23S	5S, 16S, 23S	5 S, 8S, 18S, 28S
Séquence Shine Dalgarno	Présence	Présence	Absence
Transcription			
ARN polymérase	Complexe (1 type, 12 sous-unités)	Simple (1 type, 4 sous-unités)	Complexe (3 types, nombreuses sous-unités)
Mode de fixation de l'ARN Pol sur le promoteur	Protéine TBP (TATA binding protein) Promoteur : séquence TATA située à -25 pb du site d'initiation	Facteur σ Promoteur : séquences situées à -35 pb et -10 pb du site d'initiation	Protéine TBP Promoteur : séquence TATA située à -25 pb du site d'initiation
Réplication de l'ADN			
Origine de réplication	Une ou plusieurs origines de réplication selon les espèces	Une seule origine de réplication <i>Ori C</i>	Plusieurs origines de réplication impliquant le complexe ORC
Hélicase	MCM (Minichromosome maintenance complexe)	DnaB	MCM (Minichromosome maintenance complexe)
ADN Polymérases	Famille B et D pour les <i>Euryarchaeota</i> , Famille B pour les <i>Crenarchaeota</i>	Pol I (famille A), Pol II (famille B), Pol III (famille C)	Principales ADN Pol : Pol α , Pol δ , Pol ϵ (famille B)
Facteur de processivité	PCNA	β clamp	PCNA

Tableau 1 : Comparaison des principales caractéristiques des trois domaines du vivant, d'après (Perry *et al.*, 2004) modifié.

2.2. Les principaux groupes d'*Archaea*

Les *Archaea* se composent de trois groupes phénotypiques principaux : les *Archaea* productrices de méthane (les méthanoarchéobactéries), les halophiles extrêmes (les haloarchéobactéries) et les thermophiles extrêmes.

La classification phylogénétique est plus complexe, et les *Archaea* cultivables ont été divisées en deux sous-domaines (ou phylum), sur la base de la structure de leur ARNr : les *Euryarchaeota* (Woese *et al.*, 1990) et les *Crenarchaeota* (Burggraf *et al.*, 1997). Le phylum *Crenarchaeota* est constitué essentiellement de thermophiles extrêmes, mais aussi d'espèces mésophiles et psychrophiles. Le phylum *Euryarchaeota* comporte tous les méthanogènes et les halophiles extrêmes, mais aussi certains thermophiles extrêmes.

Un grand nombre d'*Archaea* ne sont pas cultivables et parmi celles-ci, une lignée qui ne se regrouperait pas avec les deux phylums principaux a été classée dans un troisième phylum : les *Korarchaeota* (Barns *et al.*, 1996). Cependant, ce groupe pourrait être artefactuel et serait composé d'organismes présentant des gènes ARNr16S subissant des mutations rapides (Forterre *et al.*, 2002).

Plus récemment, une nouvelle espèce, *Nanoarchaeum equitans*, a été décrite (Huber *et al.*, 2002). Cette *Archaea* présente des caractéristiques étonnantes : elle serait la plus petite cellule vivante connue à ce jour et possède également le plus petit génome décrit (490 kb et environ 400 gènes) (Huber *et al.*, 2003). *Nanoarchaeum equitans* vit comme symbiote, fixée à la surface de la crénarchée hyperthermophile *Ignicoccus* sp. Elle peut être produite en co-culture avec *Ignicoccus* sp. Le génome de *N. equitans* est entièrement séquencé et annoté. Les gènes impliqués dans des mécanismes informationnels (réplication, réparation, cycle cellulaire) sont présents, mais les gènes impliqués dans certains mécanismes métaboliques comme la biosynthèse des acides aminés, des nucléotides et des lipides sont absents. Étant donné la petite taille du génome, cette absence n'est pas surprenante. Beaucoup de métabolites requis pour *N. equitans* doivent être apportés par la cellule hôte *Ignicoccus* sp., ce phénomène indiquant que *N. equitans* entretiendrait une relation de parasitisme avec *Ignicoccus* sp. (Waters *et al.*, 2003). Au niveau phylogénétique, cette nouvelle espèce représenterait un phylum très ancien : les Nanoarchaeota (Huber *et al.*, 2002).

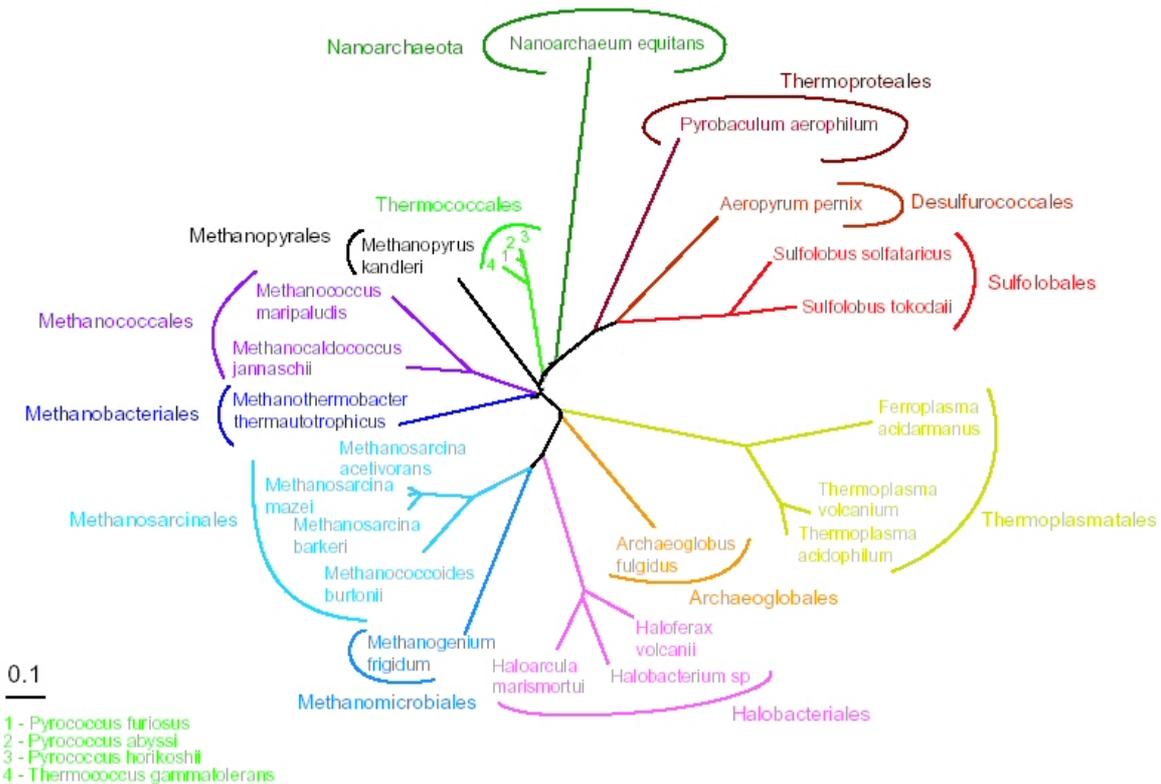


Figure 4 : Arbre phylogénétique non enraciné représentant les grands groupes d'*Archaea* d'après (Brochier *et al.*, 2005).

2.3. *Archaea* et origine de la vie

La phylogénie et les analyses des génomes disponibles des *Archaea* cultivables suggèrent que l'ancêtre de toutes les Archées actuelles était un organisme anaérobie qui vivait probablement lorsque l'oxygène était encore absent de l'atmosphère terrestre. (Forterre *et al.*, 2002). Cet ancêtre était plus probablement un chimiolitho-autotrophe, étant donné que les chimiolitho-autotrophes sont présents chez tous les phylums d'archées. Cette hypothèse est de plus compatible avec la théorie de l'origine autotrophique de la vie (Wächtershäuser, 1992). Étant donné que les hétérotrophes hyperthermophiles sont également présents dans les deux principaux phylums d'Archées, on ne peut pas exclure que l'ancêtre des *Archaea* était un hétérotrophe. L'existence de transfert latéral de gène chez les procaryotes complexifie encore toute hypothèse sur l'origine et l'évolution des profils métaboliques chez ces organismes (Jain *et al.*, 1999) (Boucher *et al.*, 2003).

La réplication de l'ADN est un processus essentiel et caractéristique de toutes les formes de vies cellulaires. La recherche de l'origine de l'ADN et des protéines participant à la réplication de l'ADN est un sujet à la fois complexe et captivant, amenant les chercheurs à formuler hypothèses et concepts en perpétuelle évolution. LUCA, le dernier ancêtre commun de tous les êtres vivants actuels se positionnerait, au sein d'un arbre phylogénétique, au point de divergence entre les trois règnes du monde vivant. Si la plupart des chercheurs admettent

l'existence d'un monde à ARN à l'origine de la vie, à ce jour la nature du génome de LUCA (ARN ou ADN) est encore soumise à débat. Après l'invention de l'ADN, une séparation claire s'établit entre les processus de répllication du matériel informatif, et les processus de synthèse des protéines. L'arbre universel de la vie, enraciné, proposé par Carl Woese en 1990 et celui de Karl Stetter laissait suggérer que les plus proches parents de LUCA étaient des hyperthermophiles (Bactéries et Archées). A l'époque, l'hypothèse d'un LUCA hyperthermophile, vivant dans les conditions extrêmes de la Terre primitive, a été adoptée par une partie de la communauté scientifique. Di Giulio propose que LUCA, tout comme l'ancêtre des Bactéries et des Archées était un hyperthermophile, la vie serait donc apparue dans un environnement avec une température élevée (Di Giulio, 2003). Cette hypothèse est toutefois soumise à débat. En effet, Patrick Forterre a mis en évidence une protéine spécifique des hyperthermophiles, c'est à dire absente des génomes des organismes mésophiles et modérément thermophiles : la reverse gyrase (Forterre, 2002b). En analysant la phylogénie et le contexte génomique du gène de la reverse gyrase, il a mis en évidence que cette enzyme avait été transférée des Archées aux Bactéries par transfert latéral de gène, suggérant ainsi que l'ancêtre commun des bactéries n'était pas un hyperthermophile (Forterre *et al.*, 2000). Par contre, Forterre et Brochier suggèrent que la reverse gyrase était présente chez l'ancêtre commun des *Archaea* et proposent que cette lignée serait née dans des environnements avec une température élevée.

La machinerie moléculaire de la répllication présente des similarités fonctionnelles chez les trois domaines du vivant, ce qui laisserait suggérer une origine commune de ce mécanisme. Cependant, un travail cherchant à déterminer quelles étaient les protéines ubiquitaires dans les trois domaines du vivant a permis de mettre en évidence entre 60 et 80 protéines universelles (Koonin, 2003) et étonnement, les protéines impliquées dans le métabolisme de l'ADN sont très peu représentées. Les trois protéines essentielles à la fourche de répllication (ADN polymérase, primase et hélicase) des Archées sont homologues avec celles des Eucaryotes mais pas avec celles des Bactéries. Suite à ces observations, Mushegian et Koonin ont proposé que la répllication de l'ADN ait été inventée deux fois (Mushegian and Koonin, 1996), une fois dans la branche des Bactéries, une fois dans la branche commune des Archées et des Eucaryotes. Leipe et Koonin ont également proposé que LUCA contenait à la fois de l'ADN et de l'ARN, et que c'est le système de répllication de l'ADN qui a évolué indépendamment dans la lignée des Bactéries d'une part, et dans la lignée des Archées et des Eucaryotes d'autre part (Leipe *et al.*, 1999). Par ailleurs, en s'intéressant aux protéines de la répllication chez les virus, Forterre propose une autre théorie. L'ADN serait apparu d'abord chez les virus, puis aurait été transféré aux cellules au cours de l'évolution en même temps que les trois protéines essentielles à la répllication. Deux transferts indépendants auraient eu lieu, l'un dans la branche des ancêtres des Bactéries, l'autre dans la branche commune des ancêtres des Archées et des Eucaryotes (Forterre, 2002a).

2.4. Ecologie des *Archaea*

Les *Archaea* sont très différentes des autres organismes, mais elles présentent également une grande diversité entre elles. On retrouve chez les archées les mêmes types morphologiques existant chez les bactéries (bacilles, coques, présence ou absence de flagelles, agrégation en amas) mais aussi des morphologies inconnues chez les bactéries comme les formes polygonales des archées halophiles. Il existe également de nombreux groupes métaboliques, des chimiolithoautotrophes aux chimiohétérotrophes. Les *Archaea* peuvent être aérobies, anaérobies facultatives ou anaérobies strictes. Dans le milieu naturel, les *Archaea* peuvent être isolées à partir d'une grande diversité d'habitats.

Les *Archaea* méthanogènes tel que *Methanobacterium*, *Methanosarcina* sont rencontrées dans des environnements anaérobies contenant des quantités élevées de matière organique facilement dégradable comme les sédiments des lacs d'eau douce, les marécages et les rizières. Le méthane peut aussi être formé par des *Archaea* méthanogènes de la flore anaérobie du gros intestin des animaux. *Methanobrevibacter ruminantium* est le méthanogène prédominant du rumen des bovins, mais plusieurs autres espèces ont été identifiées (Whitford *et al.*, 2001). L'accroissement de la quantité de méthane présent dans l'atmosphère a conduit les scientifiques à rechercher ses origines et il a été montré que les méthanogènes colonisant les ruminants contribuent de façon significative à la présence de méthane atmosphérique (Johnson and Johnson, 1995). 10 à 20% du méthane émis dans l'atmosphère proviendrait du rumen des herbivores. *Methanobrevibacter smithii* est le méthanogène le plus abondant de l'intestin humain (Miller *et al.*, 1982). Des méthanogènes sont également présents au niveau de la plaque dentaire (Belay *et al.*, 1988) et dans le vagin (Belay *et al.*, 1990). Beaucoup de protozoaires anaérobies du rumen, de l'eau douce et des marécages, et même des ciliés marins sont associés à des symbiotes méthanogènes. Les méthanogènes sont les *Archaea* les plus fréquemment rencontrés en association avec les animaux (Protozoaires, Invertébrés, Arthropodes, Ruminants, Oiseaux, l'Homme) (Lange *et al.*, 2005) mais notons toutefois que des études ont proposés la présence d'autres *Archaea* chez quelques animaux. Des créonarchées pourraient être présentes dans l'intestin humain (Rieu-Lesme *et al.*, 2005).

Il est surprenant de noter que malgré l'association ubiquitaire des *Archaea* avec les animaux, aucune Archée pathogène n'ait été formellement identifiée. Le débat est cependant ouvert et des recherches sur la pathogénicité éventuelle des archées sont actuellement en cours (Cavicchioli *et al.*, 2003; Eckburg *et al.*, 2003). Des études proposent que des archées puissent être impliquées, de façon indirecte, chez l'Homme, dans des lésions de tissus de patients atteints de parodontite (Lepp *et al.*, 2004).

Des études récentes, basées sur l'amplification de l'ARNr 16S par PCR sur des échantillons environnementaux du sol et de l'eau, ont montré que les archées mésophiles étaient plus largement répandues que ce que l'on pensait auparavant (Robertson *et al.*, 2005). A titre d'exemple, des études ont montrées que des archées mésophiles du phylum Crenarchaeota représenteraient près de 20% de la biomasse totale du picoplancton océanique

(Karner *et al.*, 2001) et pourraient contribuer de façon significative au cycle du carbone océanique (Herndl *et al.*, 2005).

Les *Archaea* halophiles extrêmes ne peuvent se multiplier qu'à des concentrations en NaCl supérieures à 1,8 M et beaucoup d'espèces sont capables de croître à des concentrations supérieures à 4 M. Ces micro-organismes sont fréquemment rencontrés dans des environnements très salés comme le littoral marin, les marais salants, la Mer Morte, des lacs salés (Grand Lac Salé de l'Utah, Lac Magadi au Kenya) (Oren, 1994). Les Halobacteriales (*Halobacterium*, *Halococcus*, *Halorubrum*, *Natrinema*, *Natronococcus*) sont caractéristiques de ce type d'environnement. Certaines espèces possèdent des caroténoïdes et présentent une pigmentation rouge à jaune qui est responsable de la coloration de certains lacs (Lac Magadi au Kenya par exemple) et parfois de la couleur rouge de la saumure des salines lors de la préparation du sel marin.

Règne	Ordre	Genres	Habitat	Morphologie	Température (°C)		pH optimum
					Optimum	Maximum	um
Crenarchaeota							
Igneococcales							
		<i>Pyrodictium</i>	Cheminées hydrothermales profondes et cotières	Disque	105	110	6
		<i>Pyrolobus</i>	Cheminées hydrothermales profondes	Coque	106	113	5,5
		<i>Desulfurococcus</i>	Sources chaudes et solfatares	Sphère	85	95	6
		<i>Staphylothermus</i>	Bouche hydrothermale sous marine	Sphère	92	98	6-7
Thermoproteales							
		<i>Thermoproteus</i>	Sources chaudes et solfatares	Bâtonnet	88	96	6
		<i>Thermofilum</i>	Solfatares	Bâtonnet	88	95	5,5
		<i>Pyrobaculum</i>	Cheminées hydrothermales profondes et solfatares	Bâtonnet	95	100	6
Sulfolobales							
		<i>Sulfolobus</i>	Solfatares acides	Coque irrégulier	75-85	87	1-5
		<i>Acidianus</i>	Solfatares acides	Coque irrégulier	85-90	95	2
		<i>Desulfurolobus</i>	Solfatares acides	Coque irrégulier	80	87	2,5
		<i>Stygiolobus</i>	Sources chaudes	Coque irrégulier	80	89	3
Euryarchaeota							
Thermococcales							
		<i>Pyrococcus</i>	Cheminées hydrothermales profondes et cotières	Coque	100	106	6-8
		<i>Thermococcus</i>	Cheminées hydrothermales profondes et cotières	Coque	88	98	6-7
Archaeoglobales							
		<i>Archaeoglobus</i>	Cheminées hydrothermales profondes et cotières	Coque	83	95	7
		<i>Ferroglobus</i>	Bouche hydrothermale sous marine	Coque irrégulier	85	95	7
Thermoplasmatales							
		<i>Thermoplasma</i>	Solfatares, tas de charbon	Coque, filament	59	63	2
		<i>Picrophilus</i>	Solfatares acides	Coque	60	65	0,7

Tableau 2 : Habitats et caractéristiques générales des *Archaea* hyperthermophiles. D'après (Vetriani, 2001).

Les *Archaea* hyperthermophiles sont rencontrées dans des environnements où l'énergie géothermique est disponible comme les sources chaudes, les solfatares, les geysers,

les sédiments marins chauffés par géothermie et les cheminées hydrothermales sous-marines. Par exemple, les espèces des genres *Pyrococcus*, *Methanopyrus*, *Thermococcus*, *Sulfolobus*, *Pyrodictium*, *Archaeoglobus* sont rencontrés dans ce type d'environnement (Tableau 2). Ces espèces possèdent des températures optimales de croissance supérieures à 80°C. *Pyrolobus fumarii* est à ce jour l'hyperthermophile la plus extrême, avec un optimum de croissance de 106°C, elle est encore capable de se multiplier jusqu'à 113 °C (Blochl *et al.*, 1997).

Les organismes qui se développent dans des environnements extrêmes sont qualifiés d'extrémophiles. Ce qualificatif englobe les résistances à des conditions physiques (par exemple, la température, la pression, des radiations) et des résistances à des conditions géochimiques (par exemple la dessiccation, la salinité, le pH, le potentiel redox). L'étude des organismes extrémophiles est un champ de recherche passionnant à divers titres : l'évolution, les stratégies d'adaptation et le potentiel industriel et biotechnologique des enzymes de l'extrême.

2.5. Adaptation à la thermophilie

Les organismes thermophiles peuvent vivre et se multiplier à des températures comprises entre 50 et 70°C, alors que les organismes hyperthermophiles présentent des optimum de croissance à des températures comprises entre 80 et 110°C. Les hyperthermophiles les plus extrêmes sont des *Archaea* (figure 5). Afin de vivre et croître à des températures si élevées, les composants cellulaires de ces micro-organismes, comme les protéines, les acides nucléiques et les membranes cellulaires, doivent être adaptées et stables aux températures proche de 100°C (Stetter, 1999). Les bases biochimiques de cette thermoadaptation ne sont encore que partiellement connues.

Les lipides membranaires doivent diminuer la fluidité membranaire afin d'éviter la fonte de la bicouche lipidique par la chaleur. Les archées présentent par exemple des liaisons éther dans les lipides (contrairement aux liaisons ester des bactéries mésophiles), ce qui leur confère une résistance remarquable contre les hydrolyses dues aux hautes températures, mais aussi contre les pH acides (Stetter, 1999).

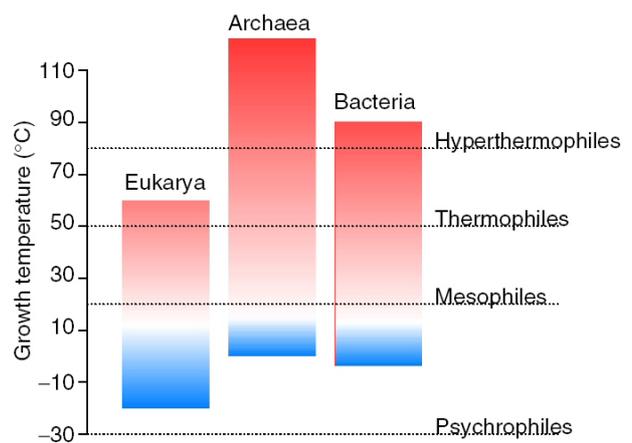


Figure 5 : Distribution phylogénétique de la thermophilie. D'après (Hickey and Singer, 2004).

Hors du contexte cellulaire, l'ADN double brin et la structure secondaire de l'ARN ribosomique peuvent se dénaturer aux températures de croissance des hyperthermophiles. *In vitro*, la température de dénaturation de l'ADN ou l'ARN double brin est affecté par la composition en nucléotides, et l'on pensait, par inférence, que les organismes vivants à une température élevée avaient sélectionné une plus haute proportion de G+C que de A+T dans leur génome, car reliées par trois liaisons hydrogènes qui stabiliseraient les brins d'ADN complémentaires. Ainsi, la structure secondaire des ARN serait stabilisée par un contenu élevé en paires de base GC et par des modifications post-transcriptionnelles (Kawai *et al.*, 1991). Or, à l'analyse d'un grand nombre de génomes séquencés, on pense maintenant qu'il n'y a pas de corrélation évidente entre le contenu en G+C du génome et la température optimale de croissance des organismes (Galtier and Lobry, 1997; Hurst and Merchant, 2001). Par exemple, l'archée *Pyrococcus abyssi* ou la bactérie *Aquifex aeolicus*, toutes deux hautement thermophiles, présentent un contenu en G+C inférieur à 50%, comme de nombreux micro-organismes mésophiles. D'autres mécanismes doivent donc participer à la stabilisation de la structure double brin de l'ADN. La sélection de certains dinucléotides semble contribuer à la thermostabilité (Nakashima *et al.*, 2003). En comparant les génomes d'organismes mésophiles à ceux d'organismes thermophiles, il a été observé une augmentation du taux de bases puriques dans le brin codant, chez les thermophiles (Singer and Hickey, 2003). Des protéines participent également à la stabilisation de la structure double brin de l'ADN. Certaines, qui peuvent être des homologues fonctionnelles aux histones eucaryotes, compactent l'ADN pour former des structures similaires aux nucléosomes des eucaryotes (Reddy and Suryanarayana, 1989; Sandman and Reeve, 2005). Les archées hyperthermophiles présentent de plus une topologie particulière de l'ADN. Une girase inverse, l'ADN topoisomérase de type I introduit des super-tours positifs dans l'ADN et cela contribue à stabiliser la double hélice d'ADN dans des conditions de températures élevées (Forterre *et al.*, 1989; Forterre *et al.*, 1996). Cet ADN girase inverse est caractéristique des hyperthermophiles, elle est absente chez tous les organismes présentant une température optimale de croissance inférieure à 80°C.

Les protéines sont thermostables, c'est à dire qu'elles présentent la capacité de résister et de conserver leurs fonctions à des températures élevées. Cependant, il n'existerait pas de mécanisme universel permettant d'expliquer cette thermostabilité. Il est surprenant d'observer que les protéines homologues mésophiles et hyperthermophiles diffèrent par leur température d'activité, mais sont très similaires en ce qui concerne leurs séquences en acides aminés et leurs structures tridimensionnelles. Cependant, le repliement des protéines semble être un facteur essentiel de la thermostabilité. A des températures identiques, les protéines hyperthermophiles seraient plus rigides, c'est à dire moins flexibles, que leurs homologues mésophiles. Par contre, la flexibilité conformationnelle de ces protéines est similaire à leurs températures physiologiques respectives (Zavodszky *et al.*, 1998). D'autres propriétés communes aux protéines thermostables ont été identifiées, comme par exemple une

augmentation des liaisons ioniques et hydrophobes déterminant la structure tertiaire (Trivedi *et al.*, 2006). Un fort taux d'acides aminés chargés, particulièrement à la surface des protéines, augmente les interactions ioniques, permettant également d'améliorer la thermostabilité de ces protéines (Fukuchi and Nishikawa, 2001; Szilagyí and Zavodszky, 2000). L'augmentation, parmi les thermophiles, de résidus chargés positivement (arginine et lysine) et négativement (acide glutamique) suggère que les liaisons ioniques entre ces résidus de charges opposées stabilisent les protéines à haute température (Tekaiá *et al.*, 2002). Une étude propose, de plus, qu'il y aurait une réduction significative de la fréquence des acides aminés thermolabiles, comme l'histidine, la glutamine et la thréonine, chez les organismes thermophiles (Singer and Hickey, 2003). Des changements subtils dans la séquence en acide aminé des protéines semblent donc être suffisants pour provoquer des changements dans la structure tridimensionnelle et la thermostabilité des protéines. Il existe également des protéines chaperones, de type protéines de choc thermique hsp (*heat-shock proteins*), qui participent à la stabilisation et au repliement des protéines chez les thermophiles (Shockley *et al.*, 2003). Un groupe particulier de protéines chaperones n'est exprimé qu'aux températures de croissances les plus élevées. Elles forment un complexe protéique cylindrique appelé thermosome (Phipps *et al.*, 1991). La structure de celui de *Thermoplasma acidophilum*, l'homologue à la chaperone eucaryote CCT/TRiC a été déterminée (Ditzel *et al.*, 1998). Un nouveau type de chaperones, importantes pour les hyperthermophiles, à de plus été mis en évidence (Makarova *et al.*, 2003). Elles seraient impliquées dans la réorganisation de la chromatine.

2.6. Potentialités biotechnologiques des Archées

Les micro-organismes extrémophiles produisent des enzymes, des catalyseurs biologiques, qui sont fonctionnelles dans des conditions extrêmes. Ces enzymes (parfois appelées extrémozymes) peuvent être valorisables car elles présentent des applications lors de processus industriels comme la transformation de produits chimiques, alimentaires ou pharmaceutiques (Niehaus *et al.*, 1999; van den Burg, 2003). Le marché des enzymes industrielles issus de micro-organismes est estimé à 3,8 milliards d'euros (Antranikian and Egorova, 2007). Les produits dérivés issus de ces enzymes sont estimés à plus de 140 milliards d'euros. La majorité de ces enzymes sont isolées de bactéries ou de champignon, mais les enzymes des archées comme les protéases, les lipases, les enzymes dégradant les polymères comme les cellulases, les chitinases, les amylases présentent également des applications industrielles innovantes (Egorova and Antranikian, 2005). Ces enzymes sont stables et actives à des températures élevées. Dans ces conditions, la solubilité, et donc l'accessibilité des substrats sont améliorés. L'élévation de la température permet aussi de diminuer la viscosité des composés organiques. L'industrie de l'amidon est un secteur particulièrement demandeur en enzymes thermostables. L'amidon est un polysaccharide

constitué de chaîne de glucose, trouvé dans des organes de réserve des plantes pour stocker de l'énergie. L'amidon est utilisé par l'industrie qui peut le transformer en un certain nombre de sous-produits comme des sirops de glucose, du maltose, l'éthanol, le mannitol, le sorbitol, des dextrines, de l'éthanol. Les procédés de transformation de l'amidon s'effectuent à des températures relativement élevées, comprises entre 55 et 110°C. De nombreuses α -amylases et β -amylases thermostables ont été caractérisées chez les archées au cours des quinze dernières années (Chung *et al.*, 1995; Dong *et al.*, 1997; Jorgensen *et al.*, 1997). D'autres enzymes comme des endoglucanases (Bauer *et al.*, 1999), des pullulanases (Brown and Kelly, 1993), des cellulases ont également été décrites et pourraient être utilisées dans des processus de bioconversion.

Une des avancées les plus importantes en biologie moléculaire au cours des 20 dernières années correspond au développement de la réaction de polymérisation en chaîne (PCR) (Mullis *et al.*, 1986) (Saiki *et al.*, 1988) très utilisée en génie génétique et en diagnostic. La première procédure décrite utilisait le fragment Klenow de l'ADN polymérase I de *E. coli*, mais cette enzyme était thermolabile et devait être rajouté durant chaque cycle suivant les étapes de dénaturation et d'hybridation des amorces. L'introduction d'une ADN polymérase thermostable facilita la procédure. C'est l'ADN polymérase I de la bactérie *Thermus aquaticus*, la Taq polymérase, qui était la première ADN polymérase thermostable caractérisée (Chien *et al.*, 1976; Kaledin *et al.*, 1980; Lawyer *et al.*, 1989). Plusieurs ADN polymérases thermostables ont depuis été décrites (Perler *et al.*, 1996) et valorisées pour des applications en PCR, notamment l'utilisation d'ADN polymérases hautement fidèles permettant de réduire les erreurs générées lors de l'amplification. Des polymérases thermostables avec une activité 3'-5' exonucléase de correction sur épreuve ont été décrites, notamment chez les *Archaea* comme la Pwo (Frey and Suppmann, 1996) de *Pyrococcus woesei*, la Pfu (Lundberg *et al.*, 1991) de *Pyrococcus furiosus*, la Vent pol (Cariello *et al.*, 1991) de *Thermococcus litoralis*, la KOD1 DNA pol (Takagi *et al.*, 1997) de *Pyrococcus* sp. Afin d'optimiser, encore l'efficacité de l'amplification de l'ADN par PCR, des champs de recherche important sont tout à fait d'actualité : par exemple, des ADN polymérases mutantes ou chimériques ont été produites afin de répondre à certaines exigences pratiques de la biologie moléculaire (Loh and Loeb, 2005). La caractérisation des ADN polymérases thermostables translésionnelles et des systèmes de réparation de l'ADN des archées hyperthermophiles pourraient aussi être utilisés pour la réplication d'ADN endommagés et permettraient ainsi la récupération d'ADN fossiles ou d'ADN dégradés dans le secteur médico-légal. La recherche de nouveaux partenaires des polymérases, la caractérisation des protéines accessoires⁴ associées aux polymérases lors de la réplication de l'ADN, et leur utilisation lors des processus d'amplification de l'ADN *in vitro* permettrait de rendre les polymérases plus performantes. Cette approche permettrait même de parvenir à l'obtention

⁴ La notion de protéines accessoires des enzymes de la réplication sera définie dans le chapitre concernant la réplication de l'ADN

d'un complexe de réplication minimal actif *in vitro*. L'avantage espéré par cette approche est d'améliorer la processivité de la réaction de polymérisation, c'est à dire synthétiser ou amplifier *in vitro* de longs fragments d'ADN, tout en augmentant encore la fidélité.

Les enzymes des psychrophiles présentent aussi des applications dans des processus industriels effectués à faible température, comme par exemple les amylases et lipases, dans le domaine des détergents. L'industrie de la pâte à papier ainsi que certains processus alimentaires nécessitent également des enzymes capables de fonctionner à faible température. Les enzymes extraites des organismes halophiles et alcalophiles, notamment des protéases, présentent également des propriétés intéressantes (Marhuenda-Egea and Bonete, 2002).

En plus des enzymes dégradant les polymères et des enzymes modifiant l'ADN, il est intéressant de noter que les archées présentent aussi d'autres intérêts biotechnologiques. Quelques métabolites archéens comme des exopolysaccharides, des lipides peuvent présenter des applications industrielles potentielles. A titre d'exemple, les lipides d'archées peuvent être une source de liposomes (archaeosomes) qui pourraient être utilisés comme adjuvant lors de vaccination (Krishnan *et al.*, 2001). Les archées produisent également des agents antimicrobiens comme les archaeocines, les sulfulobicines et les halocines (Prangishvili *et al.*, 2000) (O'Connor and Shand, 2002). Par ailleurs, la biomasse archaeale est utilisée lors des processus de traitement anaérobies des eaux usées, mais également dans la conversion de déchets organiques en méthane qui peut lui même être valorisable en tant que source d'énergie (production de biogaz).

2.7. *Pyrococcus abyssi*

Les archées du genre *Pyrococcus* appartiennent au phylum des Euryarchaeota et à l'ordre des thermococcales. Le genre regroupe quatre espèces : *P. furiosus* (Fiala and Stetter, 1986), *P. woese* (De Wachter *et al.*, 1989), *P. horikoshii* (Gonzalez *et al.*, 1998), *P. abyssi* (Erauso *et al.*, 1993).

Pyrococcus abyssi, la « boule de feu des abysses », a été collectée au niveau d'une cheminée de source hydrothermale à 2000 m de profondeur, au Nord du bassin Fidjien, dans l'Océan Pacifique. *P. abyssi* est un micro-organisme anaérobie strict, chimiohétérotrophe, réduisant le soufre, hyperthermophile avec un optimum de croissance de 96°C (Erauso *et al.*, 1993), a pression atmosphérique, la croissance est possible entre 67 et 102°C, le temps de génération étant de 50 min à 95°C. Lorsque les cellules sont cultivées à la pression *in situ* (20 MPa), la température optimale de croissance est de 100°C. Les cellules se présentent sous la forme de coques de 0,8 à 2 µm de diamètre, mobiles grâce à des flagelles (figure 6). Le génome de la souche GE5 est entièrement séquencé (Cohen *et al.*, 2003) (GenBank AL096836). Il est composé d'un chromosome circulaire de 1,76 Mb, et d'un

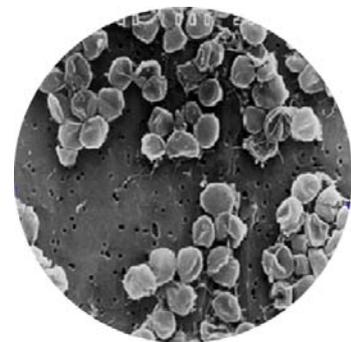


Figure 6 : *Pyrococcus abyssi* observé en microscopie électronique à balayage (Ifremer).

plasmide pGT5 de 3,5 kb (Charbonnier *et al.*, 1992). La teneur en GC est de 44,7%. L'annotation du génome a mis en évidence 1764 cadre de lecture ouverts dont environ 50% n'avait pas de fonction identifiée (Cohen *et al.*, 2003). Le génome est disponible sur :

www-archbac.u-psud.fr/genomes/newpab/newpab.html

Sur le plan métabolique, *P. abyssi* est hétérotrophe, elle nécessite des composés organiques comme des peptides ou des acides aminés comme source de carbone. La croissance est stimulée par la présence de soufre élémentaire. Dans ces conditions, *P. abyssi* peut produire de l'acétate, de l'isovalérate, du propionate, H₂, CO₂ et de l'H₂S. La culture de *P. abyssi* peut être réalisée en milieu liquide, en batch ou en culture continue (Godfroy *et al.*, 2000) mais aussi en milieu solide (Erauso *et al.*, 1995).

P. abyssi a été choisie comme un modèle d'étude de la réplication de l'ADN chez les *Archaea*, notamment en raison des conditions de culture relativement plus simple par rapport à d'autres *Archaea*, mais aussi en raison de la thermostabilité de ses protéines et des applications biotechnologiques potentielles qui pourraient en résulter (Dib *et al.*, 1998) (Ladrat *et al.*, 1996; Mueller *et al.*, 2006). En effet, les constituants liés à la réplication de l'ADN chez *P. abyssi* présente un fort potentiel biotechnologique dans le domaine de la PCR. A titre d'exemple, une des ADN polymérase de cet organisme est désormais commercialisé depuis 2002 sous le nom d'IsisTM par la société Q biogene® (Dietrich *et al.*, 2002). Par ailleurs, comme nous le verrons dans la suite de ce manuscrit, *Pyrococcus abyssi* réplique son génome, dans un contexte procaryote, avec des protéines de type eucaryote. Ces protéines pourraient jouer un rôle identique, chez toutes les espèces, de *Pyrococcus* à l'Homme. En paraphrasant une citation de Jacques Monod, « ce qui est vrai pour *Escherichia coli* est vrai pour l'éléphant », Patrick Forterre et Hannu Myllykallio proposent « ce qui est vrai pour *Pyrococcus* est (presque) vrai pour l'éléphant » (Forterre *et al.*, 2001). Il n'est donc pas aberrant de penser que des protéines qui auraient des fonctions inconnues chez l'Homme, pourraient avoir des homologues chez les *Archaea*. Un projet étudie justement un ensemble de protéines conservées chez les archées et les eucaryotes, absentes chez les bactéries, et de fonction inconnue : les protéines PACE (*Proteins from Archaea Conserved in Eucarya*) (Matte-Tailliez *et al.*, 2000). De plus, les données préliminaires génétiques et biochimiques de ces PACE chez l'Homme, indiquent que certaines seraient impliquées dans la cancérogenèse. L'étude de la réplication de l'ADN chez *Pyrococcus abyssi* permettrait donc ainsi de contribuer aux avancées des connaissances des mécanismes informationnels et de leur régulation chez l'Homme.

* *
*

III. Maintenance génomique chez les *Archaea*

Le génome, regroupant l'ensemble du matériel génétique, dont les gènes, d'une cellule vivante, est constitué d'ADN. L'ADN est le support moléculaire de l'hérédité. Ce génome doit être relativement stable et transmis de façon suffisamment fidèle afin d'assurer la survie des cellules. La réplication de l'ADN est le processus clef qui permet d'assurer le maintien et la transmission de l'information génétique, l'évolution et la propagation des êtres vivants. Malgré sa fiabilité, ce processus peut introduire des mutations. Afin de remédier à ce problème, il existe au sein de la cellule un système de maintenance permettant de corriger les erreurs issues de la réplication mais aussi de réparer des lésions résultant d'agressions chimiques, physiques ou biologiques. Nous proposerons ici un aperçu des mécanismes moléculaires et des protéines impliquées dans la réplication d'une part et dans la réparation d'autre part chez les bactéries et les eucaryotes et nous présenterons leurs équivalents chez les archées.

3.1. La réplication de l'ADN

La réplication, ou la duplication de l'ADN est un processus qui permet à l'ADN de se « reproduire » : deux molécules d'ADN sont obtenues à partir d'une seule.

C'est en 1944 que l'ADN a été identifié comme la molécule responsable de la transmission d'une information génétique. Avery, MacLeod et McCarty ont démontré que l'ADN extrait d'une souche virulente de pneumocoque était capable de transformer une souche avirulente en la rendant pathogène (Avery *et al.*, 1944). Grâce aux techniques de diffraction au rayon X, Watson et Crick ont proposé en 1953 le modèle de l'ADN en double hélice antiparallèle, polynucléotidique (Watson and Crick, 1953). Ils ont de plus suggéré qu'un brin d'ADN pourrait servir de matrice pour conduire à la synthèse de son brin complémentaire. Grâce à leurs travaux sur *Escherichia coli*, le biochimiste Arthur Kornberg et ses collègues décrivent en 1956 une enzyme capable de catalyser la synthèse d'ADN : l'ADN polymérase (Friedberg, 2006; Kornberg *et al.*, 1956).

La réplication du génome conduit à l'obtention de deux doubles hélices. Un brin de chaque molécule est de l'ADN parental et l'autre brin est le brin complémentaire nouvellement synthétisé lors de la réplication. De ce fait, on dit que la réplication de l'ADN est semi-conservative (Meselson and Stahl, 1958). La réplication est dite bidirectionnelle, car elle a lieu dans les deux sens le long de la molécule d'ADN, formant ainsi deux fourches de réplication (Cairns and Davern, 1967).

La réplication de l'ADN peut être partagée en trois étapes :

- l'initiation : il y a reconnaissance de l'origine de réplication et séparation des deux brins d'ADN au niveau de cette origine.

- l'élongation : cette phase correspond à la synthèse de l'ADN. Au cours de cette étape, un grand nombre de protéines (ADN polymérase et facteurs accessoires) interagissent de façon stable ou transitoire, au niveau de la fourche de réplication, afin de former un complexe : le complexe de réplication ou replisome.
- La terminaison : la réplication s'arrête lorsque deux fourches de réplication se rencontrent ou lorsque que le replisome rencontre un complexe spécifique de terminaison de la réplication.

La complexité de la réplication, la diversité des protéines impliquées dans ce phénomène, proviennent probablement de la nécessité d'une très grande fidélité dans la réplication de l'ADN, en permettant ainsi de conserver l'intégrité du génome transmis de génération en génération. Nous exposerons dans les parties suivantes, les mécanismes et les protéines impliquées dans ces différentes étapes chez les bactéries et les eucaryotes et nous présenterons leurs homologues archéens. Le tableau 3 permet de comparer les protéines impliquées dans la machinerie moléculaire de la réplication, dans les trois domaines du vivant.

Fonctions	Bacteria	Eukarya	Archaea <i>Crenarchaea</i>	<i>Euryarchaea</i>	Gène chez <i>P. abyssi</i>
Origine de Réplication	Une	Plusieurs	Une ou plusieurs selon les espèces		
Initiation	DnaA	6 protéines ORC	Cdc6	Cdc6	PAB2265
Hélicase	DnaB	Dna2, MCM	6 MCM	MCM	PAB2373
Stabilisation ADN simple brin	SSB (monomère)	RPA (hétérotrimère)	SSB/RPA	RPA	PAB2163 PAB2164 PAB2165
Création amorce ARN	DnaG (primase)	Sous-unité de ADN Pol α Primase	Primase eukaryotic-like	Primase eukaryotic-like	PAB2235 PAB2236
Polymérase principale de la réplication	Pol III (10 sous-unités) Core $\alpha\epsilon\theta$	Polymérases famille B : α, δ et ϵ	Polymérases famille B	Polymérases famille B et famille D	PAB1128 PAB2404
Protéine accessoire réplisome	Absente	Petite sous-unité de l'ADN Pol δ		Petite sous-unité de Pol D	PAB2266
Facteur de processivité	Sous-unité β de Pol III	PCNA	PCNA	PCNA	PAB1465
Facteur de chargement	Sous-unité γ	RFC	RFC	RFC	PAB0068 PAB0069
Maturation fragments d'Okazaki élimination ARN	Absent (ADN Pol I)	Fen1 RNase H	Fen1	Fen1	PAB1877
Maturation fragments d'Okazaki Ligation	Ligase (dépendante du NAD)	Ligase (dépendante de l'ATP)	Ligase (dépendante de l'ATP et NAD)	Ligase (dépendante de l'ATP et NAD)	PAB2002

Tableau 3 : Comparaison de la réplication chez les bactéries, les eucaryotes et les archées et protéines chez *P. abyssi*. D'après (Forterre *et al.*, 2001) et (Böhlke *et al.*, 2002) modifié. Les *Archaea* présentent la particularité d'avoir un mode réplication de l'ADN impliquant des protéines de « type eucaryote » dans un contexte procaryote.

3.1.1. L'initiation de la réplication

L'initiation de la réplication implique la reconnaissance d'une origine de réplication par un complexe de protéines (les protéines OBP : origin binding proteins). Chez le chromosome des bactéries, il n'y a qu'une seule origine de réplication (OriC) qui consiste en une région riche en A-T. Cette séquence est reconnue par une protéine DnaA (Messer *et al.*, 2001). Chez beaucoup d'espèces bactériennes, le gène codant DnaA est adjacent à l'origine de réplication (Marians, 1992). Plusieurs monomères DnaA vont se fixer à l'origine de réplication. La formation de ce complexe permet l'ouverture de la double chaîne d'ADN et l'entrée de l'hélicase DnaB, associé à DnaC. L'hélicase DnaB peut commencer à dérouler l'ADN, en présence d'ATP, de l'ADN gyrase et de protéines SSB.

A l'inverse des bactéries, les chromosomes eucaryotes présentent de nombreuses origines de réplication : les séquences ARS : (*autonomous replicating sequence*) qui sont en permanence occupées par un complexe de reconnaissance de l'origine (ORC pour *origin recognition complex*). Ces ORC contiennent six polypeptides différents : Orc1-6. Avant le démarrage de la synthèse d'ADN, les protéines ORC doivent recruter un certain nombre de protéines pour former le complexe pré-répliatif. Comme chez les bactéries, l'étape clef de cette phase est le recrutement d'une hélicase. Chez les eucaryotes, c'est la protéine MCM (*minichromosomal maintenance*), composée de six sous-unités (Mcm2-7) qui joue ce rôle (Robinson and Bell, 2005). Les protéines MCM, organisées en complexe hexamérique, forment un anneau. Le recrutement des protéines MCM nécessitent la présence de deux autres facteurs, Cdc6 et Cdt1 (Bell and Dutta, 2002; Stillman, 2005). Par un mécanisme encore mal connue, ces protéines utilisent l'hydrolyse de l'ATP pour faire passer l'ADN à travers le trou central de l'hexamère MCM.

Chez les *Archaea*, le nombre d'origines de réplication varie selon les espèces. *Pyrococcus abyssi* réplique son chromosome de manière bidirectionnelle à partir d'une seule origine de réplication, comme chez les bactéries (Matsunaga *et al.*, 2001; Myllykallio *et al.*, 2000). Les études *in silico*, se basant sur la séquence complète de cet organisme et les comparant avec les séquences connues des origines de réplication eucaryotes ont permis d'identifier une seule origine chez *P. abyssi* (Myllykallio *et al.*, 2000). Des études *in vivo*, ont ensuite confirmé cette hypothèse (Matsunaga *et al.*, 2001). Cependant plusieurs origines de réplication ont été identifiées chez certaines espèces d'archées, comme *Halobacterium* (Zhang and Zhang, 2003), *Sulfolobus* (Lundgren *et al.*, 2004; Robinson *et al.*, 2004), *Haloferax volcanii* (Norais *et al.*, 2007). Des homologues des protéines Cdc6 et MCM eucaryotes ont été identifiées chez la plupart des archées. MCM est capable d'interagir fonctionnellement *in vivo* avec Cdc6 (De Felice *et al.*, 2003; Kasiviswanathan *et al.*, 2005). Au niveau du génome, l'origine de réplication de *P. abyssi* est proche du gène codant Orc1/Cdc6. Cette caractéristique se retrouve également chez les bactéries, où le gène *dnaA* est retrouvé proche

de l'origine (Matsunaga *et al.*, 2001). Chez la crénarchée *Sulfolobus solfataricus*, il y a trois gènes *Orc1/Cdc6* codant les protéines Cdc6-1, Cdc6-2 et Cdc6-3. Les deux origines identifiées chez cet organisme, OriC1 et OriC2, sont localisées en amont des gènes Cdc6-1 et Cdc6-3 respectivement (Robinson *et al.*, 2004). Aucun homologue des protéines DnaB bactériennes et Cdt1 eucaryotes, participant au chargement de l'hélicase au niveau de l'origine de réplication, n'ont été identifiés chez les archées. Cela suggère que les protéines Orc/Cdc6 effectuent seules les fonctions des protéines ORC, Cdc6 et Cdt1 eucaryotes. Cependant, il faut noter qu'il y a une faible similitude, autour de 10%, entre les protéines Cdt1 humaines et de levures. L'analogue fonctionnel de Cdt1 pourrait exister chez les archées mais n'aurait pas encore été identifié (Barry and Bell, 2006).

L'assemblage du complexe de pré-réplication constitue probablement le point clé de la régulation. A leur niveau, les origines de réplication reçoivent « l'autorisation » de démarrer la réplication. Nous ne détaillerons pas ici l'ensemble des mécanismes et des protéines intervenant dans cette régulation. Les brins d'ADN appariés doivent être ensuite séparés. Les topoisomérase de type I et de type II, en modifiant la topologie de l'ADN, favorisent cette séparation. Cette dissociation permet l'entrée des ADN polymérase et des protéines formant le réplisome, et le début de la synthèse d'ADN. Chez les bactéries, la fourche de réplication se déplace à une vitesse comprise entre 500 et 1000 bases par seconde, alors que chez les eucaryotes la vitesse est de 10 à 50 bases par seconde. Chez *Pyrococcus abyssi*, la vitesse de progression de la fourche est estimée à 300 bases par seconde (Myllykallio *et al.*, 2000).

3.1.2. Le réplisome : l'élongation de l'ADN

L'élongation de l'ADN progresse dans le sens 5'→3' et produit un brin complémentaire et anti-parallèle au simple-brin matrice. Cependant, de part sa structure, les deux brins de l'ADN sont enroulés dans des sens opposés. De ce fait, selon le brin d'ADN répliqué, les mécanismes de réplication sont différents.

On peut distinguer (figure 7) :

- Un brin direct ou brin précoce (*leading strand*) qui est le brin complémentaire du brin parental orienté 3'→5'. Le sens de progression de la fourche de réplication est le même que celui de la synthèse du brin d'ADN. La synthèse a donc lieu de façon continue dans le sens 5'→3'.
- Un brin retardé ou tardif (*lagging strand*) qui est le brin complémentaire du brin parental orienté 5'→3'. Dans ce cas, le sens de progression de la fourche de réplication est opposé à celui de sa copie. L'ADN polymérase n'étant capable d'ajouter un nucléoside triphosphate qu'à l'extrémité 3'OH d'un polynucléotide existant, le brin ne peut donc être synthétisé que de façon discontinue, toujours dans le

sens 5'→3', avec formation de petits fragments, appelés fragments d'Okazaki (Kurosawa *et al.*, 1975)⁵.

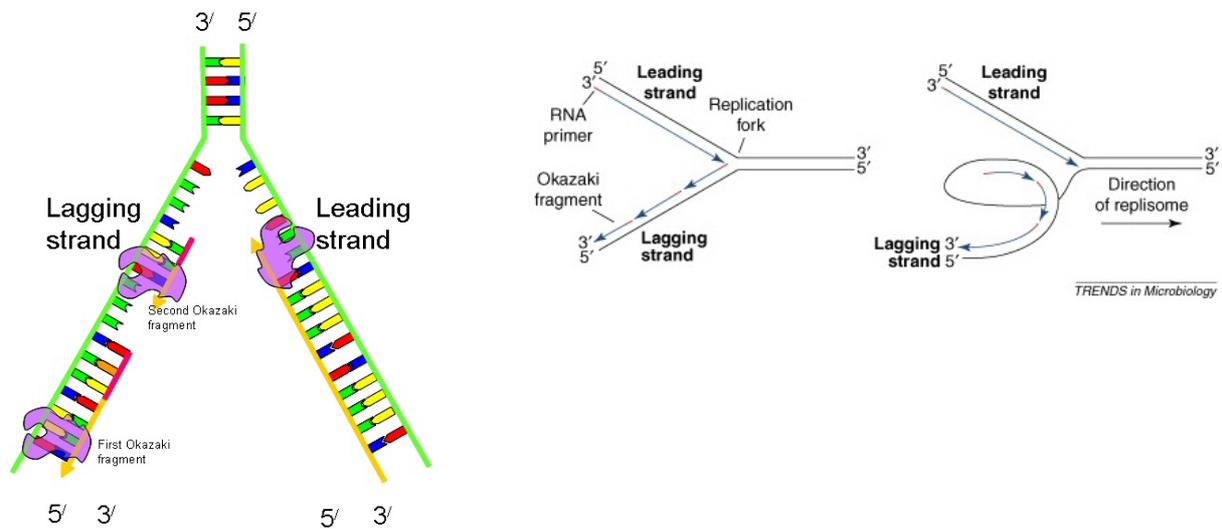


Figure 7 : Schéma général de la fourche de réplication de l'ADN. Le brin précoce (*leading strand*) est synthétisé de façon continue dans le sens 5' vers 3', le brin tardif (*lagging strand*) de façon discontinue sous forme de courts fragments d'Okazaki.

3.1.3. Les protéines de liaison à l'ADN simple brin

Après ouverture des brins d'ADN et naissance de la fourche de réplication, il existe des protéines qui empêcheront les deux brins de se réassocier avant d'avoir été répliqués. Ce sont les protéines affines de l'ADN simple brin (SSB pour *single-strand binding protein*). Ces protéines collaborent au déroulement de l'ADN et protègent l'ADN simple brin des dégradations par les nucléases, des modifications chimiques durant la réplication. Les protéines SSB bactériennes sont des homotetramères alors que chez les eucaryotes, ce sont des hétérotrimères appelés RPA (pour *replication protein A*). Chez les *Archaea*, il y a une variété d'arrangement possible. A titre d'exemple, chez *Pyrococcus furiosus*, il y a trois SSBs, nommés RPA41, RPA32 et RPA14 formant un hétérotrimère avec une haute affinité pour l'ADN simple brin (Komori and Ishino, 2001). D'autre part, les RPA interagissent avec diverses protéines impliquées dans la réparation, la recombinaison et la réparation de l'ADN.

3.1.4. L'ADN Primase

Dans les trois domaines du vivant, et quel que soit le brin, pour commencer la synthèse d'ADN, les ADN polymérases ont besoin d'une petite région double-brin qui sert d'amorce. Les ADN polymérases ont en effet besoin d'un groupement 3'OH libre afin d'initier la polymérisation de l'ADN. A cet effet, il existe dans la cellule des ARN

⁵ Pour information, nous pouvons noter ici qu'il existe d'autres systèmes de réplication ne faisant pas intervenir de brins directs et retardés, ni de fragments d'Okazaki : c'est le cas du mode de réplication des ADN circulaires des mitochondries et des chloroplastes, mais aussi du mode de réplication dit en cercle roulant existant chez les bactériophages.

polymérase ADN dépendante capables de débiter la synthèse de novo sans besoin d'une extrémité 3'OH libre, afin de fournir une amorce pour les polymérase répliquatives. Cette protéine est appelée ADN primase.

La primase des bactéries est DnaG. Elle est capable de synthétiser de l'ARN *de novo* dans la direction 5' vers 3'. DnaG interagit avec la réplique bactérienne (l'ADN polymérase III) et cette interaction limite la taille des amorces synthétisées entre 9 et 14 nucléotides (Zechner *et al.*, 1992). Chez les Eucaryotes, l'activité primase est portée par un complexe formé de l'ADN polymérase α et de l'ADN primase, appelé collectivement Pol α /Primase. C'est un hétérotétramère composé de quatre sous-unités de 165 kDa (p180), 67-86 kDa (p70), 58 kDa (p58) et 48 kDa (p48) (Arezi and Kuchta, 2000). Les sous-unités p48 et p58 portent l'activité primase. C'est la petite sous-unité p48 (PriS) qui porte l'activité catalytique, tandis que la grande sous-unité (PriL) n'est pas catalytique. Au niveau de l'origine de réplique (pour le brin direct) ou au niveau de sites préférentiels (pour le brin retardé), la sous-unité p48 synthétise une amorce ARN de 6 à 15 nucléotides, selon les espèces étudiées. Ensuite, la sous-unité Pol α synthétise un fragment ADN adjacent de 30 désoxynucléotides pour générer une amorce ADN qui peut être élongé par la réplique eucaryote (l'ADN polymérase δ ou l'ADN polymérase ϵ).

Les Archaea présentent des homologues des protéines PriS et PriL eucaryotes mais, l'équivalent de la sous-unité Pol α ne semble pas être présent. Initialement, à la suite des travaux sur *Methanocaldococcus jannaschii*, on pensait que la primase d'Archaea était une enzyme monomérique, homologue à la petite sous-unité eucaryote (Desogus *et al.*, 1999). Les travaux suivant sur *Methanocaldococcus*, mais aussi chez *Archaeoglobus fulgidus*, *Halobacterium* sp. et chez *Pyrococcus* sp. ont permis de mettre en évidence des homologues des petites et grandes sous-unités primase eucaryotes (Bocquier *et al.*, 2001; Matsui *et al.*, 2003). Chez *Pyrococcus furiosus*, la grande sous-unité Pfup46 régule l'activité de la petite sous-unité catalytique Pfup41 (Liu *et al.*, 2001). Étonnamment, la petite sous-unité primase catalytique de *Sulfolobus solfataricus* et *Pyrococcus* sp. peuvent synthétiser des amorces ARN et ADN *in vitro* (Lao-Sirieix and Bell, 2004; Le Breton *et al.*, 2007; Liu *et al.*, 2001; Matsui *et al.*, 2003). Pfup41 synthétise même préférentiellement des oligonucléotides ADN. Cependant, l'addition de la sous-unité PriL augmente l'activité de polymérisation de l'ARN et diminue l'activité de polymérisation de l'ADN.

La petite sous-unité PriS des Archées contient un motif catalytique conservé qui présente une similarité structurelle à celui trouvé chez les polymérase de la famille X comme la polymérase β humaine. (Augustin *et al.*, 2001; Lao-Sirieix *et al.*, 2005a). Les polymérase de cette famille interviennent dans des processus de réplique, de réparation et de recombinaison. Étant donné, qu'à l'exception de *M. thermotrophicum*, les Archées ne possèdent pas de polymérase appartenant à la famille X, la similarité entre ces protéines et la

primase suggère que la primase des Archées pourrait jouer également un rôle dans des mécanismes de réparation (Lao-Sirieix *et al.*, 2005b).

3.1.5. Les ADN Polymérase

Il existe plusieurs ADN polymérase dans les trois domaines du vivant qui interviennent dans tous les processus du métabolisme de l'ADN comme la réplication, la réparation et la recombinaison. Décrire la somme des connaissances amassées au cours des dernières années concernant cette famille d'enzymes sortirait du cadre de ce manuscrit. Nous décrirons ici quelques généralités sur les polymérase des trois domaines.

3.1.5.1. Topologie générale des ADN polymérase

La première activité ADN polymérase a été mise en évidence en 1956 chez *E. coli* (Bessman *et al.*, 1956) et correspondait à l'enzyme appelée par la suite ADN polymérase I. La portion C-terminale de cette protéine, appelée enzyme de klenow a été la première ADN polymérase cristallisée (Ollis *et al.*, 1985). La structure a depuis été résolue (Beese *et al.*, 1993; Freemont *et al.*, 1988). La forme générale, ressemblant à une main, a été décomposée en sous-domaine, *fingers* (doigts), *palm* (la paume) et *thumb* (pouce) (figure 8). Cette organisation a été observée dans presque toutes les structures d'ADN polymérase cristallisées ultérieurement, et cela malgré les faibles pourcentages d'identité de séquence entre ces différentes enzymes. L'activité catalytique de l'ADN polymérase est localisée au niveau du sous-domaine *palm*, dans une crevasse formée par les sous-domaines *fingers* et *thumb* (figure 8). Le domaine *finger* est important pour la reconnaissance et la liaison des nucléotides, le domaine *thumb* est important pour la liaison à l'ADN (Rothwell and Waksman, 2005).

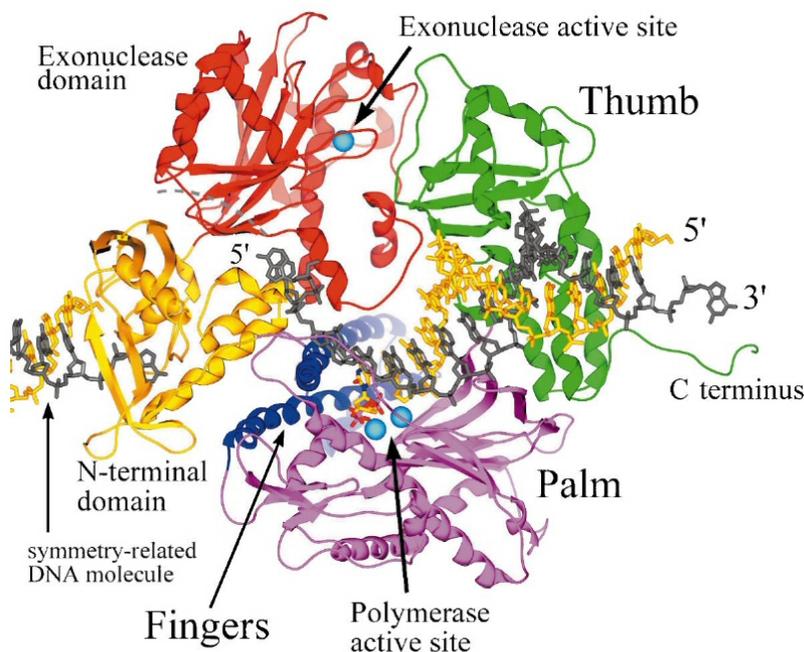


Figure 8 : Structure de l'ADN Polymérase répliquative du bactériophage RB69 (appartenant à la famille B des ADN polymérase) avec l'ADN. Les sous domaines *fingers*, *palm* et *thumb* sont colorés en bleu, magenta et vert respectivement. Le domaine exonuclease est représenté en rouge. La matrice ADN est en gris. D'après (Franklin *et al.*, 2001).

3.1.5.2. Différentes familles d'ADN polymérases

Si toutes les ADN polymérases présentent une similarité fonctionnelle, il existe de grandes différences au niveau de la fidélité lors de la polymérisation de l'ADN (Kunkel, 2004). Les ADN polymérases peuvent être distinguées sur le plan de leur organisation structurale et au niveau de leur implication dans divers mécanismes. Elles peuvent en effet intervenir dans la réplication de l'ADN mais aussi dans des mécanismes de réparation de l'ADN. En plus de présenter une activité de polymérisation de l'ADN 5'→3', certaines polymérases possèdent une activité exonucléasique 3'→5' leur permettant de relire le brin d'ADN en cours de synthèse et de corriger les erreurs éventuelles (*proof reading*). Ces ADN polymérases sont en général hautement fidèles et présentent un taux d'erreur de 10⁻⁷ à 10⁻⁸ par paire de bases répliquées. Certaines présentent également une activité exonucléase 5'→3' permettant le déplacement d'une cassure simple brin (*nick translation*). Chez les trois domaines du vivant, les polymérases peuvent être divisées en différentes familles sur la base de leur structure primaire : les familles A, B, C, D, X, Y (Burgers *et al.*, 2001) (Ohmori *et al.*, 2001). La famille A regroupe les ADN polymérases homologues de l'ADN polymérase I de *E. coli*, comme la Pol γ (Ropp and Copeland, 1996) et la Pol θ (Sharief *et al.*, 1999), alors que la famille B regroupe les ADN polymérases homologues de l'ADN polymérase II de *E. coli* comme les ADN polymérases α , δ , ϵ et ζ . La famille C est celle de la polymérase III des bactéries. La famille X inclut les ADN polymérases β (Nagasawa *et al.*, 2000), λ (Garcia-Diaz *et al.*, 2000), μ (Dominguez *et al.*, 2000) et σ (Wang *et al.*, 2000) qui participent à des mécanismes de réparation. La famille Y regroupe les ADN polymérases η , ι , κ et Rev1 (Ohmori *et al.*, 2001). Les ADN polymérases de cette famille ne présentent pas d'activité de correction sur épreuve et possèdent un taux d'erreur relativement élevé (10⁻¹ à 10⁻³ par paire de bases répliquées). Elles peuvent recopier de l'ADN endommagé, ce sont des ADN polymérases translésionnelles. La famille D comprend seulement des ADN polymérases identifiées chez les euryarchées (Cann and Ishino, 1999).

3.1.5.1. Répartition dans les trois domaines du vivant

Chez les bactéries, cinq ADN polymérases ont été caractérisées. L'ADN polymérase I est impliquée dans la réparation de l'ADN et secondairement dans la réplication. L'ADN polymérase II participe également au mécanisme de réparation de l'ADN. C'est l'ADN polymérase III qui est la réplicase bactérienne. Cette enzyme possède dix sous-unités, elle est également appelée l'holoenzyme Pol III. Les sous-unités $\alpha\epsilon\theta$ sont la partie centrale de l'ADN polymérase III. C'est la sous-unité α qui possède l'activité polymérase. Les sept autres sous-unités modulent l'activité de la polymérase (Johnson and O'Donnell, 2005) (figure 13).

Chez les eucaryotes, plusieurs ADN polymérases ont également été isolées (Hubscher *et al.*, 2002). Les ADN polymérases α , β , δ , et ϵ ont une localisation nucléaire. La polymérase

α est associé à la fonction primase comme nous l'avons vu précédemment. La polymérase δ et ϵ sont les répliques principales qui interviennent durant la réplication de l'ADN. Le rôle respectif de ces polymérases dans la coordination de l'élongation de l'ADN n'a pas encore été clarifié. Cependant, les hypothèses actuelles proposent que chacune soit spécifique de la synthèse d'un des deux brins. L'ADN polymérase δ interviendrait dans la synthèse du brin avancé et l'ADN polymérase ϵ dans la synthèse du brin retardé (Fukui *et al.*, 2004) (Garg and Burgers, 2005). La polymérase β , quand à elle, est impliquée dans des mécanismes de réparation (Wilson, 1998). L'ADN polymérase γ , localisé dans les mitochondries, intervient dans la réplication de l'ADN de ces organites.

Chez les *Archaea*, à ce jour des ADN polymérases appartenant à trois familles distinctes ont été identifiées : des ADN polymérases appartenant à la famille B, un autre type d'ADN polymérases, regroupées dans la famille D, qui paraît être une famille spécifique aux Euryarchaea. Les polymérases de la famille D sont constituées de deux sous-unités, DP1 et DP2 (Cann *et al.*, 1998), et une ADN polymérase de la famille Y (Kulaeva *et al.*, 1996) identifiée chez *Sulfolobus*. Les polymérases des familles B et D présentent toutes les deux à la fois une activité polymérase 5'-3' et une activité exonucléase 3'-5'. En raison de l'absence d'outils génétiques chez les Euryarchaea, il était difficile d'identifier laquelle de ces polymérases était la réplique principale lors de la réplication de l'ADN. Une étude récente propose que les deux ADN polymérases soient requises dans ce mécanisme. La polymérase D étant capable d'élonguer une amorce ARN et possédant une activité de déplacement de brin ARN, en présence du PCNA, interviendrait dans la synthèse du brin retardé (Henneke *et al.*, 2005). Les deux ADN polymérases pourraient interagir de façon séquentielle avec le PCNA chargé, lors de la réplication de l'ADN des Euryarchaea (Rouillon *et al.*, 2007).

3.1.6. Les facteurs accessoires

Le PCNA et le RFC : « the loader of the rings »⁶

Chez les bactéries, les eucaryotes et les archées, la processivité et l'efficacité de la synthèse de l'ADN nécessite une protéine en forme d'anneau, appelée protéine à pince coulissante ou facteur de glissement (*DNA sliding clamp*) qui encercle l'ADN et qui maintient la polymérase sur la matrice ADN (Kelman and O'Donnell, 1995). Cette protéine est chargée sur l'ADN par une autre protéine, le facteur de chargement (*clamp loader*) (Kelman and O'Donnell, 1994). Il a toutefois été observé, chez les archées, que le facteur de glissement était aussi capable de se charger spontanément sur des matrices ADN (Rouillon *et al.*, 2007). Ces protéines sont fonctionnellement conservées dans les trois domaines du vivant.

Chez les bactéries, c'est la sous-unité β de la polymérase III qui forme une structure circulaire encerclant les brins d'ADN et facilitant la progression de la polymérase. Le

⁶ D'après (Trakselis and Bell, 2004)

complexe γ qui fait partie de l'holoenzyme est responsable du chargement et du décrochement de la sous-unité β dans une réaction dépendante de l'ATP.

Chez les eucaryotes, la protéine à pince coulissante, ou facteur de processivité est le PCNA (*Proliferating Cell Nuclear Antigen*). Cette protéine a été décrite à l'origine, en 1978, comme autoantigène lors d'une maladie auto-immune, le lupus érythémateux (Miyachi *et al.*, 1978), mais sa présence est en réalité essentielle dans toutes les cellules, elle a été identifiée deux ans plus tard comme une protéine exprimée durant la phase S du cycle cellulaire (Bravo and Celis, 1980). Le PCNA a été identifié chez tous les eucaryotes, avec une séquence en acide aminé très conservée, des organismes unicellulaires à l'Homme (Kelman and O'Donnell, 1995). Le PCNA est chargé sur l'ADN grâce au RF-C (*Replication Factor C*), également conservé chez tous les eucaryotes (Cullmann *et al.*, 1995). Des protéines homologues au PCNA et à son facteur de chargement RF-C sont identifiées chez toutes les archées séquencées. Elles possèdent des structures très conservées dans les deux domaines. Il a même été montré que les protéines d'archées complètent presque parfaitement le système eucaryote humain. En effet, le PCNA de *Pyrococcus furiosus* interagit fonctionnellement avec l'ADN Pol δ de thymus de veau et stimule son activité (Ishino *et al.*, 2001). Le RF-C humain stimule l'activité de synthèse de la Pol δ en présence du PCNA de *P. furiosus*, ce qui démontre que le RF-C humain est capable de charger le PCNA de *P. furiosus* (Henneke *et al.*, 2002).

Le Facteur de Réplication C (RF-C)

Le RF-C est constitué de cinq sous-unités (heteropentamère), et chacune de ses sous-unités possède une activité ATPase qui est stimulé par la liaison à l'ADN et au PCNA. Les cinq sous-unités sont différentes, mais elles partagent des domaines conservés. Le RF-C appartient à la superfamille des protéines AAA+ (*ATPase associated with a variety of cellular activity*) qui regroupe des protéines chaperonnes capables de fixer et d'hydrolyser l'ATP et de participer à l'assemblage ou au désassemblage d'autres protéines (Neuwald *et al.*, 1999). La fixation de l'ATP entraîne un changement conformationnel du RF-C, ce qui lui permet de fixer le PCNA, de catalyser son ouverture, et de le charger sur l'ADN (Davey *et al.*, 2002) (O'Donnell *et al.*, 2001). Des homologues du RF-C ont été identifiés chez toutes les archées séquencées (Cann *et al.*, 2001a) (Edgell and Doolittle, 1997). Chez *Sulfolobus solfataricus*, chez *Pyrococcus furiosus*, chez *Archaeoglobus fulgidus*, le RF-C est un hétéropentamère comme chez les eucaryotes (Cann *et al.*, 2001b; Pisani *et al.*, 2000; Seybert *et al.*, 2002). L'interaction physique du PCNA avec le RF-C a été bien caractérisée, notamment chez l'espèce *Pyrococcus furiosus* (Matsumiya *et al.*, 2002). La liaison de l'ATP au RF-C est requise pour le chargement du PCNA, mais il a été montré chez *Pyrococcus abyssi* que l'hydrolyse de l'ATP n'était pas nécessaire à ce chargement (Henneke *et al.*, 2002). Des études *in vitro* ont aussi montré que chez certaines espèces, le PCNA pouvait se charger de lui-même sur l'ADN (Cann *et al.*, 2001b). Chez d'autres espèces comme *Methanobacterium thermoautotrophicum*, le RF-C apparaît être un hexamère (Kelman and Hurwitz, 2000).

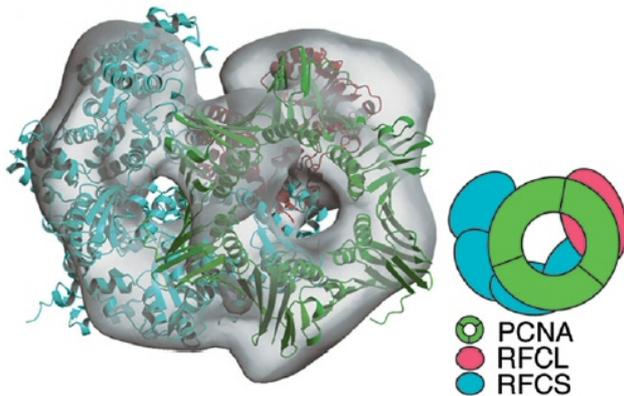


Figure 9 : Structure tridimensionnelle du complexe RFC/PCNA de *Pyrococcus furiosus* en microscopie électronique superposée aux structures cristallines du PCNA et du RFC. D'après (Miyata *et al.*, 2004). Le PCNA est schématisé en vert, les petites et grande sous-unités du RFC sont en bleu et rose respectivement.

Le PCNA et ses partenaires : la communauté de l'anneau

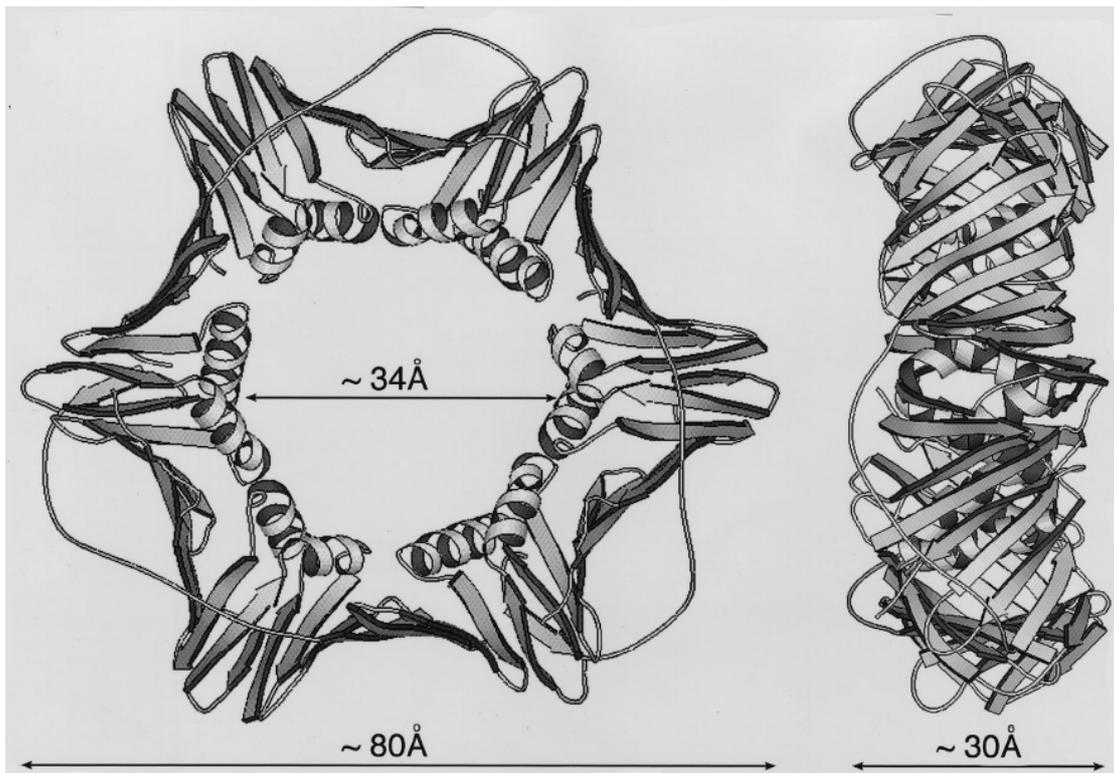


Figure 10 : Structure du PCNA humain, vue de face et vue de profil. D'après (Kelman, 1997).

Le PCNA est une protéine trimérique, en forme d'anneau (Krishna *et al.*, 1994) (Kelman, 1997) (figure 10). La structure du PCNA montre que la surface externe est composée essentiellement de feuillets β alors que la surface interne, entourant l'ADN, est

composée d'hélice α . La topologie du PCNA lui permet donc d'encercler le brin d'ADN et de coulisser le long de ce brin. Le PCNA joue un rôle très important lors de la réplication. Tout d'abord, il est capable d'interagir avec les ADN polymérases, et permet le recrutement de ces polymérases nécessaires à l'élongation de l'ADN. En maintenant les polymérases à proximité de l'ADN, le PCNA améliore leur processivité. En plus des polymérases, le PCNA est capable d'interagir avec un grand nombre d'autres protéines ; Giovanni Maga qualifie le PCNA de : « *dancer with many partners* » (Maga and Hubscher, 2003). Beaucoup de protéines interagissant avec le PCNA possèdent un motif de liaison au PCNA appelé PIP-box, comportant une séquence consensus de type Qxxhxxaa, où h représente un résidu avec une chaîne modérément hydrophobe (L, I, M), a représente un résidu aromatique hautement hydrophobe (F, Y) et x tout type d'acide aminé (Jonsson *et al.*, 1998; Warbrick, 2000). Un autre motif de fixation au PCNA, la KA-Box, a été identifié dans certaines protéines (Xu *et al.*, 2001). A titre d'exemple, chez les eucaryotes, sont capable d'interagir avec le PCNA, l'ADN polymérase δ , l'ADN polymérase ϵ , l'ADN polymérase β , mais aussi d'autres ADN polymérases translésionnelles, le RF-C, l'ADN ligase 1, Fen1, les topoisomérases I et II, la protéine p21, des cyclines intervenant dans le cycle cellulaire, etc (Maga and Hubscher, 2003; Vivona and Kelman, 2003). Le PCNA présente un rôle capital lors de la réplication du génome, mais aussi dans d'autres événements du métabolisme de l'ADN comme la réparation et la recombinaison, ainsi que dans la régulation du cycle cellulaire. Le PCNA interagit aussi avec des protéines participant à des processus post-réplicatifs comme l'assemblage de la chromatine et la cohésion des chromatides sœurs (Maga and Hubscher, 2003). Le PCNA semble donc capable de recruter des protéines à un moment précis, à un endroit précis, de faciliter des interactions entre une protéine et son substrat. Il coordonne l'action de ces protéines et les différences d'affinité des protéines pour le PCNA pourraient participer à cette coordination par des mécanismes de compétition pour les sites de fixation au PCNA (Maga and Hubscher, 2003; Prospero, 2006). D'autres phénomènes comme des phosphorylations par des protéines kinases et des phosphatases, des modifications du PCNA (ubiquitination ou SUMOylation) participent également au changement de partenaires du PCNA et donc à l'orchestration des différents processus de maintenance de l'ADN (Moldovan *et al.*, 2007).

La structure du PCNA des *Archaea* est superposable à celle des facteurs de glissement eucaryotes et bactériens (Matsumiya *et al.*, 2001) (figure 11). Chez presque toutes les euryarchées, comme chez les Eucaryotes, il n'y a qu'un seul homologue du PCNA qui forme un homotrimère. *Thermococcus kodakarensis* est la seule euryarchée connue présentant deux homologues du PCNA (Fukui *et al.*, 2005). Par contre, la majorité des Crénarchées pour lesquels le génome est disponible présentent plusieurs homologues du PCNA. Par exemple, il y a deux homologues au PCNA chez *Sulfolobus solfataricus* (De Felice *et al.*, 1999) et trois chez *Aeropyrum pernix* (Daimon *et al.*, 2002). Le PCNA serait hétérotrimérique chez *Sulfolobus solfataricus* (Dionne *et al.*, 2003), mais il pourrait y avoir des formes homotrimériques et hétérodimériques chez *Aeropyrum pernix* (Daimon *et al.*, 2002).

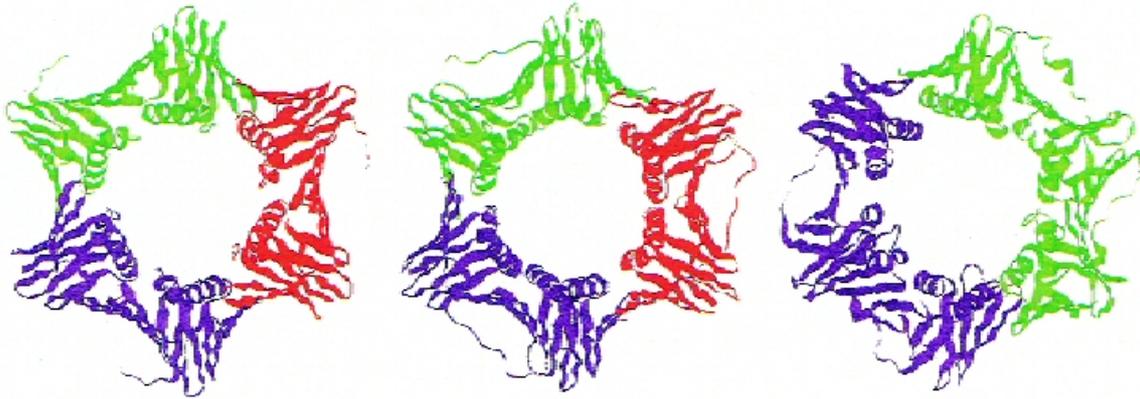


Figure 11 : comparaison dans les trois domaines du vivant : à gauche, le PCNA trimérique de *Pyrococcus furiosus* (Matsumiya *et al.*, 2001), au centre, le PCNA trimérique de *Sacharomyces cerevisiae* (Krishna *et al.*, 1994), à droite l'analogue bactérien, la sous unité β dimérique de *Escherichia coli* (Kong *et al.*, 1992).

3.1.7. Les protéines interagissant avec le PCNA et intervenant dans la maturation des fragments d'Okazaki.

Lors de la réplication de l'ADN, une fois le PCNA chargé sur l'ADN, il recrute l'ADN polymérase répliquative, et l'amorce peut être élongée. Durant la réplication, le PCNA peut toutefois lier d'autres types de protéines. Les interactions bien documentées durant ce mécanisme sont celles avec Fen1 et l'ADN ligase 1 (Vivona and Kelman, 2003).

La protéine flap endonucléase 1 (Fen1) (Harrington and Lieber, 1994) joue un rôle très important dans la maturation des fragments d'Okazaki, en enlevant les amorces ARN et ADN durant la synthèse du brin retardé (Hubscher and Seo, 2001; Lieber, 1997). Fen1 est une endonucléase capable de couper le brin d'ADN libre (*flap*) engendré par le déplacement de brin de l'amorce en aval, par l'ADN polymérase répliquative. Suite à la dégradation de l'amorce ARN/ADN, la polymérase répliquative synthétise l'ADN jusqu'à sa rencontre avec l'extrémité 5' créée par Fen1. La ligase 1 relie ensuite les deux fragments d'ADN ensemble. L'amorce peut aussi être dégradée grâce à une ARN exonucléase, la RNaseHIII, mais toutefois Fen1 interviendra pour éliminer le dernier ribonucléotide de l'amorce. Nous pouvons préciser ici, pour information, que la protéine Fen-1 participe en outre à des mécanismes de réparation de l'ADN (Henneke *et al.*, 2003), mais nous y reviendrons ultérieurement au cours de ce manuscrit.

La protéine Fen1 est donc capable d'interagir avec le PCNA, et de plus elle est stimulée par celui-ci (Tom *et al.*, 2000). Les structures de plusieurs Fen1 humaines ou archées ont été résolues, soit seules, soit en complexe avec le PCNA et le substrat ADN (Chapados *et al.*, 2004; Hosfield *et al.*, 1998). Fen1 a été identifiée dans les génomes d'archées disponibles, elle présente une séquence en acide aminé similaire à celle des eucaryotes (Matsui *et al.*, 1999; Shen *et al.*, 1998). Les études biochimiques de Fen1 chez *Methanococcus jannaschii*

(Rao *et al.*, 1998) et chez *Pyrococcus horikoshii* (Matsui *et al.*, 1999) ont révélées de plus des propriétés biochimiques similaires à celles connues chez les eucaryotes.

3.1.8. Autres protéines : les GINS

Le complexe protéique hétérotétramérique GINS (*go, ichi, nii, san* ; cinq, un deux, trois en japonais ou Sld5, Psf1, Psf2, Psf3) a tout d'abord été découvert, récemment, chez l'eucaryote, le xénope et la levure (Gambus *et al.*, 2006; Kubota *et al.*, 2003; Takayama *et al.*, 2003). C'est le composant de la machinerie de réplication le plus récemment identifié. Il est requis pour l'initiation de la réplication et également pour la progression de la fourche de réplication (Labib and Gambus, 2007). Chez la levure, ce sont des protéines essentielles et elles sont capables d'interagir avec le MCM et Cdc45, à l'origine de réplication et seraient importantes pour le recrutement des ADN polymérases. Les GINS seraient recrutées à l'origine de réplication et se déplaceraient avec la fourche de réplication (Calzada *et al.*, 2005). Chez les *Archaea*, un homologue au complexe GINS a été identifié lors d'une recherche de protéines interagissant avec le MCM par double-hybride chez *Sulfolobus solfataricus* (Marinsek *et al.*, 2006). L'interaction avec le MCM a été démontrée *in vivo* et *in vitro*. Les investigations ultérieures ont montrés qu'une des sous-unités des GINS archéens, codés par le gène *gins23* présente plus d'homologie avec les protéines eucaryotes codées par les gènes *psf2* et *psf3*, alors que *gins15* est plus proche de *sld5* et *psf1*. Les études d'interaction additionnelles ont montré que Gins23 est capable d'interagir avec la primase, suggérant que les GINS pouvaient créer un pont entre le MCM et la primase (figure 9). Chez l'Homme, l'interaction des protéines GINS avec l'ADN polymérase primase α a également été démontrée (De Falco *et al.*, 2007).

* *
*

Les figures 12 à 16 proposent des modèles des protéines et mécanismes mis en jeu lors de la réplication de l'ADN dans les différents domaines du vivant.

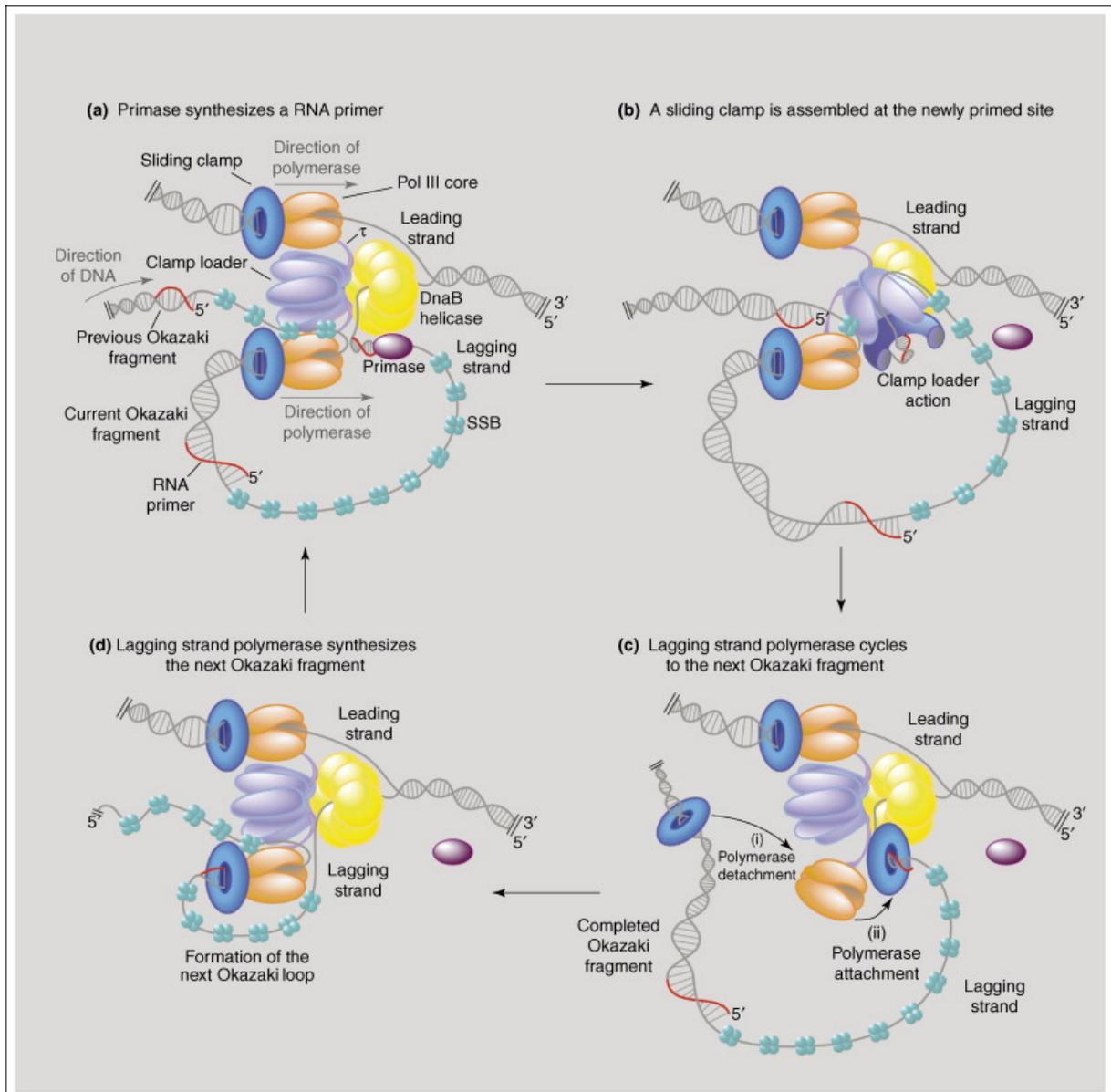


Figure 12 : Dynamique des protéines au niveau de la fourche de réplication de *E. coli*. La synthèse du brin direct (*leading strand*) est un processus continu, dans le sens de progression de la fourche de réplication. La synthèse du brin retardé (*lagging strand*) est discontinue, avec la formation de petits fragments d'Okazaki. L'hélicase DnaB permet l'ouverture de la double chaîne d'ADN. Les protéines SSB stabilisent l'ADN simple brin et empêchent la formation de structure secondaire. (a) Un fragment d'Okazaki est initié quand la primase (DnaG) synthétise une courte amorce ARN. (b) Le complexe γ (facteur de chargement ou *clamp loader*) ouvre et assemble le facteur de processivité (β *clamp*) au nouveau site amorcé. (c) L'ADN polymérase du brin retardé se dissocie du β *clamp* du fragment d'Okazaki achevé et s'assemble au β *clamp* au nouveau site amorcé. (d) L'ADN polymérase initie la synthèse des fragments d'Okazaki sur le brin retardé. D'après (Pomerantz and O'Donnell, 2007).

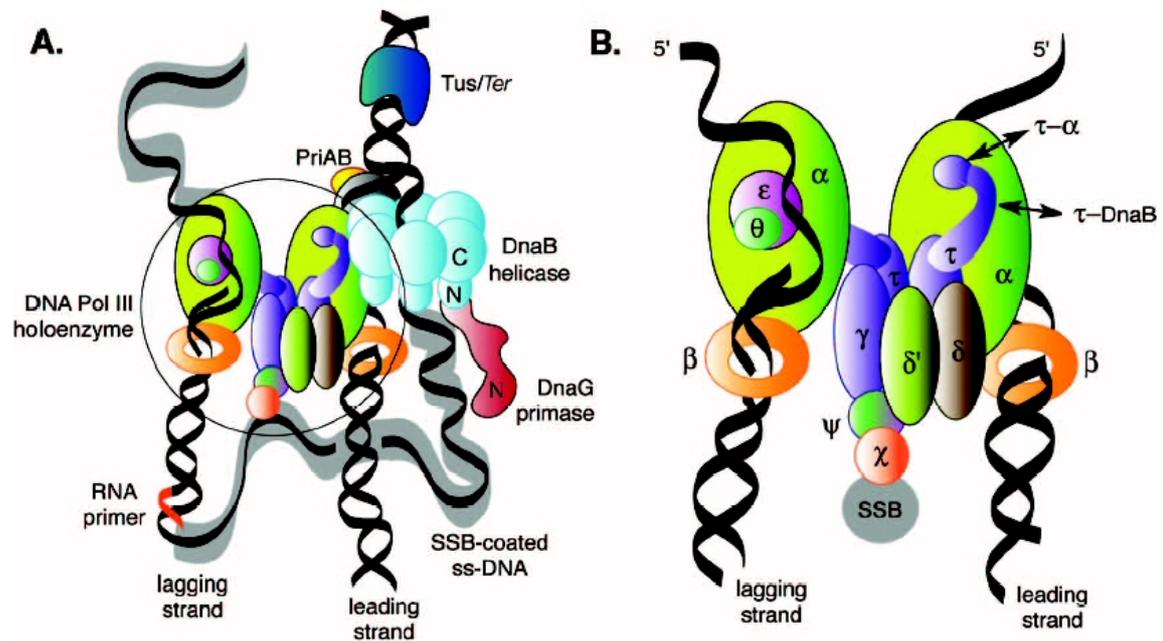


Figure 13 : Représentation schématique du réplisome de la bactérie *Escherichia coli* (A) et de sa réplicase, l'holoenzyme ADN polymérase III. D'après (Schaeffer *et al.*, 2005).

L'hélicase DnaB permet l'ouverture des brins, les protéines SSB stabilisent l'ADN simple brin, la primase DnaG synthétise l'amorce ARN. L'ADN polymérase III intervient l'élongation de l'ADN. Le facteur β améliore sa processivité.

(B) L'ADN polymérase de *E. coli* est constituée d'une partie centrale, le domaine core, constitué des sous-unité $\alpha\epsilon\theta$, portant l'activité polymérase et exonucléase, du facteur de processivité β et du facteur de chargement $\tau, \gamma, \delta, \delta', \chi$ et ψ .

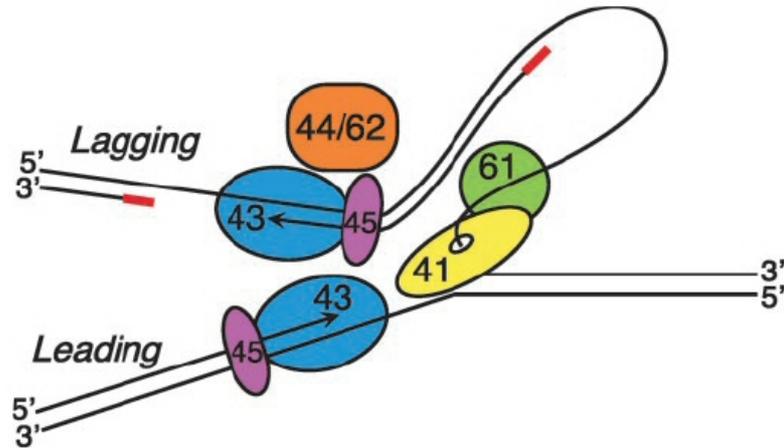


Figure 14 : représentation schématique de la fourche de réplication chez le bactériophage T4 (Joyce, 2004). Le génome du phage T4 code lui même les protéines intervenant dans sa réplication. Deux ADN polyméras (gp43) synthétise l'ADN, une au niveau du brin direct, l'autre au niveau du brin retardé, en complexe avec le facteur de processivité (gp45) et le facteur de chargement (gp44/62). L'hélicase (gp41) est associé avec la primase (gp61). La protéine gp32 (non représentée ici) s'associe et stabilise l'ADN simple brin.

Model of a eukaryotic replication fork

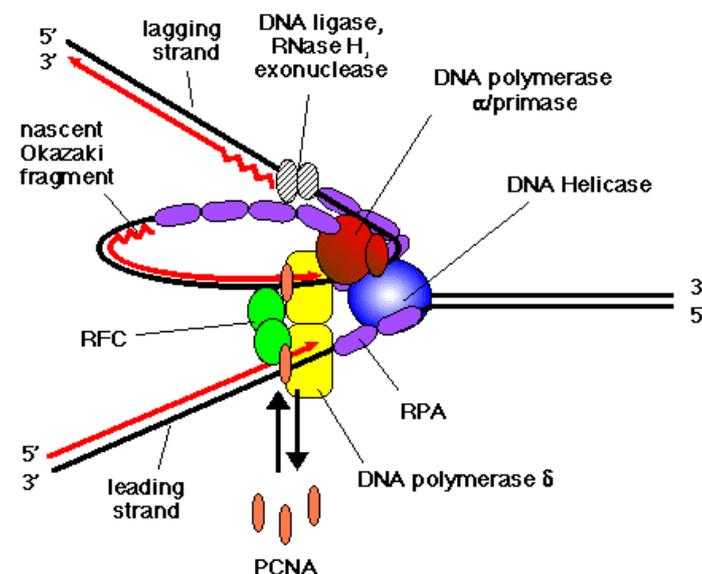


Figure 15 : Modèle simplifié de la réplication de l'ADN chez l'eucaryote. L'hélicase MCM permet l'ouverture de la double chaîne d'ADN. Les protéines RPA interviennent pour stabiliser l'ADN simple brin. Le complexe α -primase synthétise l'amorce ARN. Le facteur de processivité PCNA est chargé sur l'ADN grâce à son facteur de chargement RFC. Les ADN polyméras Pol ϵ et Pol δ synthétise l'ADN au niveau du brin avancé et du brin retardé. Les protéines Fen1 et RNase H permettent l'élimination ultérieure de l'amorce.

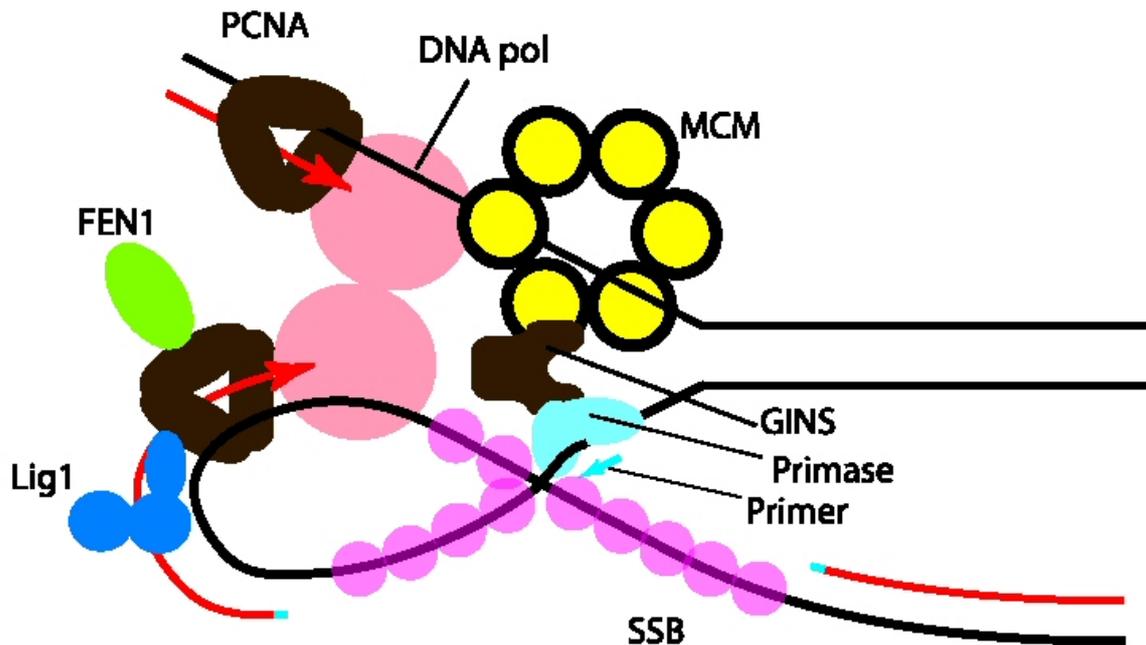


Figure 16 : Modèle proposant l'architecture de la fourche de réplication chez l'*Archaea Sulfolobus*, d'après (Barry and Bell, 2006). Le brin d'ADN parental est indiqué en noir et le brin nouvellement synthétisé est indiqué en rouge. Les amorces ARN synthétisées par la primase sont indiquées en bleu. L'ADN simple brin est lié aux protéines SSB (nomenclature chez *Sulfolobus*). L'hélicase MCM interagit avec le complexe de protéines GINS, interagissant lui-même avec la primase. L'ADN polymérase élongue l'amorce et interagit avec le PCNA trimérique. Le PCNA sert de plaque-tournante permettant l'assemblage d'autres acteurs comme Fen1 ou Lig1. Le facteur de chargement du PCNA, le RFC, n'est pas représenté ici.

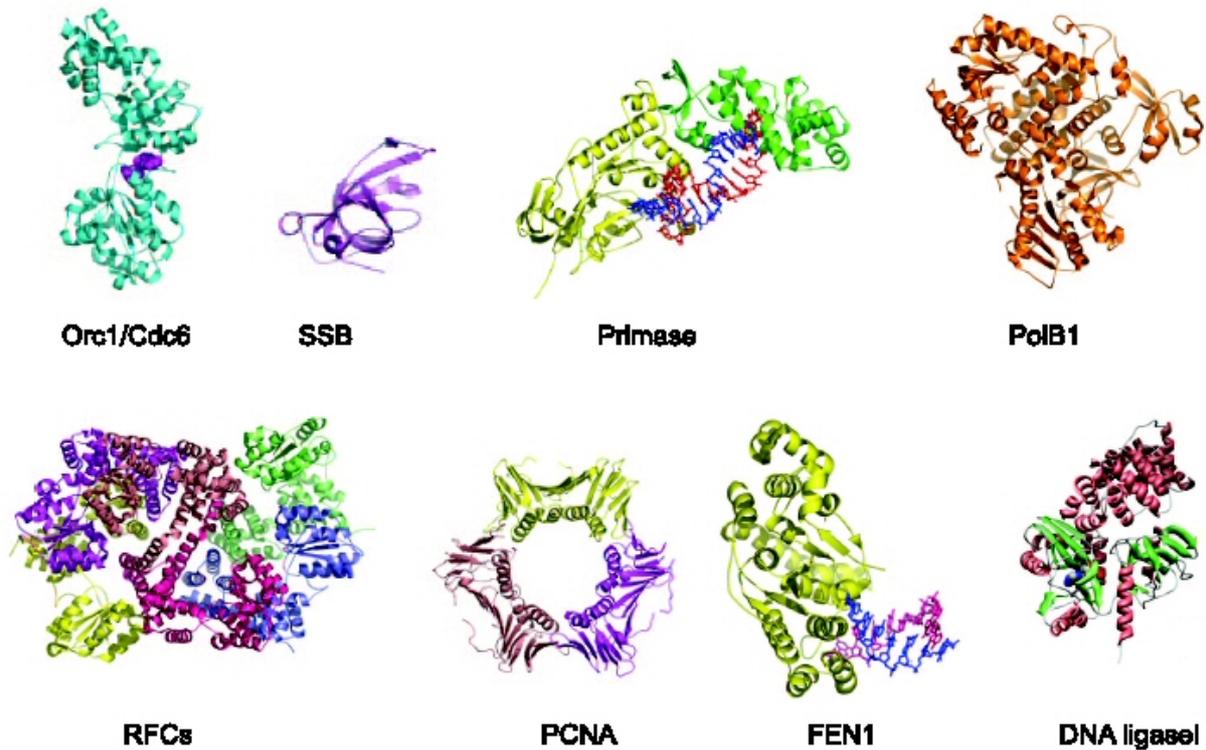


Figure 17 : Structure des différentes protéines intervenant dans la réplication de l'ADN chez les *Archaea*. les figures ont été obtenues en utilisant Pymol, les codes PDB sont indiqués. Cdc6 de *Pyrobaculum*, 1FNN ; SSB de *Sulfolobus*, 1O71 ; petite sous unité du RFC de *Pyrococcus*, 1IQP ; Polymérase B1 de *Sulfolobus*, 1S5J ; PCNA de *Pyrococcus*, 1ISQ ; Fen1 de *Archaeoglobus*, 1RXW. D'après (Barry and Bell, 2006).

* *
*

Nous avons donc vu comment la cellule pouvait dupliquer fidèlement son génome. Cependant, la réplication ou d'autres mécanismes peuvent entraîner des modifications, mutations ou lésions, de l'ADN. Il existe donc des mécanismes permettant le maintien de la stabilité du génome. Ce sont les mécanismes de réparation de l'ADN. Les liens avec le processus de la réplication sont parfois évidents. Nous l'avons vu précédemment, certaines protéines intervenant dans la réplication de l'ADN participent également à des mécanismes de réparation. Seuls quatre mécanismes seront abordés dans le chapitre suivant : la réparation par excision de nucléotide (NER), la réparation par excision de bases (BER), la réparation des mésappariements (MMR) et la réparation des cassures double brin (DSBR).

* *
*

3.2. Processus de réparation de l'ADN chez les *Archaea*

Une grande variété de lésions de l'ADN peuvent être induites par des agents externes (rayonnement ionisant, agents chimiques) ou produites par l'activité normale de la cellule (radicaux libres, instabilité des bases azotés de l'ADN : déamination, dépurination, oxydation) (figure 18). Des erreurs lors de la réplication de l'ADN peuvent introduire des mutations (mésappariements, insertion ou délétion de bases). Ces dommages peuvent entraîner des altérations de la structure de l'ADN pouvant bloquer la réplication de l'ADN ou la transcription. Les mutations défavorables peuvent entraîner la mort de la cellule.

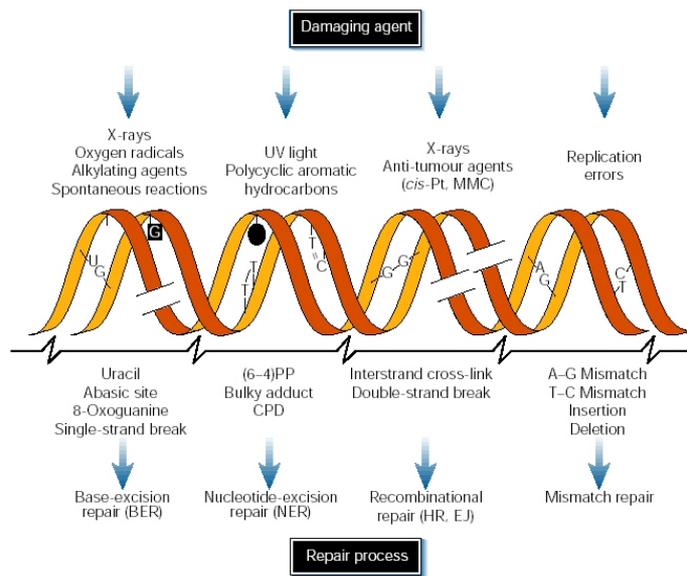


Figure 18 : schéma montrant la diversité des lésions de l'ADN induit par différents agents, et les mécanismes de réparation de l'ADN les plus pertinents pour enlever ces lésions, chez les eucaryotes. D'après (Hoeijmakers, 2001).

Il existe toute une diversité de protéines garantissant l'intégrité et la stabilité de l'ADN. L'étude des mécanismes de la réparation de l'ADN est actuellement un champ de recherche très compétitif, en utilisant comme modèle d'étude *Escherichia coli* et *Saccharomyces cerevisiae*. Trois grands systèmes ont été décrits : la réversion du dommage, la réparation des cassures double brin et la réparation par excision (incluant les mécanismes de réparation par excision de bases, BER ; la réparation par excision de nucléotides, NER ; et la réparation des mésappariements, MMR). Par contre, l'étude de ces mécanismes chez les *Archaea* en est encore à ses balbutiements, le retard étant dû en partie à la difficulté de réaliser des systèmes génétiques chez ces organismes. Les récentes découvertes montrent des similarités entre les systèmes eucaryotes et archées (Grogan, 2004). Toutefois, il est intrigant de remarquer chez les archées, l'absence de certaines protéines de réparation de l'ADN, par ailleurs très conservées chez d'autres organismes. L'analyse des génomes archéens séquencés montre qu'il manque des composants cruciaux impliqués dans les processus de NER et MMR. De nouvelles protéines, de nouvelles voies de réparation pourraient être découvertes chez les archées, en particulier chez ceux vivant dans des environnements atypiques. Ainsi, l'étude du

maintien de la stabilité et de la réparation de l'ADN chez les hyperthermophiles est une activité qui éveille la curiosité. En effet, l'augmentation de la température augmente les processus de lésions spontanées de l'ADN. Cependant, le taux de mutation n'est pas particulièrement élevé chez les hyperthermophiles, proche de celui rencontré chez *E. coli*, ce qui suggère qu'il existe des systèmes de réparation efficaces chez ces organismes (Grogan *et al.*, 2001).

3.2.1. La réparation par excision de nucléotide (*nucleotide excision repair NER*)

La réparation par excision de nucléotide intervient pour corriger des lésions de l'ADN comme les dimères de pyrimidine. Lors de ce mécanisme, l'ADN contenant la lésion est excisé puis la brèche est comblée par une ADN polymérase qui utilise le brin non endommagé comme matrice. La bactérie *E. coli* utilise un complexe de protéines, l'endonucléase UvrABC, qui est capable de cliver environ 12 nucléotides du brin contenant le dimère de pyrimidine (Van Houten *et al.*, 2005), alors que chez les eucaryotes, une machinerie plus complexe crée une brèche d'environ 28 nucléotides. Chez les eucaryotes, la reconnaissance de la lésion requiert la protéine XPC-RAD23B (Rad4-Rad23 chez la levure). Les protéines TFIIH, XPA, RPA et XPG (TFIIH, Rad14, RPA et Rad2 chez la levure) sont nécessaires pour former un complexe de préincision. Le complexe TFIIH, et notamment deux de ses sous-unités, les hélicases XPB et XPD (Ssl2 et Rad3 chez la levure) interviennent pour ouvrir l'ADN au niveau de la lésion. La protéine XPG (Rad2 chez la levure) coupe l'ADN en 3' de la lésion et ERCC1-XPF (Rad1-Rad10 chez la levure) clive du côté 5'. La brèche ainsi formée est ensuite comblée par la polymérase δ ou ϵ avec l'intervention des protéines PCNA, RFC et RPA. L'ADN ligase va ensuite ligaturer les fragments adjacents (Friedberg *et al.*, 2006).

Chez les *Archaea*, des homologues des protéines bactériennes UvrABC sont retrouvées seulement chez certaines archées mésophiles, et cela pourrait être le reflet d'un transfert latéral de gènes avec un donneur bactérien (Grogan, 2000). Pour la majorité des *Archaea*, et notamment pour les hyperthermophiles, une autre voie doit être supposée. Beaucoup d'archées présentent des homologues aux nucléases XPF-ERCC1 et Fen1/XPG eucaryotiques, aux hélicases XPB et XPD, rendant possible un mécanisme NER de type eucaryote fonctionnel (Kelman and White, 2005). Par contre, il manque certains homologues de reconnaissance de la lésion comme XPA et XPC chez les archées, un homologue à XPF est aussi absent chez les thermoplasmatales. Nous pouvons préciser ici, que les protéines XPA et XPC ne sont pas présentes chez tous les eucaryotes ce qui laisse penser que ces protéines sont apparues ultérieurement, dans certaines branches d'eucaryotes, au cours de l'évolution. Il a été suggéré que, tout comme la réplication de l'ADN, le mécanisme de réparation NER chez les archées est une version simplifiée du mécanisme eucaryote (White, 2003).

3.2.2. La réparation par excision de bases (*Base excision repair BER*)

Le système de réparation par excision de bases (BER) reconnaît plus spécifiquement des bases modifiées ou endommagées, résultats de processus chimiques, comme l'oxydation ou l'alkylation de bases, la méthylation, la déamination des cytosines en uracile. Ces lésions sont fréquemment rencontrées chez tous les types d'organismes, souvent générées lors du métabolisme cellulaire, et leur fréquence peut augmenter chez les hyperthermophiles. Plusieurs réactions successives interviennent lors du BER. Il est initié par une ADN glycosylase qui reconnaît et enlève la base modifiée et génère ainsi un site AP (apurique ou apyrimidique). Le site AP est ensuite reconnu par une AP endonucléase (Ape1) qui enlève le sucre en 5', créant une cassure de brin. Cette cassure présente un groupe hydroxyl en amont (3'-OH) et un groupe déoxyribose phosphate en aval (5'-deoxyribose-5-phosphate ou 5'dRP) (Demple *et al.*, 1991). Chez les mammifères, il peut ensuite y avoir deux voies distinctes de BER impliquant des enzymes différentes : le SN-BER (*single nucleotide repair*) concerne le remplacement d'un nucléotide et le LP-BER (*long patch repair*) concerne le remplacement de 2 à 8 nucléotides (Sung and Demple, 2006). La figure 19 schématise les deux voies du BER chez l'eucaryote.

Dans le SN-BER, chez les eucaryotes, c'est la polymérase β qui ajoute d'une part à l'extrémité 3'-OH un nucléotide à la place de l'ancienne lésion et qui d'autre part excise le groupe 5'dRP grâce à son activité dRPase. Par conséquent, une extrémité 5'-phosphate est ainsi générée qui sera ligaturée à l'extrémité 3'-OH par une ADN ligase.

Dans le LP-BER, une polymérase ajoute plusieurs nucléotides à l'extrémité 3'-OH et génère ainsi un déplacement de brin. Ce déplacement de brin entraîne ainsi la formation d'un brin d'ADN libre (structure *flap*) contenant le groupe 5'dRP et ce brin est clivé grâce à la nucléase Fen1 (Klungland and Lindahl, 1997; Prasad *et al.*, 2000). L'identité de la polymérase impliquée dans ce déplacement de brin n'est pas encore bien définie, plusieurs ADN polymérases ont été proposées pour ce mécanisme : Pol β , Pol δ ou Pol ϵ (Sung and Demple, 2006). Une étude a reconstitué le mécanisme du LP-BER en utilisant les enzymes purifiées Ape1, Pol β , Pol δ , le PCNA, Fen1 et l'ADN ligase I ou Pol δ remplace Pol β quand le PCNA est absent (Klungland and Lindahl, 1997).

Dans la cellule eucaryote, la protéine XRCC1 en interagissant avec les autres protéines participant au BER, coordonne les activités enzymatiques de ces protéines (Nazarkina Zh *et al.*, 2007).

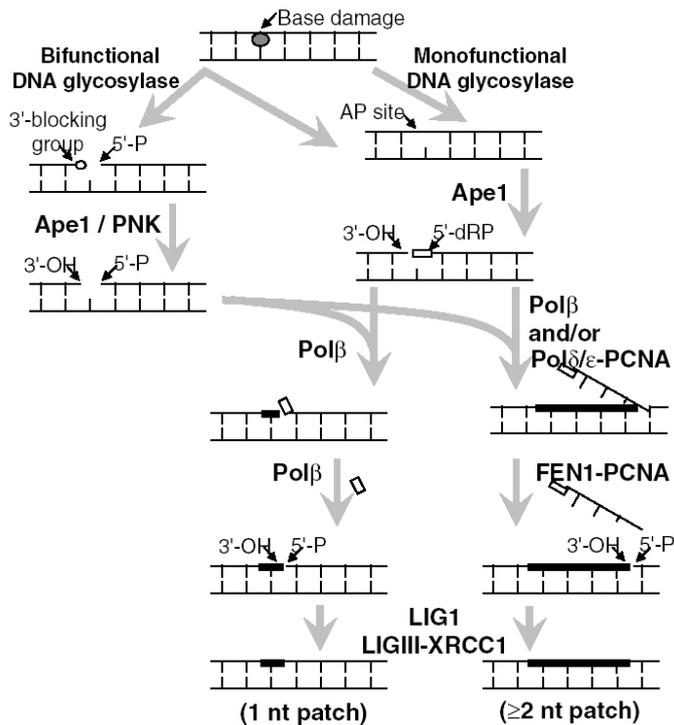


Figure 19 : système de réparation par excision de bases (BER) chez les eucaryotes. Les étapes impliquées dans les voies du SN-BER et du LP-BER sont discutées dans le texte. D'après (Sung and Demple, 2006).

Le système du BER a été reconstitué *in vitro* chez la crénarchée *Pyrobaculum aerophilum* en utilisant les protéines recombinantes uracile ADN glycosylase, AP endonucléase, une ADN polymérase de la famille B et l'ADN ligase (Sartori and Jiricny, 2003). L'étude a montré que l'utilisation de ces protéines a permis de catalyser la réparation d'un mésappariement G-U. Les données suggèrent que le système du BER des archées serait similaire à celui rencontré chez les bactéries et les eucaryotes. Cette voie serait très conservé dans les trois domaines du vivant.

3.2.3. La réparation des mésappariements (*Mismatch Repair MMR*)

Le système de réparation des mésappariements de bases intervient notamment lors de la mésincorporation de bases, au moment de la réplication de l'ADN. Le système MMR nécessite la détection du mésappariement, l'excision d'un brin et resynthèse de ce brin par copie du brin complémentaire. Le mécanisme le mieux connu est le système de réparation généralisée (SRM) Mut de *E. coli*. Il fait intervenir 12 protéines dont trois sont spécifiques à ce mécanisme : MutS, MutL et l'endonucléase MutH (Marti *et al.*, 2002). D'autres protéines interviennent comme l'hélicase UvrD, l'adénine méthyl-transférase Dam, quatre exonucléases (I, VII, X et RecJ), SSB, l'ADN polymérase III et l'ADN ligase. Plusieurs protéines homologues à MutS et MutL (mais pas MutH) sont retrouvées également chez les eucaryotes (les composants centraux sont MSH2, MSH3 et MSH6 homologues à Mut S ; MLH1 homologue de MutL), mais étonnamment ces protéines sont absentes de la plupart des génomes d'archées. Elles sont retrouvées seulement chez quelques archées mésophiles et

pourraient correspondre tout comme UvrABC à des transferts latéraux de gènes à partir de bactéries (White, 2007). Chez les bactéries, des mutations spontanées peuvent se produire à une fréquence de 10^{-6} à 10^{-7} par nucléotides répliqués. Lorsque les gènes *mutS* et *mutL* sont inactivés, cette fréquence de mutation est augmentée par 100, ce qui montre l'importance du système MMR (Matic *et al.*, 1997). L'absence d'un tel système chez les *Archaea* serait donc problématique. Il a été suggéré qu'il n'y ait pas de système de type MMR fonctionnel chez la crénarchée *Pyrobaculum aerophilum* (Fitz-Gibbon *et al.*, 2002). Cependant, en analysant le faible taux de mutation chez *Sulfolobus acidocaldarius*, d'autres études suggèrent que cette faible fréquence d'erreur nécessite un système de réparation de type MMR fonctionnel (Grogan *et al.*, 2001). Une étude récente propose l'implication des protéines SSB chez *Sulfolobus* dans la détection des mésappariements *in vitro* (Cubeddu and White, 2005). Les protéines impliquées dans un mécanisme de type MMR où dans un système alternatif reste donc encore à découvrir grâce notamment à des approches génétiques.

3.2.4. Réparation des cassures double brin (*Double strand break repair DSBR*)

Les cassures double brin de l'ADN peuvent être une lésion très sérieuse pour les cellules. Elles peuvent être accidentelles mais elles peuvent aussi être programmées par la cellule lors de processus de recombinaison. Les mécanismes de recombinaison homologue (RH) et de réparation des cassures doubles brins (DSBR) peuvent être équivalents dans beaucoup de situations pour aboutir à la formation d'une jonction de Holliday (Liu and West, 2004) (figure 21). Ces systèmes de RH et DSBR peuvent avoir un rôle important pour réinitialiser la fourche de réplication lors d'un blocage dû à une cassure (McGlynn, 2004).

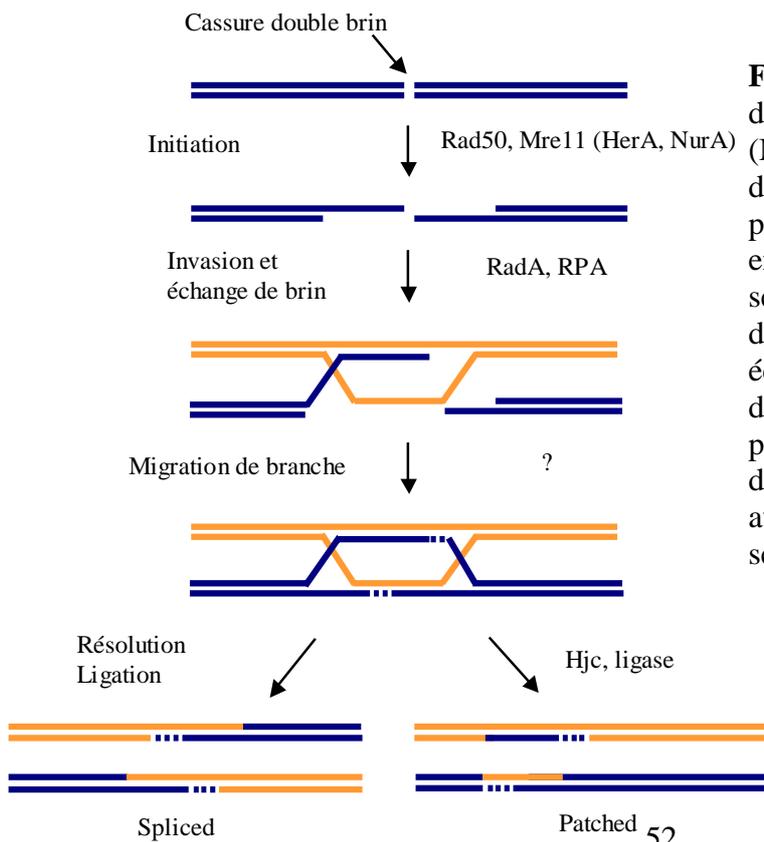


Figure 21 : Réparation des cassures double brin. Une exonucléase (Mre11/Rad50) génère des extrémités d'ADN simple brin. RadA/Rad51 polymérise sur l'ADN simple brin et envahit un ADN bicaténaire de séquence homologue, provoquant le déplacement du brin identique. Cet échange de brin entraîne la formation d'une jonction de Holliday. La polymérisation comble les brèches d'ADN et les jonctions sont résolues avec ou sans échange réciproque des séquences adjacentes.

Tous les génomes des archées séquencées présentent des homologues des protéines eucaryotes intervenant le DSBR comme RadA/Rad51, Mre11 et Rad50 (homologue respectivement à RecA, SbcD et SbcC des bactéries), ainsi que d'autres enzymes intervenant dans la résolution de la jonction de Holliday Hjc (Komori *et al.*, 1999; White, 2007). Les protéines Archées RadA, Mre11 et Rad50 présentent une structure très similaires avec leurs homologues eucaryotes.

Chez les eucaryotes, les protéines Mre11 et Rad50 forment un complexe avec Nbs1 et sont impliquées dans divers processus incluant la DSBR, la recombinaison non homologue (NHEJ), la détection de l'ADN endommagé, et les points de contrôle lors du cycle cellulaire (Assenmacher and Hopfner, 2004; D'Amours and Jackson, 2002). Mre11 est une ADN nucléase. Cette activité est portée par un motif phosphoestérase dans la partie N-terminale de la protéine (figure 22). L'activité nucléase a lieu dans le sens 3'-5' sur l'ADN simple brin comme sur l'ADN double brin. Mre11 présente également un site de liaison à l'ADN, qui est stimulé par Rad50 seul ou en complexe avec Nbs1. La protéine Rad50 possède à chacune de ses extrémités les motifs *walker A and B* permettant la liaison aux nucléotides. Ces motifs sont séparés par deux régions de structure non ordonnée (figure 22). Rad 50 est aussi capable de lier l'ATP grâce à un domaine ABC (ATP Binding Cassette), cette liaison induit un changement conformationnel générant deux motifs ATPases (figure 22). Ce changement conformationnel ATP dépendant peut stimuler l'activité nucléase de Mre11 (Hopfner *et al.*, 2000b; Hopfner *et al.*, 2001).

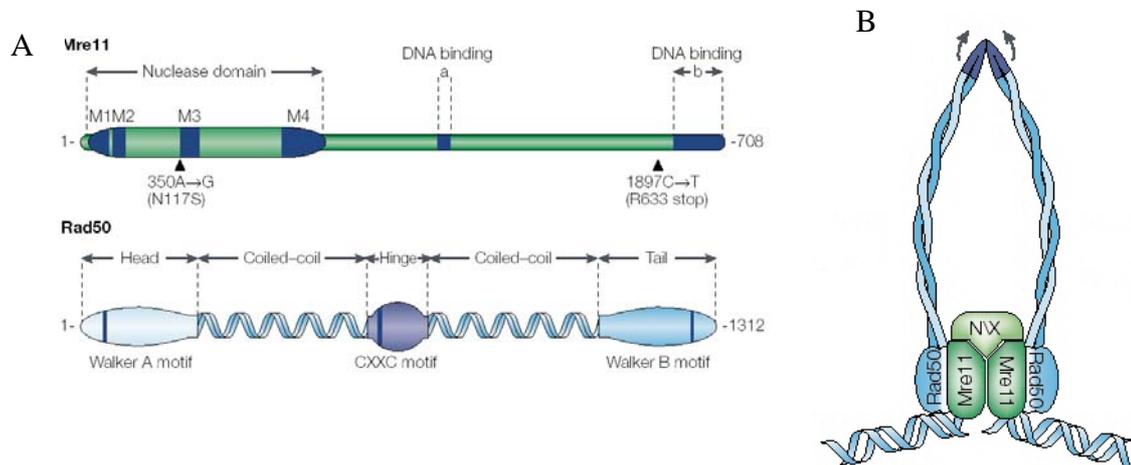


Figure 22 : A. Différents motifs sur les protéines Mre11 et Rad50 (D'Amours and Jackson, 2002). B. Architecture du complexe Mre11 chez l'eucaryote. Le complexe est composé des dimères Rad50 et Mre11 associé avec une seule protéine Nbs1/Xrs2. Le dimère Rad50 est flexible, la liaison de l'ATP induit un changement conformationnel (D'Amours and Jackson, 2002).

Dans les génomes de *Pyrococcus* sp., les protéines Rad50 et Mre11 sont codées par deux gènes organisés en opéron avec une autre protéine NurA, une nucléase, impliqué également dans le traitement des cassures double brin (Cohen *et al.*, 2003; Constantinesco *et al.*, 2002). Chez les archées hyperthermophiles, Mre11, Rad 50 et NurA sont cotranscrits avec un quatrième gène codant une hélicase HerA (Constantinesco *et al.*, 2004). Les protéines Rad50 et Mre11 de *Pyrococcus furiosus* présentent les mêmes activités catalytiques *in vitro* que leurs équivalents eucaryotes (Hopfner *et al.*, 2000a). Le complexe Mre11/Rad50 présente une activité endonucléase simple brin et une activité exonucléase double brin ATP dépendante.

* *
*

Au sein des cellules, la réparation de l'ADN est un phénomène étroitement associée à la réplication. D'une part, des erreurs peuvent apparaître lors de la réplication du génome, et d'autre part, les dommages sur l'ADN peuvent bloquer la fourche de réplication et entraîner la mise en place des mécanismes de réparation. Les protéines impliquées dans les mécanismes de réplication et de réparation de l'ADN peuvent être communes (tableau 4). Il y a ainsi une véritable intrication des complexes protéiques mis en jeu lors de ces processus.

Protéines	Rôle dans la réparation de l'ADN
Pol δ / Pol ϵ	Comblent une brèche ADN lors du BER, NER, DSB, MMR
PCNA	Facteur de processivité de Pol δ Etape d'excision dans le LP-BER Stabilisation du complexe NER Reconnaissance du complexe MMR
RF-C	Chargement du PCNA
RP-A	Reconnaissance de lésions dans le NER Stabilisation du complexe DSB Modulation de Mu α et MutL α lors du MMR
Fen1	Clivage de la structure flap dans le BER Rôle dans le DSB
ADN ligase I	Ligation des produits dans le NER Médiateur dans le BER

Tableau 4 : rôle des protéines caractéristiques de la réplication de l'ADN chez les eucaryotes intervenant dans des processus de réparation de l'ADN. D'après (Touaille and Hubscher, 2004).

* *
*

IV. Dis-moi avec qui tu interagis, je te dirais qui tu es...

4.1. Interactome : De l'interaction aux réseaux

L'ensemble des protéines d'une cellule qui réalise le programme génétique, à un instant donné, constitue le protéome (Kahn, 1995). La structure tridimensionnelle d'une protéine est à la base des propriétés particulières de cette protéine et lui permet d'assumer sa fonction biologique. La liaison d'une protéine avec un ligand spécifique a également un grand rôle dans la fonction biologique de la protéine. L'interaction d'une protéine avec une autre est un processus dynamique, mettant en jeu à l'interface protéine-protéine des phénomènes de complémentarité, la création de liaisons hydrogène, d'interactions électrostatiques et la participation de domaines hydrophobes. Les protéines peuvent s'associer deux par deux ou former de larges complexes multiprotéiques et ces interactions permettent de contrôler divers processus cellulaires. Les interactions protéines-protéines peuvent modifier leur propriétés cinétiques, la liaison avec un substrat, leur conformation dans l'espace. Les interactions peuvent être stables, ou bien transitoires. Les protéines peuvent s'associer ou se dissocier *in vivo* selon le contexte afin d'exécuter un processus cellulaire. La richesse de ces interactions, leurs dynamiques dans l'espace et dans le temps permettent en partie d'expliquer la diversité et le fonctionnement du vivant.

Une des stratégies de la protéomique est de cartographier les interactions protéine-protéine, représentées sous forme de réseau d'interaction. Il est possible de parler d'interactome. Ces représentations permettent de visualiser les relations existantes entre les protéines.

La description de la dynamique des interactions protéines-protéines au sein de la cellule nous permettra de mieux comprendre son fonctionnement. L'analyse de ces données peut par exemple permettre de proposer des fonctions pour des protéines dont on ignorait le rôle, mais aussi révéler des rôles inattendus pour des protéines que l'on pensait connaître. Certaines de ces protéines peuvent avoir de nombreuses connections, elles sont une plateforme d'interaction (nœud ou *hub*) et peuvent être ainsi au cœur d'un réseau fonctionnel. Ces modules regroupent des protéines impliquées dans des processus cellulaires particuliers (Hartwell *et al.*, 1999). Ainsi, les protéines impliquées dans des processus communs ont tendance à être en interaction directe les unes avec les autres.

Plusieurs stratégies permettant d'identifier des partenaires protéiques interagissant avec une protéine d'intérêt ont été décrites au cours des dernières années. La technique du double hybride est très employée pour mettre en évidence des interactions entre protéines. En utilisant cette technique, des travaux sur la levure *Saccharomyces cerevisiae*, initiés par deux consortia de laboratoires ont permis l'analyse de nombreux complexes multi-protéiques et la réalisation de réseau d'interaction chez cet organisme (Gavin *et al.*, 2002; Ho *et al.*, 2002) (figure 24). Une étude récente présente une analyse des interactions protéines-protéines chez l'euryarchée *Pyrococcus horikoshii* (Usui *et al.*, 2005). Grâce à une technique de double-

hybride, à partir de 960 protéines solubles, les auteurs ont mis en évidence 107 interactions protéine-protéines fiables dont 56 hétéro-interactions et 51 auto-interactions. Quelques unes de ces interactions ont été validées par des techniques de *pull-down*. Lors de cette approche, 13 protéines avec une fonction inconnue étaient capables d'interagir avec des protéines annotées. Cette information a été utile pour la prédiction des fonctions et l'annotation de ces 13 protéines. D'autres approches génétiques comme le TAP-tag (*tandem affinity purification*) sont également très employées. Cette stratégie permet de purifier *in vivo*, à partir d'une protéine étiquetée, des complexes protéiques. Combinée à la spectrométrie de masse, les protéines interagissantes ainsi récupérées peuvent être identifiées (Rigaut *et al.*, 1999).

Sans être exhaustif, nous pourrions citer d'autres exemples de techniques permettant de détecter des interactions protéine-protéine : l'analyse en spectrométrie de masse de complexes purifiés, la capture de partenaire grâce à des techniques de chromatographie d'affinité, de co-immunoprécipitation ou de *pull-down*. De nombreux articles décrivent ces méthodes (Cho *et al.*, 2004; Drewes and Bouwmeester, 2003; Phizicky and Fields, 1995; von Mering *et al.*, 2002). Des prédictions d'interaction protéines-protéines *in silico* en analysant les génomes sont également employées. Par exemple, chez les procaryotes, les gènes organisés en opéron codent souvent des protéines interagissant entre elles (analyse du contexte génomique). L'analyse des profils phylogénétiques peut aussi permettre d'inférer des interactions entre protéines d'un organisme à un autre. Les protéines homologues présentent en effet une séquence, une structure et possiblement une fonction similaire. L'analyse de la co-présence ou la co-absence de gènes dans des génomes entièrement séquencés est un indicateur d'une interaction potentielle entre les produits de ces gènes (Pellegrini *et al.*, 1999). Par ailleurs, l'analyse des bases de données expérimentales d'interactions entre protéines permet de retrouver des séquences signatures de l'interaction (figure 23). Ces séquences peuvent être recherchées par bioinformatique pour prédire des partenaires interactants putatifs dans la cellule (Sprinzak and Margalit, 2001).

4.2. Réseau d'interaction au niveau de la réplication et la maintenance de l'ADN

De nombreux travaux ont permis de progresser dans la connaissance des interactions entre protéines intervenant dans la réplication, entre protéines de la réplication et celles de la réparation, et cela à plus ou moins grande échelle.

Par exemple, une étude a été entreprise à grande échelle, afin d'identifier des protéines s'associant physiquement à des protéines de la réplication connues chez la bactérie Gram positif *Bacillus subtilis*, par l'intermédiaire de screening grâce à la technique de double hybride chez la levure (Noirot-Gros *et al.*, 2002). Afin d'éviter l'introduction de faux positifs engendrés par cette approche, chaque interaction identifiée a été testée expérimentalement. Le réseau obtenu était ainsi composé de 91 interactions spécifiques faisant intervenir 69 protéines (figure 25). Cette étude confirme le lien entre la réplication de l'ADN avec les processus de réparation et de recombinaison de l'ADN.

Concernant les eucaryotes, une étude génétique remarquable a permis d'élucider un réseau de gènes permettant d'assurer la stabilité du génome et la fidélité de la réplication chez la levure *Saccharomyces cerevisiae* (Budd *et al.*, 2005). Cette approche a de plus permis d'améliorer les connaissances sur la protéine essentielle Dna2, une hélicase/nucléase intervenant dans la réplication, mais aussi dans les processus de maturation des fragments d'Okazaki, dans la réparation de l'ADN et dans la dynamique de la chromatine.

Plusieurs études intéressantes ont cherchées à identifier des protéines capable d'interagir avec le PCNA et intervenant dans la réplication de l'ADN ou dans le cycle cellulaire. L'une d'elle (Loor *et al.*, 1997), chez l'eucaryote, a permis la purification de plusieurs protéines par chromatographie d'affinité en utilisant le PCNA comme appât fixé sur une colonne de sépharose. Après élution, les ADN polymérase Pol α , Pol δ , Pol ϵ , le PCNA, le RFC, l'ADN ligase I, le RPA, les topoisomérases I et II ainsi que des protéines du cycle cellulaire ont été mises en évidence grâce à des anticorps spécifiques par Western blot. Une autre approche comparable avait pour but l'identification des protéines se liant au PCNA dans un extrait de cellules de lignées humaines (Ohta *et al.*, 2002). Cette étude a couplé la purification par chromatographie liquide, le PCNA étant fixé sur une résine, et l'identification des partenaires capturés, par spectrométrie de masse (tableau 5). Ce travail a permis de mettre en évidence des interactions déjà connues, mais aussi d'identifier un nouveau partenaire du PCNA, la protéine CHL12, impliqué dans le contrôle du cycle cellulaire et la cohésion des chromosomes.

Plus récemment, une colonne d'affinité a été utilisée pour la recherche de partenaires interactants de MGMT (O6-méthyl-guanine-DNA-méthyltransferase), une protéine intervenant dans la réparation de l'ADN, dans un extrait de lignée cellulaire cancéreuse humaine HT29 (Niture *et al.*, 2005). Les protéines capables de se lier de façon spécifique à MGMT fixée sur la colonne ont été identifiées à la fois par spectrométrie de masse MS/MS et western blot. Mcm2, PCNA, Orc1 et Pol δ , des protéines intervenant dans la régulation du cycle cellulaire et des histones ont été ainsi mis en évidence comme partenaire de MGMT.

Chez les *Archaea*, un travail a permis de caractériser un réseau d'interaction avec les protéines de la réplication de l'ADN chez *Archaeoglobus fulgidus* (Mutz *et al.*, 2002). A l'aide d'une technique de double hybride chez la levure, l'approche a mis en évidence un réseau de 30 protéines potentiellement impliquées dans la réplication ou la réparation de l'ADN chez cet organisme (tableau 6 et figure 26). Cette approche a permis par exemple de retrouver les protéines liant le PCNA, comme les ADN polymérase et le facteur de chargement du PCNA. Des protéines capables de lier l'ADN ont également été identifiées, comme le RP-A, ainsi que des protéines impliquées dans le métabolisme de l'ADN comme des ADN endonucléases et des ADN exonucléases, des protéines impliquées dans le métabolisme des nucléotides.

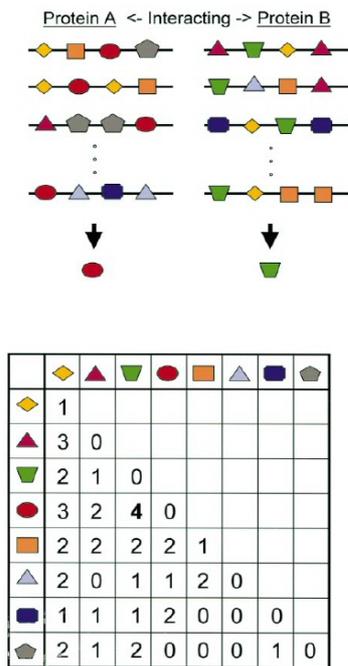


Figure 23 : représentation schématique de l'analyse permettant de détecter des séquences signatures au sein des protéines interagissantes. Dans le schéma du haut, chaque rangée correspond aux séquences de protéines qui interagissent expérimentalement avec les protéines A et B. chaque séquence est caractérisée par des signatures, schématisé ici par des formes colorées. Le schéma du bas représente la table de combinaison des signatures permettant d'identifier la séquence signature de l'interaction pour chacune des protéines A et B. (Sprinzak and Margalit, 2001)

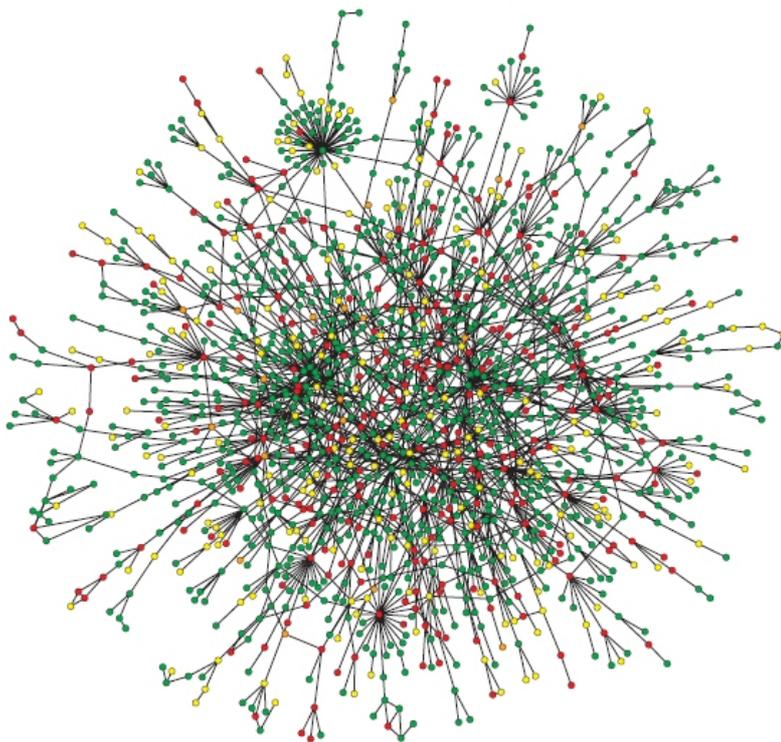


Figure 24 : Exemple de réseau d'interaction de protéines obtenue par analyse systématique des complexes protéiques, chez la levure *S. cerevisiae*. D'après (Jeong *et al.*, 2001). Certaines zones du réseau apparaissent plus denses que d'autres. Une protéine peut avoir un grand nombre de partenaires ou la densité des connections entre les voisins d'une protéine peut être importante. Les protéines avec une grande densité de connections peuvent appartenir à un même module fonctionnel.

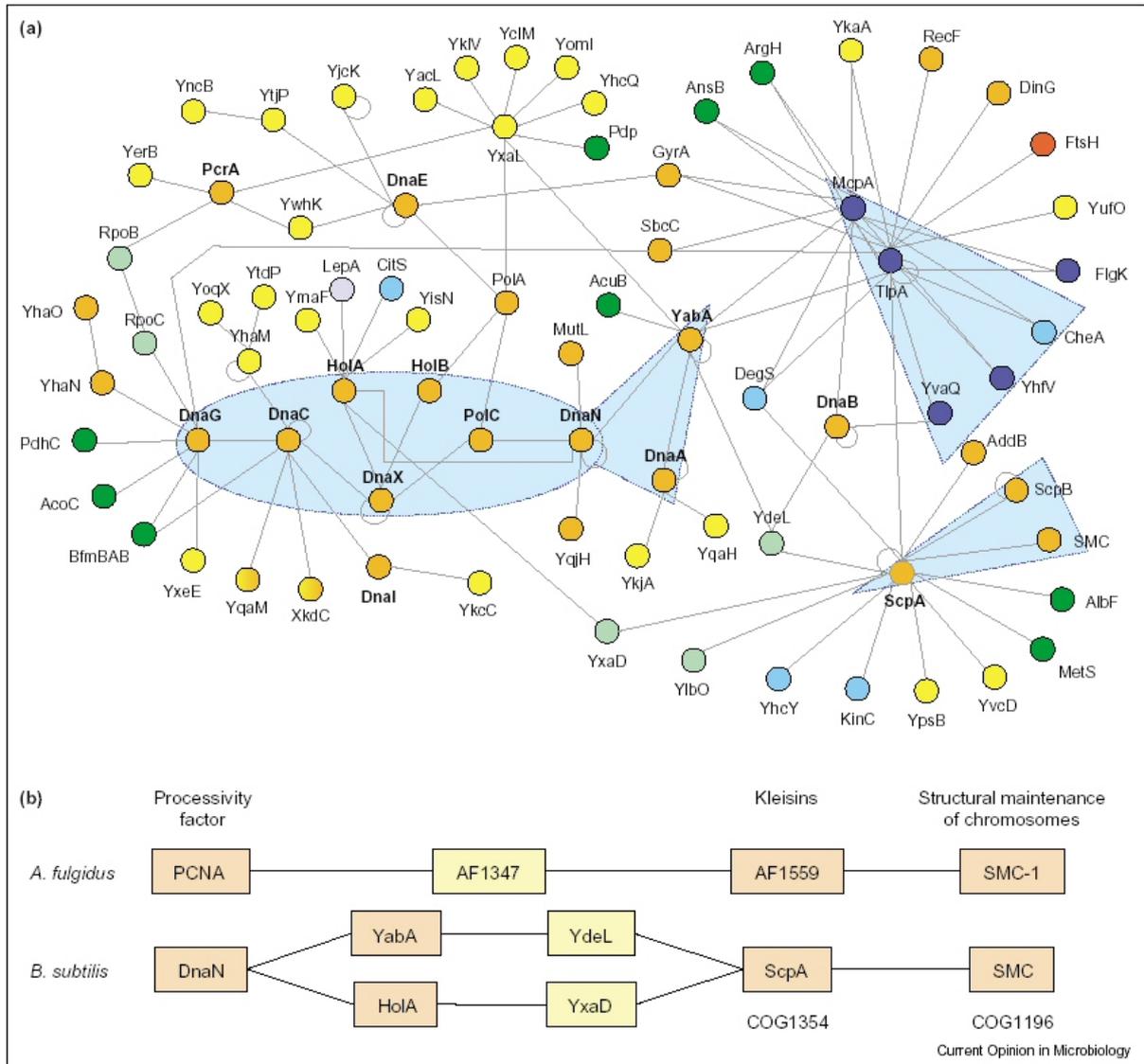


Figure 25 : Réseau d'interaction de protéines montrant le lien entre la réplication de l'ADN et la dynamique du chromosome (recombinaison et réparation) D'après (Noirot and Noirot-Gros, 2004). (a) Réseau d'interaction chez *Bacillus subtilis* composé de 112 interactions spécifiques entre 78 protéines fusionnant deux travaux : (Dervyn *et al.*, 2004; Noirot-Gros *et al.*, 2002). La couleur des protéines indiquent leur catégorie fonctionnelle. Les protéines oranges interviennent dans les processus de réplication, réparation et recombinaison de l'ADN. Les protéines bleues foncées interviennent dans la mobilité et la chimiotaxie, les bleues claires dans la transduction du signal. Les protéines vertes foncées interviennent dans le métabolisme, les vertes claires dans la transcription, les violettes dans la synthèse des protéines, les rouges dans le stress et la protéolyse, les jaunes n'ont pas de fonctions connues. (b) Alignement du réseau d'interaction de *Bacillus subtilis* avec celui de *Archaeoglobus fulgidus* (Motz *et al.*, 2002).

Gel slice number	gi	Mass	Total score	Peptides matched	Annotation
8	1709123	152858	214	11	DNA mismatch repair protein MSH6 (MUTS-ALPHA 160-kDa subunit)
11	4505933	123635	148	8	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
	4557719	101736	77	5	DNA ligase I
12	4505933	123635	467	21	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
	4557719	101736	326	13	DNA ligase I
13	4505933	123635	1525	57	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
	4557719	101736	176	9	DNA ligase I
14	4505933	123635	445	16	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
	4557719	101736	385	16	DNA ligase I
	13623351	108738	338	13	Protein for IMAGE:3959495; CHL 12
15	13623351	108738	468	17	Protein for IMAGE:3959495; CHL 12
	4505933	123635	99	4	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
16	4505933	123635	495	22	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
17	4505933	123635	488	21	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
18	4557761	104743	281	10	mutS homolog 2; mutS (E. coli) homolog 2
19	1285967	68462	108	7	(AK019460) putative
21	4505933	123635	110	6	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
23	4505933	123635	161	9	Polymerase (DNA-directed), delta 1, catalytic subunit (125 kDa)
24	3123056	54079	131	7	KIAA0039; delta 3 (66 kDa)
25	3123056	54079	615	33	KIAA0039; delta 3 (66 kDa)
26	3123056	54079	229	10	KIAA0039; delta 3 (66 kDa)
	5453840	72371	78	4	(NM_006386) DEAD/H box polypeptide 17
31	7657383	54252	100	6	Non-Pou domain-containing octamer binding protein
33	5453924	51289	137	4	Polymerase (DNA-directed), delta 2, regulatory subunit (50 kDa)
34	5453924	51289	568	24	Polymerase (DNA-directed), delta 2, regulatory subunit (50 kDa)
35	4758356	42593	239	9	Flap structure-specific endonuclease 1; FEN1
	5453924	51289	147	4	Polymerase (DNA-directed), delta 2, regulatory subunit (50 kDa)
36	4758356	42593	380	22	Flap structure-specific endonuclease 1; FEN1
37	13129096	44844	227	9	Hypothetical protein MGC5528
	4758356	42593	195	7	Flap structure-specific endonuclease 1; FEN1
	13278939	41190	132	4	(BC004223) Unknown (protein for IMAGE:3533309)
38	4758356	42593	36	1	Flap structure-specific endonuclease 1; FEN1
40	14741989	35244	259	9	Replication factor C (activator 1) 2 (40 kDa)
41	4506489	40556	256	13	Replication factor C (activator 1) 3 (38 kDa)
	6224979	33924	155	6	Uracil-DNA glycosylase
	11990420	154157	66	4	(AB014772) MOP-3
42	6677723	38497	177	7	Replication factor C (activator 1) 5 (36.5 kDa)
43	4506491	39682	203	6	Replication factor C (activator 1) 4 (37 kDa)
44	4506491	39682	218	11	Replication factor C (activator 1) 4 (37 kDa)
47	4758356	42593	207	12	Flap structure-specific endonuclease 1; FEN1
50	4758356	42593	212	9	Flap structure-specific endonuclease 1; FEN1

Tableau 5 : Exemple de données primaires correspondant aux protéines identifiées comme liées au PCNA sur colonne de chromatographie d'affinité à partir d'un lysat cellulaire de lignée humaine. Les protéines sont identifiées en spectrométrie de masse. D'après (Ohta *et al.*, 2002).

Gene	Function
1. AF0111	Related to inosine-5'-monophosphate dehydrogenase (IMPDH)*
2. AF0130	Putative histone deacetylase, complexes with DP-2 in <i>P. furiosus</i> (HDAC)*
3. AF0335	PCNA, sliding clamp to tether polymerases to DNA
4. AF0382	Replication protein A, 26-kD subunit single-stranded DNA binding complex (RPA-26)*
5. AF0472	Unknown function
6. AF0497	DNA polymerase B1 (polB)
7. AF0699	Putative ssDNA-specific exonuclease, homology to AF0735, RecJ related, (RecJ-2)*
8. AF0735	recJ related, homology to AF0699, complexes with DP-2, (RecJ-1)*
9. AF0780	Replication protein A, 36-kD subunit single-stranded DNA binding complex (RPA-36)*
10. AF1045	Methyl-accepting chemotaxis protein (tlpC-2)
11. AF1194	Unknown, complexes with DP-2 in <i>P. furiosus</i>
12. AF1195	Replication factor C, 53-kD subunit (RFC-53)
13. AF1272	Phosphoribosylaminoimidazolesuccinocarboxamide synthase (purC)
14. AF1347	Unknown function, homology to AF1590*
15. AF1558	Chromosome segregation protein (smc1), complexes with DP-2 (SMC-1)
16. AF1559	Unknown function, putative same operon as 1558
17. AF1650	Tungsten formylmethanofuran dehydrogenase (fwdB-1), complexes with DP-2 in <i>P. furiosus</i>
18. AF1692	Endonuclease III (Endo-III)
19. AF1722	DNA polymerase, archaeal type II, large subunit, (DP-2)
20. AF1790	DNA polymerase, archaeal type II, small subunit, (DP-1)
21. AF2060	Replication factor C, 35-kD subunit (RFC-35)
22. AF2215	Methyl-malonyl-CoA-mutase (mmCoA-mut)
23. AF0103	Tentative homology to histidine kinases (HisK-1)*
24. AF0225	Branched-chain amino acid ABC transporter, permease protein (braE-1)
25. AF0264	DNA repair protein RAD2 (RAD2)
26. AF0621	Ribonuclease HIII (RNaseHIII)
27. AF1332	Histidine kinase (HisK-2)
28. AF1452	Histidine kinase (HisK-3)
29. AF1741	Orotate phosphoribosyl transferase (pyrE)
30. AF2111	Function unknown
31.	Transposases (tp1, tp2)

Tableau 6 : Annotation des protéines identifiées dans le réseau de *Archaeoglobus fulgidus* (Motz *et al.*, 2002).

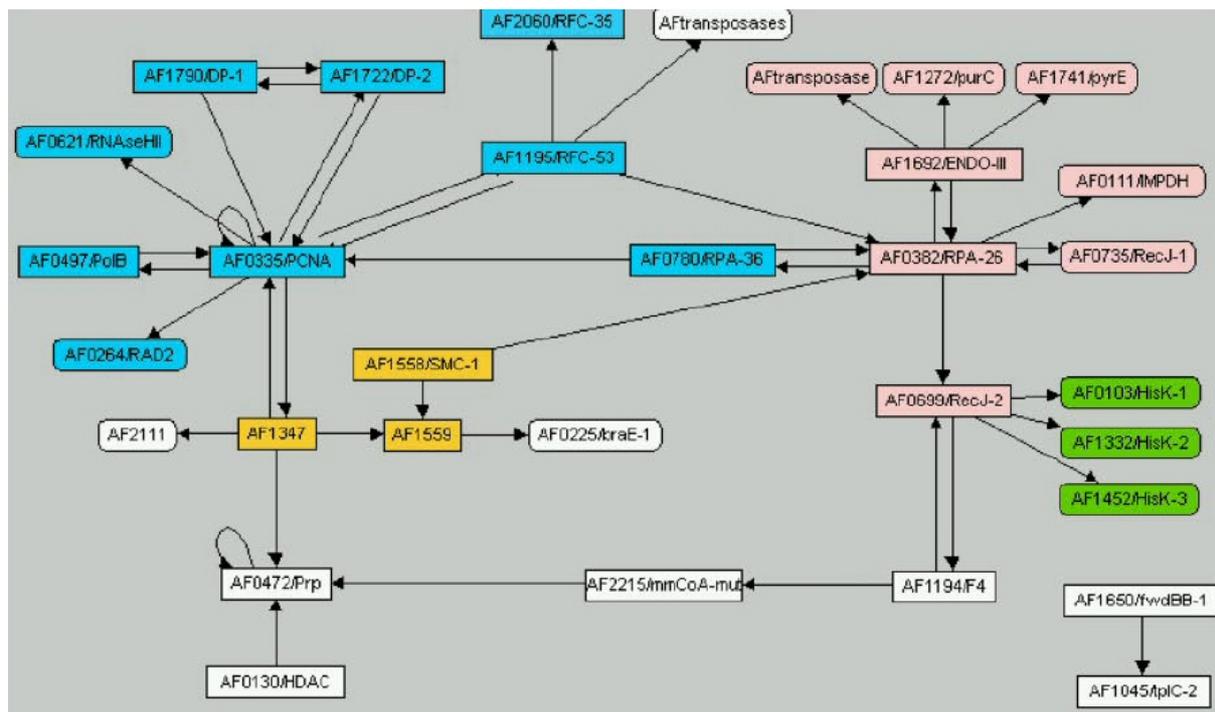


Figure 26 : Réseau d'interaction identifié chez *Archaeoglobus fulgidus*. D'après (Motz *et al.*, 2002). Les doubles flèches indiquent des interactions réciproques. Le groupe de protéines centrées autour du PCNA et intervenant plutôt dans la réplication de l'ADN est indiqué en bleu. Les protéines colorées en rose interviennent plutôt au niveau de la réparation de l'ADN.

IV. Présentation de l'étude

La comparaison des génomes de divers organismes a permis de montrer que les protéines impliquées dans des processus métaboliques sont homologues chez les Archées et Bactéries, tandis que les Archées et les Eucaryotes présentent plus de similitudes en ce qui concerne les complexes multi-moléculaires impliqués dans les mécanismes informationnels (réplication de l'ADN, transcription et traduction). Cette observation a conduit à poser l'hypothèse que les Archées pourraient s'avérer être un excellent modèle pour comprendre les mécanismes enzymologiques lors de la réplication de l'ADN chez les Eucaryotes. L'étude du réplisome des Archées hyperthermophiles présente de plus un intérêt sur le plan biotechnologique ; les propriétés des enzymes thermostables peuvent en effet être utilisées pour améliorer les performances de réplication de l'ADN *in vitro* (PCR). Les conditions environnementales du biotope des archées hyperthermophiles pourraient entraîner l'apparition de lésions au sein de l'ADN. L'ADN des hyperthermophiles est en effet exposé à des températures élevées susceptible d'accroître le taux de réaction de décomposition (dépurination et déamination) (Lindahl, 1993). Cependant, des travaux récents suggèrent que le taux de mutation des archées hyperthermophiles ne semble pas plus élevé que celui des autres micro-organismes. Cependant, la compréhension des mécanismes et protéines impliqués dans la réparation de l'ADN des archées reste encore limitée (Kelman and White, 2005), d'autant plus que certaines voies et protéines de la réparation, présentes dans les autres domaines du vivant, semblent manquer chez les archées. Ces observations soulèvent des questions fondamentales. Comment les hyperthermophiles, et particulièrement *Pyrococcus abyssi*, parviennent à dupliquer leur génome tout en maintenant leur intégrité génomique, dans ces conditions particulières ?

Les travaux précédemment réalisés au laboratoire de Microbiologie des Environnements Extrêmes ont permis de caractériser plusieurs protéines participant à la réplication de l'ADN chez *P. abyssi* : les protéines appartenant aux familles B et D des ADN polymérases impliquées dans la phase d'élongation (Gueguen *et al.*, 2001), le facteur de processivité PCNA et son facteur de chargement le RF-C (Henneke *et al.*, 2002), l'ADN primase (Le Breton *et al.*, 2007). Nous avons également à notre disposition les enzymes impliquées dans l'initiation et la maturation de la synthèse des fragments d'Okazaki : l'endonucléase Fen1, la ligase 1 et RNase HII (équipe de Hannu Myllykallio). Ces protéines ne constituent probablement qu'une partie du réplisome. Par ailleurs, les récentes approches, en utilisant des organismes modèles eucaryotes, ont montré que les voies métaboliques de la réplication de l'ADN s'intriquent fonctionnellement avec celles de la réparation de l'ADN. La recherche et la caractérisation de protéines complémentaires, grâce aux outils de la protéomique (interaction protéine-protéine et spectrométrie de masse) nous permettrait d'approfondir nos connaissances sur les mécanismes moléculaires mis en jeu au niveau de la maintenance génomique chez les archées.

L'étude présentée dans ce manuscrit présente la caractérisation d'un réseau d'interactions au niveau de la maintenance génomique de *P. abyssi* et propose une fonction physiologique pour certains des nouveaux complexes mis en évidence.

Résultats et Discussion

I. Le PCNA acteur central de la maintenance génomique

Comme décrit dans l'introduction, le PCNA interagit avec de multiples partenaires, il est connu pour avoir un rôle clé dans la maintenance du génome en coordonnant l'action de ces nombreux partenaires dans les mécanismes de réplication, de réparation et de recombinaison (Maga and Hubscher, 2003). Le PCNA recombinant a été utilisé comme appât pour capturer des partenaires dans l'extrait cellulaire total de *P. abyssi*. Sur le plan méthodologique, le PCNA est fixé, par l'intermédiaire d'une queue histidine, sur des billes magnétiques. Après incubation avec l'extrait cellulaire et lavage extensif, les partenaires potentiels sont récupérés. La figure 27 représente le type de résolution que l'on peut obtenir après élution et migration du complexe en électrophorèse et coloration au bleu de Coomassie. Les partenaires identifiés par spectrométrie de masse sont référencés dans le tableau 7. Le témoin négatif permet de vérifier la spécificité des interactions. L'identification d'une protéine par empreinte peptidique est considérée comme fiable lorsque le pourcentage de couverture dépasse 20%, et que le score mascot est supérieur à 66, cela en fonction de la base de données NCBIInr interrogée pour cette étude.

Cinq partenaires ont été identifiés pour le PCNA selon cette approche : l'ADN Polymérase B (*Pab1128*), la petite sous-unité du RFC (*Pab0068*), la grande sous-unité du RFC (*Pab0069*), Mre11 (*Pab0811*) et Rad50 (*Pab0812*).

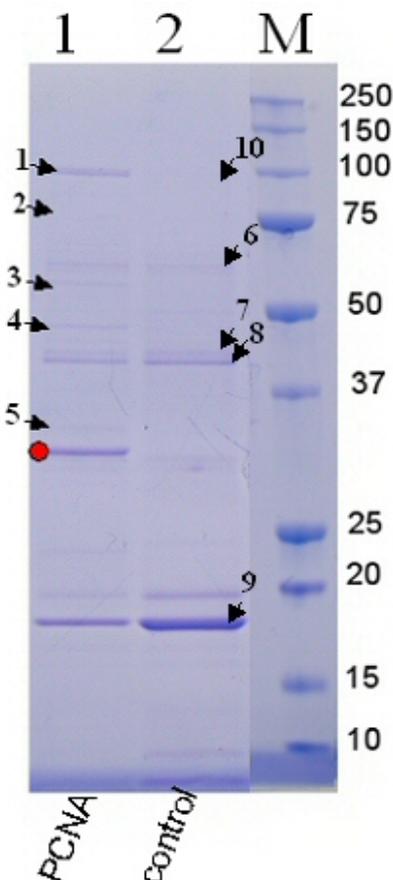


Figure 27 : identification de partenaires potentiels de différents du PCNA. (piste 1), 20 µg de PCNA ont été fixé sur des billes magnétiques par l'intermédiaire d'une queue histidine. Après incubation avec le lysat cellulaire, et plusieurs étapes de lavages, les complexes ont été séparés par électrophorèse et visualisés après coloration au bleu de Coomassie. Le cercle rouge correspond au PCNA, les flèches noires correspondent aux partenaires potentiels, identifiés par spectrométrie de masse, et référencés dans le tableau x. La spécificité de ces interactions est évaluée par rapport au contrôle négatif (billes, sans appât fixé, incubées avec l'extrait cellulaire ; piste 2).

Appâts	Partenaires interactants						
	N° spot	Protéines	Accession number	Masse Moléculaire	Score Total	Peptides matched	% Coverage
No (negatif control)	6	Fe-S oxydoreductase	gi 5457532	65 953	153	21/58	39
	7	Hypothetical protein	gi 5458650	45 316	65	12/63	29
	8	Hypothetical protein	gi 5459109	44 798	77	20/100	38
	9	MoaB-2 molybdenum cofactor biosynthesis	gi 5458945	18 359	139	11/46	73
	10	Acetyltransferase (GNAT) family	gi 5458694	94 666	95	10/18	15
DNA polymerase sliding clamp	4	Mre11 DNA repair prot	gi 5458642	49 085	208	29/100	57
PCNA gi 5458797	1	Rad50 purine ntpase	gi 5458643	104 021	167	21/36	27
	2	Pol I DNA Polymerase	gi 5459139	89 668	127	23/71	33
	5	Replication factor C small subunit	gi 5457543	37 667	121	23/47	14
	3	Replication factor C large subunit	gi 5457544	55 265	122	12/23	24

Tableau 7 : tableau récapitulatif des partenaires potentiels identifiés par spectrométrie de masse pour le PCNA, et des protéines identifiées dans le contrôle négatif. Les protéines retrouvées dans le contrôle négatif ne sont pas incluses dans la liste des partenaires potentiels.

L'interaction du PCNA avec la polymérase B et le RFC était connue et a été décrite chez *P. abyssi* auparavant (Henneke *et al.*, 2002; Henneke *et al.*, 2005). L'interaction du PCNA avec la polymérase B de *P. abyssi* a déjà été mise en évidence par des expériences de résonance plasmonique de surface (SPR). Cette interaction avait alors été décrite comme étant dépendante de l'ADN (Henneke *et al.*, 2005). L'ADN présent dans les lysats cellulaires pourrait avoir une influence au niveau des résultats d'interaction mis en évidence. Toutefois, l'ADN est l'acteur central des mécanismes de réplication et de réparation étudiée. Sa présence n'est donc pas un obstacle aux interprétations présentées dans ce travail. Le PCNA, chargé par le RFC, est le facteur de processivité de l'ADN polymérase B. Ces trois protéines pourraient faire partie du complexe de réplication minimal actif, équivalent au réplisome eucaryote comprenant le PCNA, l'ADN polymérase δ et le RFC. Le fait de retrouver ces interactions bien caractérisées contribue donc à valider notre approche et tend à légitimer les nouveaux partenaires potentiels identifiés.

D'autre part, en utilisant cette fois une approche de *pull-down* couplée à l'immuno-révélation, j'ai pu mettre en évidence des interactions entre le PCNA et les enzymes de la maturation des fragments d'Okazaki (figure 28).

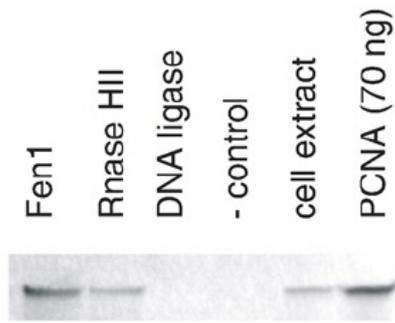


Figure 28 : visualisation de l’interaction de Fen1 et RNaseHII avec le PCNA dans le lysat cellulaire, après *pull-down*. 20 µg de Fen1, RNaseHII et ADN ligase sont fixés sur les billes magnétiques et sont incubés avec le lysat cellulaire. Le contrôle négatif correspond aux billes magnétiques seules. Le PCNA capturée est visualisé après western-blot, grâce à des anticorps anti-PCNA.

La figure 29 représente l’ensemble des partenaires du PCNA identifié au cours de ce travail.

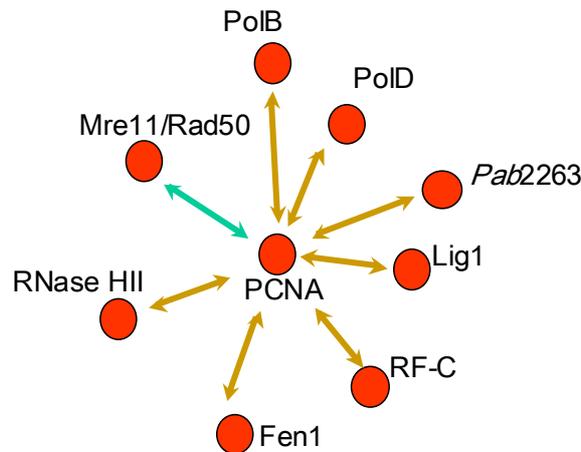


Figure 29 : réseau d’interaction montrant les protéines partenaires du PCNA identifiées au cours de ce travail, chez *P. abyssi*. Toutes ces interactions ont été observées lors des différentes expériences de *pull-down* dans l’extrait cellulaire, sauf pour les interactions PCNA-Polymérase D et PCNA-Ligase 1, identifiées seulement par SPR. Les flèches marrons indiquent que l’interaction a été validée *in vitro*, entre protéines recombinantes, grâce à la technique de résonance plasmonique de surface (SPR).

Ces résultats démontrent à nouveau le caractère central du PCNA dans les mécanismes de réplication et de réparation de l’ADN (Maga and Hubscher, 2003). Pour aller plus loin dans l’étude de la coordination du PCNA avec ses partenaires, j’ai, en collaboration avec l’équipe de H. Myllykallio, contribué à l’étude de son interaction avec les protéines Fen1, ADN ligase 1 et RNase HII. D’autre part, j’ai également participé à la caractérisation entre le PCNA et une nouvelle nucléase Pab2263. Enfin, pour cette partie, j’ai produit le complexe Mre11/Rad50 afin de confirmer l’interaction avec le PCNA et de rechercher d’autres partenaires de ce complexe.

1.1. Fen1, RNase HII et l'ADN Ligase 1 interagissent au niveau d'un site unique de fixation sur le PCNA

Lors des processus de réplication et de réparation de l'ADN, de nombreuses protéines peuvent se lier ou se dissocier du PCNA de manière hautement spécifique.

La ribonucléase RNase HII et l'endonucléase structure spécifique Fen1 sont des protéines qui participent à la maturation des fragments d'Okazaki, lors de la réplication de l'ADN. Au cours de la synthèse du brin retardé, l'amorce ARN doit être enlevée afin de rendre possible la ligation entre les deux fragments d'Okazaki adjacent, par la Ligase. La compréhension de la coordination et de l'action des protéines PCNA, Fen1, ADN ligase et RNase HII lors de la maturation des fragments d'Okazaki n'est pas encore clairement déterminée (Kao and Bambara, 2003). Les interactions entre les protéines PCNA, *PabFen1*, DNA Ligase 1 et RNase HII ont été recherchées par la technique de *pull-down* dans l'extrait cellulaire total de *P. abyssi*. Ces interactions ont également été étudiées *in vitro* grâce à la technique de la résonance plasmonique de surface. Des expériences de compétition en résonance plasmonique de surface ont révélé que les différents partenaires d'interaction du PCNA sont en compétition au niveau d'un site unique de fixation. Dans cette étude, il est également montré que l'activité de RNase HII est diminuée en présence du PCNA. Le PCNA jouerait un rôle dans la coordination des activités physiologiques de ces partenaires.

Ces observations, couplées à des expériences de génétique chez l'Archaea *Haloferax volcanii* (les outils génétiques n'étant pas développés chez *P. abyssi*) ont permis de proposer des rôles nouveaux pour les protéines Fen1 et RNaseHII dans des mécanismes de réparation de l'ADN.

La publication présentée ici expose également une nouvelle approche de protéomique permettant d'identifier des partenaires du PCNA. Une librairie de petits peptides synthétiques immobilisés sur un support a été criblée afin d'identifier leur capacité de liaison au PCNA de *P. abyssi* et au PCNA humain. Les peptides interagissant avec le PCNA ont ensuite été recherchés dans les séquences de protéines de *P. abyssi*. Ces séquences ont été retrouvées dans le RF-C, Fen1, l'ADN polymérase B, l'ADN ligase mais aussi dans la ribonucléase RNaseHII et une nouvelle protéine, *Pab2263*.



Available online at www.sciencedirect.com




A Novel Proteomic Approach Identifies New Interaction Partners for Proliferating Cell Nuclear Antigen

Laurence Meslet-Cladière^{1,2,†}, Cédric Norais^{1,2,†}, Joëlle Kuhn^{1,2}
 Julien Briffotax³, Jerry W. Sloostra⁴, Elena Ferrari⁵
 Ulrich Hübscher⁵, Didier Flament^{3,‡} and Hannu Myllykallio^{1,2,*}

¹CNRS UMR 8621
 U. Paris-Sud, 91405 Orsay
 Cedex, France

²Institute de Génétique et
 Microbiologie, U. Paris-Sud
 91405 Orsay, Cedex, France

³UMR 6197, Laboratoire
 de Microbiologie des
 Environnements Extrêmes
 IFREMER Brest, 29280
 Plouzane, France

⁴Pepsan Systems, 8219 PH
 Lelystad, The Netherlands

⁵Institute of Veterinary
 Biochemistry and Molecular
 Biology, University of
 Zurich-Irchel
 Winterthurerstrasse 190
 CH 8057 Zurich, Switzerland

*Corresponding author

During DNA replication and repair, many proteins bind to and dissociate in a highly specific and ordered manner from proliferating cell nuclear antigen (PCNA). We describe a combined approach of *in silico* searches at the genome level and combinatorial peptide synthesis to investigate the binding properties of hundreds of short PCNA-interacting peptides (PIP-peptides) to archaeal and eukaryal PCNAs. Biological relevance of our combined approach was demonstrated by identification of an inactive complex of *Pyrococcus abyssi* ribonuclease HIII with PCNA. Furthermore we show that PIP-peptides interact with PCNA largely in a sequence independent manner. Our experimental approach also identified many so far unidentified PCNA interacting peptides in a number of human proteins.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: PCNA; combinatorial peptide synthesis; ribonuclease HIII

Introduction

Proliferating cell nuclear antigen (replication clamp PCNA) was originally identified as an antigen present in the nuclei of dividing cells that reacted with an autoantibody in the sera of patients with lupus erythematosus.¹ Later studies demonstrated that PCNA is a trimeric processivity factor of eukaryotic replicative DNA polymerases.² The

outer surface of the ring-shaped PCNA trimer is negatively charged whereas the central cavity accommodating DNA carries positive charges.³ One PCNA monomer is composed of two domains that are linked by a domain-connecting loop. This loop is evolutionary conserved and plays a crucial role in mediating specific interactions of the replication clamp with its interaction partners in Okazaki fragment processing, DNA repair, translesion DNA synthesis, DNA methylation, chromatin remodeling and cell cycle regulation.⁴ In order to be functionally active, different PCNA-interacting proteins must bind and dissociate in a highly specific and ordered manner. In view of this fact it is surprising that the identified peptide domain motifs of PCNA-interacting proteins (PIP-motifs) are quite similar and have overlapping binding sites with PCNA. For instance, the sequence motif Q-x(2)-(h)-x(2)-(a)-(a) (h, moderately hydrophobic residue; a, a highly hydrophobic residue) is known to mediate interactions of ar-

† L.M.-C. and C.N. contributed equally to this work.

‡ Correspondence regarding SPR measurements should be addressed to D.F. (didier.flament@ifremer.fr).

Abbreviations used: PCNA, proliferating cell nuclear antigen; PIP-peptide, PCNA-interacting peptide; SPR, surface plasmon resonance.

E-mail address of the corresponding author:
hannu.myllykallio@igmors.u-psud.fr

chaeal and human PCNA with many partners.^{5,6} As these motifs are short, poorly conserved and can be found in literally thousands of proteins, the possibilities to identify (new) PCNA interacting proteins based solely upon *in silico* sequence comparisons are limited, making functional screens necessary.

Here we describe a combinatorial approach of *in silico* searches and combinatorial peptide synthesis to identify new PCNA interaction partners. In contrast to earlier screens that have used completely random peptide libraries, we have tested focused peptide libraries that were designed benefiting either from *Pyrococcus abyssi* or human genome. As model for the proof of principle we used the presumed enzymatic machinery for the processing of Okazaki fragments in the hyperthermophilic archaeon *P. abyssi*. Our approach led to identification of new interaction partners for archaeal PCNA and allows predicting hitherto unidentified interaction partners for human PCNA. Biological significance of our results was revealed through identification of an inactive complex of *P. abyssi* ribonuclease (RNase) HII with PCNA, indicating also a DNA repair function for RNase HII orthologs. We also show that the PIP-motif sustains a considerable amount of substitutions without having deleterious effect on the PCNA binding.

Results and Discussion

Screening of immobilized peptides with the robust archaeal *P. abyssi* PCNA

In order to experimentally identify PCNA-interacting peptides, we firstly screened a small library of 910 random X-C-X(8)-C-X peptides, immobilized onto a solid support *via* cysteine residues, with 2 μM of stable PCNA preparation from *P. abyssi*. *P. abyssi* PCNA carries a six-histidine tag at the amino-terminal end, allowing detection of the bound protein by a specific anti-His antibody. The background signal resulting from non-specific binding of *P. abyssi* PCNA corresponded to binding values smaller than 100 arbitrary units; no signal was observed when PCNA was omitted from the binding reaction. These analyses allowed the identification of 47 immobilized lead peptides with binding values higher than 500 for *P. abyssi* PCNA (Figure 1(a) and Supplementary Data, Table S1). Manual alignment showed that some of these peptides (Figure 1(b)) possess the PIP-motif.^{6,7} However, many additional peptides with high binding values did not show obvious sequence similarity to the PCNA-interacting protein motif (see Supplementary Data, Table S1), suggesting that

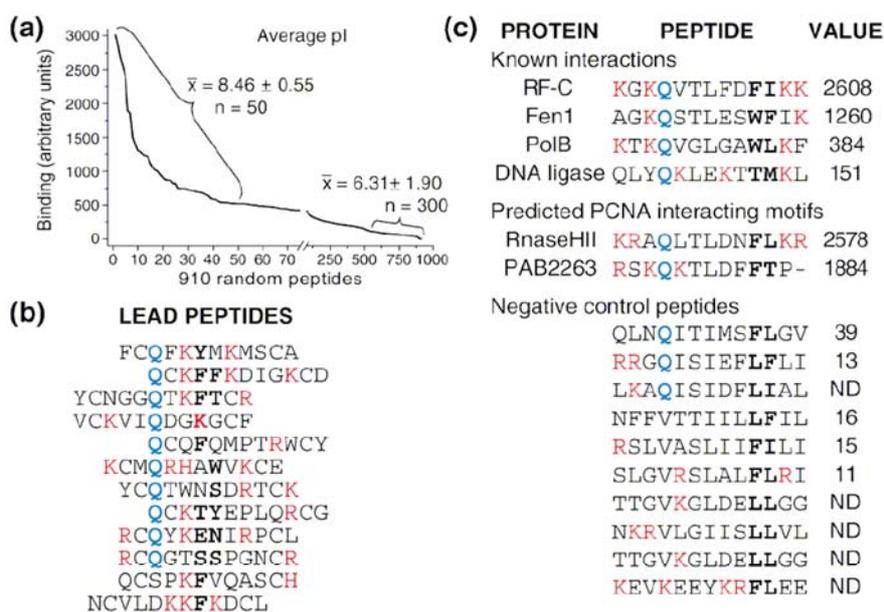


Figure 1. Screening of immobilized peptides with the robust archaeal *P. abyssi* replication clamp. (a) Screening of 910 random peptides with *P. abyssi* PCNA. Binding is indicated in arbitrarily chosen fluorescence units. Peptides were sorted according to the observed binding values. Predicted pI values for 50 high affinity peptides were statistically higher than for 300 peptides with the lowest binding values. Synthetic peptides were synthesized and screened using credit-card format mini-PEPSCAN cards (455-well plate with 3 μl wells) as described in Materials and Methods. (b) Lead peptides isolated from a random library with high affinity for *P. abyssi* PCNA. Hydrophobic residues in the vicinity of the conserved glutamine (in blue) are indicated in bold and positively charged residues are shown in red. (c) Screening of candidate *P. abyssi* PCNA interacting peptides. Sequences were identified using ScanProsite in the proteins encoded by the *P. abyssi* genome sequence. Relative binding values in arbitrary units are indicated on the right. The concentration of *P. abyssi* PCNA used for binding studies was 1 μg/ml. Several negative control peptides with their binding values are shown at the bottom of (c). ND, not detected.

physico-chemical or structural properties of these peptides played an important role in binding. This idea was supported by the fact that the average predicted isoelectric point (pI) of 50 high affinity peptides was statistically significantly higher than the pI of non-interacting peptides (8.46 versus 6.31; see Figure 1(a)). This observation indicates that the positive charge of these peptides could contribute to the observed binding. It is of note that previously performed peptide screens to identify new PCNA-binding peptides failed to identify statistically significant correlation between the PCNA binding and positive charge.^{8,9} This likely reflects the fact that not enough PCNA binding peptides were identified in earlier peptide screens and/or that binding values were not recorded for peptides with low affinities towards PCNA.

We next synthesized several control peptides found in *P. abyssi* proteins (Figure 1(c)) that were immobilized on a solid support and screened as above with 1 μ M *P. abyssi* PCNA. Naturally occurring peptides with the highest binding values contained the conserved *P. abyssi* sequence motif Q-x(2)-IL-x(2)-[WFI]-[LFT] (Figure 1(c)). We furthermore noted that this sequence motif is often flanked by positively charged residues (Figure 1(c)). Many "high affinity" peptides are present in proteins that are known to interact with the PCNA. However, we also found that peptides from *P. abyssi* ribonuclease H (RNase HII) and PAB2263, two proteins that previously have not been shown to interact with PCNA, contain a PIP-motif with a strong PCNA binding activity (Figure 1(c)).

The interaction domains of *P. abyssi* RNase HII, flap endonuclease 1 and DNA ligase with PCNA overlap

As a proof of principle, we next set out to further investigate how the above peptides when contained in archaeal proteins (their natural context) mediate interactions with PCNA. We studied the PCNA binding of the *P. abyssi* RNase HII, the structure-specific flap endonuclease 1 (Fen1) and the ATP-dependent DNA ligase using surface plasmon resonance (SPR). PCNA-interacting peptides located at the carboxyl termini of RNase HII and Fen1 have high binding values (Figure 1(c)), whereas the motif that was identified at the N terminus of DNA ligase shows moderate affinity for PCNA. In this respect it is noteworthy that very recently an additional PIP-motif "QKSFF" in the central region of *P. furiosus* DNA ligase has been described,¹⁰ suggesting that the two motifs could participate in efficient binding. First, recombinant proteins for SPR measurements, carrying a six-histidine tag at the N terminus, were produced and purified to greater than 99% homogeneity using Ni-NTA agarose and gel filtration chromatography (Figure 2(a)). SPR measurements with immobilized homotrimeric PCNA were performed to analyze the physical interactions of the purified proteins with PCNA (Figure 2(b)–(d)). The results indicated that RNase HII, Fen1 and DNA ligase physically interact with *P. abyssi* PCNA at

nanomolar concentrations (Figure 2(b)–(d), trace 1). It is also of note that the high binding values observed for the individual peptides (Figure 1(c), and data not shown) correlate with the K_D values in the nanomolar range. These values were estimated using global fitting analysis that accounts for both association and dissociation phases as well as the maximal level of complex formation. It is interesting to note that the similar SPR data have been obtained for a PAB2263–PCNA complex of unknown function (unpublished results).

To exclude that the observed SPR data result from the oligomerization of the proteins studied and/or from non-specific interactions under the experimental conditions used, we performed deletion analyses and competition experiments. RNase HII, with a 17 residue deletion at the C terminus that removed the PIP-motif, failed to interact with PCNA (Figure 2(b), inset). The competition experiments were performed in the presence of a tenfold excess of the peptide KRKQLTLDNFLKR (bold residues correspond to conserved PIP-motif residues). This peptide is a rationally designed composite peptide based on the RNase HII PIP-peptide (KRAQLILDNFLK) where the alanine at the position 3 was changed to lysine that occurs in Fen1 and PAB2263 PIP-peptides in the corresponding positions. The binding value of this peptide is similar to those of Fen1 and PAB2263 (data not shown). The addition of this peptide substantially decreased the rate of complex formation and/or accelerated dissociation of RNase HII-, Fen1- and DNA ligase–PCNA complexes (Figure 2(b)–(d), trace 3), whereas the non-binding control peptide (KEVKEEYKRFLK) had no effect on complex formation (Figure 2(b)–(d), trace 2). These results clearly indicated that the binding sites for all three proteins with PCNA do overlap considerably. To obtain further evidence for the physiological significance of the interactions detected *in vitro* between PCNA, RNase HII and Fen1, their complex formation in cell-free extracts was analyzed (Figure 2(e)). Fen1 and RNase HII were precipitated together with PCNA from cell-free extracts, whereas the lack of detectable interaction between DNA ligase and PCNA could be explained by additional PCNA interaction partners, present in cell-free extracts, preventing the binding of DNA ligase to PCNA.

The specific physical interaction between *P. abyssi* PCNA and RNase HII measured by SPR suggested that PCNA might modulate the activity of RNase HII. We therefore examined the influence of PCNA on the enzymatic properties of *P. abyssi* RNase HII on heteroduplex substrate carrying embedded ribonucleotides, which likely corresponds to the physiologically relevant RNase HII substrate.^{11,12} Upon addition of PCNA the RNase HII activity on this substrate was strongly inhibited (Figure 2(f)). To test whether this inhibition results from a physical inhibitory interaction with PCNA, RNase HII activity was measured by changing the molar PCNA/RNase HII ratio in the reaction mixtures. These studies were performed with both wild-type (WT) RNase HII and its C-terminal deletion (Δ Cter)

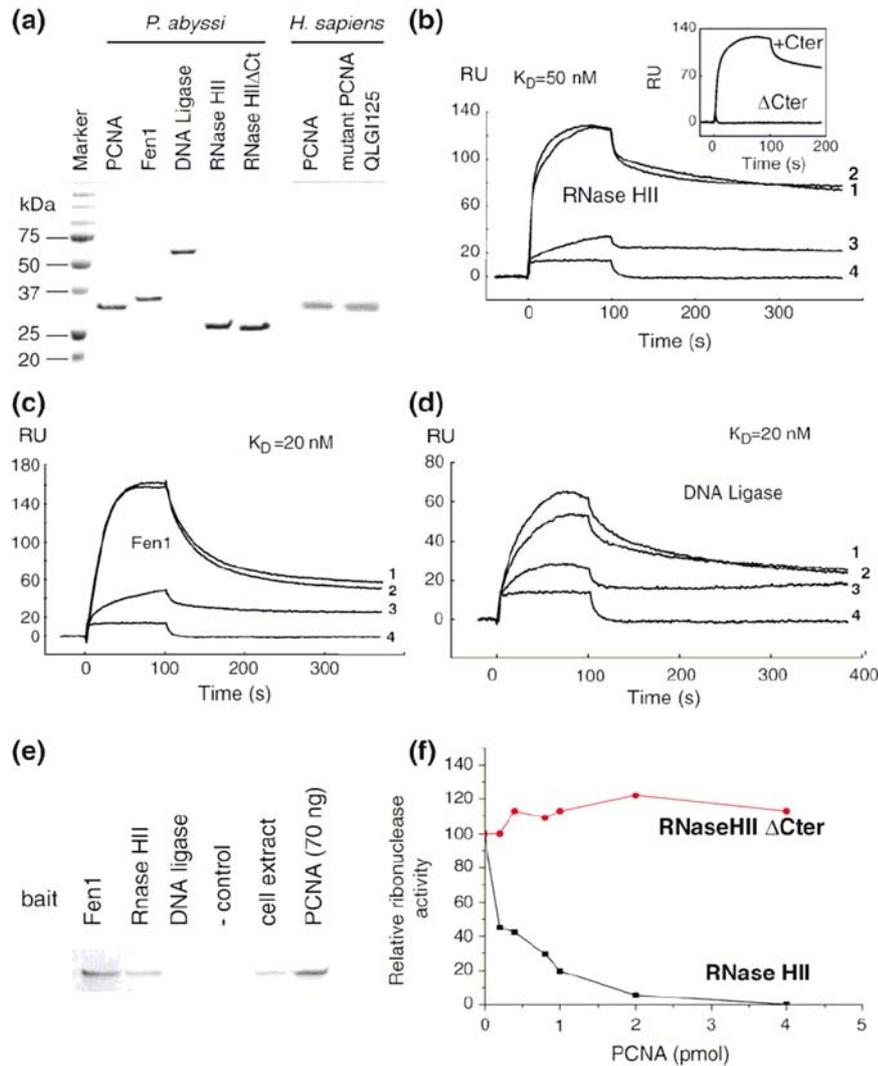


Figure 2. The interaction domains of *P. abyssi* RNase HII, Fen1 and DNA ligase with PCNA overlap (a) The proteins used in this study: 1 μ g of each purified *P. abyssi* or human DNA replication protein were separated on a SDS–13% PAGE and stained by Coomassie blue. RNase HII Δ Ct refers to the carboxy-terminal deletion derivative of *P. abyssi* RNase HII (see Material and Methods). (b)–(d) Experiments on PCNA chips. Ten pmol of RNase HII (b), Fen1 (c) or DNA ligase (d) were injected at time 0 over the PCNA chip (trace 1; panels (b), (c) and (d)). All data are indicated in RU units as a function of time. When indicated, 100 pmol of non-interacting peptide control (trace 2) or a control peptide KRKQLTLDNFLKR were included during injection (trace 3). As an additional control an interacting peptide alone over the PCNA chip was injected (trace 4). The background resulting from buffer injection was always subtracted. The inset in (b) shows that the C-terminal PIP-domain of the *P. abyssi* RNase HII is necessary and sufficient for binding to the PCNA chip. + Cter, wild-type *P. abyssi* RNase HII; Δ Cter, *P. abyssi* RNase HII lacking 17 residues from the amino terminus. (e) A pull-down experiment on magnetic beads. The 20 μ g of histidine-tagged Fen1, RNase HII or DNA ligase (from left to right) bound to magnetic beads, and a negative control that consists of beads alone, were incubated with 12.5 mg of *P. abyssi* cellular extract. Control lanes contained 7 μ g of cell free extract and 70 ng of PCNA. Pull-down assay, SDS–polyacrylamide gel electrophoresis and immunoblot using anti-PCNA antibodies were performed as described in Materials and Methods. (f) RNase H activities in dependence of PCNA. Activities of RNase HII and RNase HII Δ Cter (1 pmol) were tested on a DNA substrate radioactively labeled at 5' terminus containing a single embedded ribonucleotide (1 pmol per 20 μ l reaction) in the absence and the presence of *P. abyssi* PCNA.

derivative incapable of interacting with PCNA (Figure 2(b)). The results obtained revealed that inhibition of the ribonuclease activity of RNase HII by PCNA was dependent on the PIP-motif located at the C terminus of the enzyme (Figure 2(f)). This inhibition could be explained by the fact that the

substrate-binding domain of *Thermococcus kodakaraensis* RNase HII is located in a carboxy terminus.¹¹ In control experiments we found that the PCNA preparation used was capable of activating *P. abyssi* Fen1 and DNA ligase (data not shown), indicating that PCNA was fully functional and properly folded.

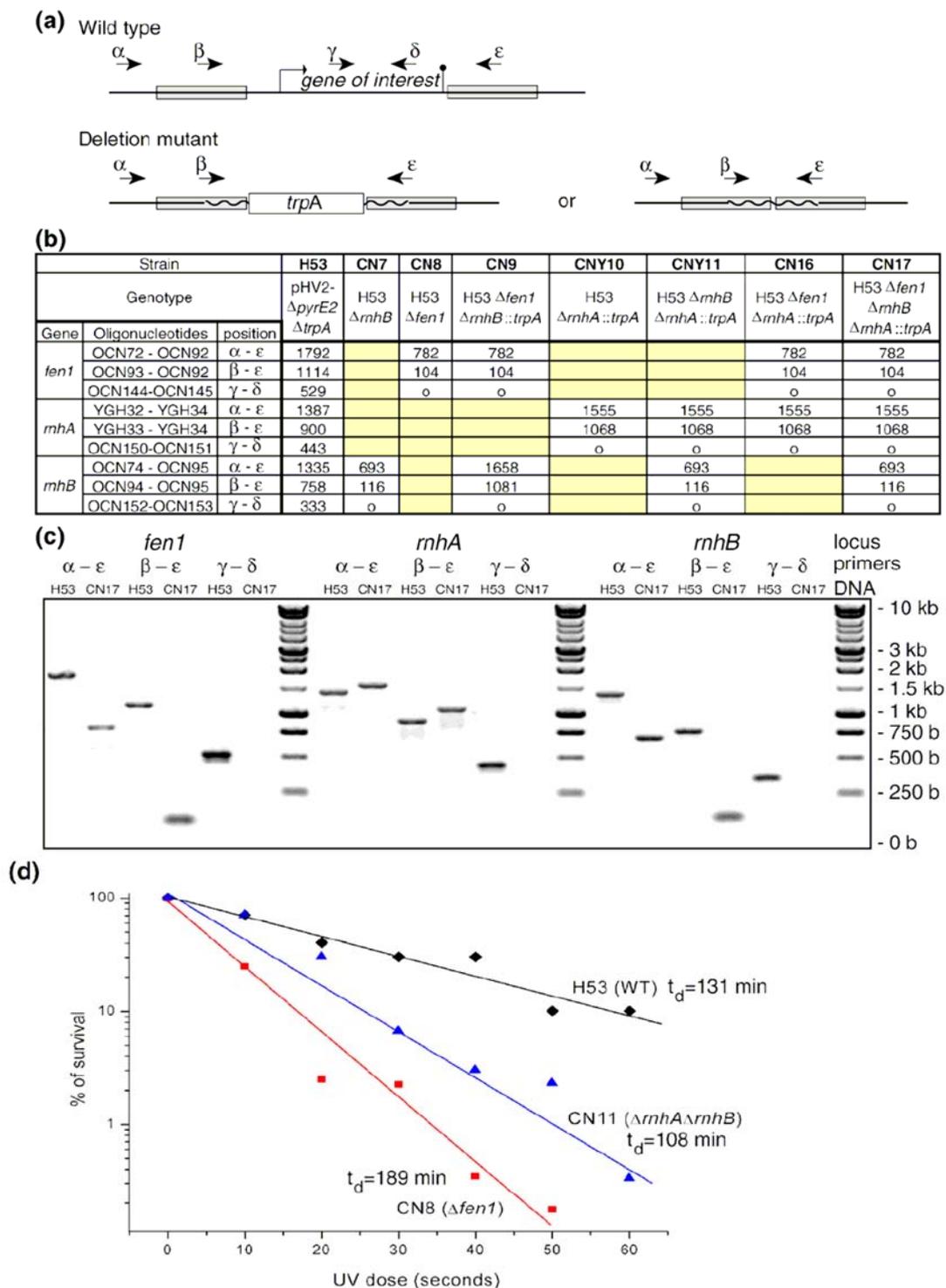


Figure 3. Deletion of *fen1*, *mhA* and *mhB* genes in *H. volcanii*. (a) Schematics indicating the annealing positions of the primers (arrows) used to confirm the deletions for each gene. Homology sequences used for the pop-in/pop-out constructs are shown by shaded boxes, the straight line corresponds to the chromosome and the undulated line to the remains of plasmidic DNA after the pop-out. (b) The expected sizes of the PCR products obtained with the different genotypes for each deletion. The panel (b) also indicates genotypes for the strains used in this work. (c) Agarose gel electrophoresis confirming gene deletions. The PCR products obtained using either the wild-type H53 or the triple mutant CN17 DNA are shown. Each product presents the expected size indicating the loss of *fen1*, *mhA* and *mhB* genes in the strain CN17. Similar results were obtained for the other strains indicated in Table 1. (d) Fraction of surviving *H. volcanii* cells on solid enriched media as a function of UV exposure is shown. One second exposure corresponds to 1.24 J/m². The strains used were H53 (wild-type), CN11 (Δ rnhA Δ rnhB), and CN8 (Δ fen1).

Archaeal RNase H and Fen1 have a key role in DNA repair

To further define the physiologically relevant activities of archaeal RNase HII and Fen1, we inactivated the structural genes for RNase HII and Fen1 (Figure 3). For this we chose the halophile, *Haloferax volcanii*, since the necessary genetic tools have not been developed for *Pyrococcus* species. It is of note that we were unable to obtain a *H. volcanii* strain carrying a deletion of the structural gene for PCNA, suggesting that this processivity factor is essential in this organism. While *P. abyssi* contains only one RNase H (family II encoded by *rnhB*), *H. volcanii* contains detectable orthologs for RNase HII and RNase HI (likely bacterial origin (encoded by *rnhA*)). Mutant strains, lacking *fen1*, *rnhA*, *rnhB*, or different combinations of these, were constructed and confirmed as indicated in Materials and Methods. On solid medium, the “wild-type” and $\Delta rnhB$ $\Delta rnhA$ strains grew similarly whereas the growth of the $\Delta fen1$ deletion strain was impaired. This visual observation was confirmed by measuring the doubling time for the parental and $\Delta rnhB$ $\Delta rnhA$ of $131(\pm 6)$ ($n=5$) and 108 (average of two measurements) min whereas the $\Delta fen1$ strain had a doubling time of 189 min (average of two measurements). We also measured the incorporation rate of extracellular tritiated uracil into alkaline resistant form (DNA) using these different mutant strains (Table 1). These measurements suggested that archaeal Fen1 has a role in DNA replication, as earlier observed for Eukarya.¹³ On the other hand, the *in vivo* role of RNase HII in DNA replication was not evident (Table 1). Considering that GC% of *H. volcanii* genome is relatively high, we can not exclude a possibility that *H. volcanii* Rnase HI participates in the processing of R-loops that could be found at some replication origins in this species.^{14,15}

We also tested our deletion strains for possible differences regarding their sensitivity to UV radiation (Table 1; Figure 3). Our results revealed that the *fen1* deletion resulted in markedly enhanced UV sensitivity, indicating that archaeal Fen1 either participates in removal of photoproducts resulting from UV irradiation and/or indicate a role in DNA replication occurring downstream of unrepaired damage on lagging strand templates.^{13,16} Our results also indicated that both single RNase H mutant strains are moderately UV sensitive (Table 1), with an additive effect observed for the $\Delta rnhA$ $\Delta rnhB$ strain (Table 1 and Figure 3(d)). Thus, substrate specificity of archaeal RNase HII orthologs, inhibition of

Table 1. Genotypes and phenotypes of various *Haloferax volcanii* strains

Name	Genotype	Incorporation rate ^a (%)	% of survival ^b ultraviolet light (112 J/m ²)
H53	Wild-type ^c	100	0.25
CN7	$\Delta rnhB$	130	0.035
CN8	$\Delta fen1$	70	0.000125
CN9	$\Delta fen1 / \Delta rnhB$	65	0.0025
CN10	$\Delta rnhA$	75	0.15
CN11	$\Delta rnhA / \Delta rnhB$	108	0.005
CN16	$\Delta fen1 / \Delta rnhA$	68	0.0003
CN17	$\Delta fen1 / \Delta rnhA / \Delta rnhB$	33	0.006

^a Relative incorporation rates of tritiated uracil (marked at the positions 5 and 6 of pyrimidine ring) into alkaline resistant form (chromosomal DNA) was measured during early exponential growth phase using TCA precipitation as described by Myllykallio *et al.*³⁰ 100% corresponds to incorporation of 0.5 fmol of tritiated uracil into alkaline resistant form in 1 h by the strain H53. Cell numbers were normalized using OD₆₀₀ values prior all labelling experiments.

^b Percentage of survival was measured using a 90 s exposure with a dose of 1.24 J/m²/s. Survival curves for H53, CN8 and CN11 strains are shown in Figure 3.

^c *Haloferax volcanii* strain H53 is referred to as “wild type” through out this work.

RNase HII activity by PCNA (Figure 2), and genetic studies (Figure 3) propose that the major physiological role of archaeal RNase HII is not to participate in DNA replication but rather, possibly together with Fen1 endonuclease,¹² in repair of ribonucleotides misincorporated into chromosomal DNA. As translesion bypass polymerases capable of ribonucleotide misincorporation are often found in archaeal species exposed to UV radiation, for instance in *H. volcanii*, the *Sulfolobales*, and *Picrophilus* species,¹⁷ translesion synthesis through the UV damaged DNA, could be one mechanism that increases ribonucleotide misincorporation in chromosomal DNA.

Predicted interaction partners of human PCNA

The above experiments performed with archaeal proteins indicated that a combination of motif searches and screening with immobilized peptides can be used to identify new PCNA interacting partners. To expand this notion to different systems, we next synthesized 98 peptides carrying the Q-x(4,5)-[HDF]-[FY] motif occurring in human proteins. These peptides were subsequently screened with human PCNA carrying a C-terminal hexahistidine tag.⁷ Figure 4(a) shows that, despite all 98 peptides screened contained the aforementioned sequence

Figure 4. Predicted interaction partners of human PCNA. (a) Screening of 98 naturally occurring peptides present in human proteins annotated in Swiss-Prot database containing the sequence motif x(3)-Q-x(4,5)-[HDF]-[FY]-x(3). The screening was performed with human PCNA (1 µg/ml). The average pI value for 25 top binders was 8.40 ($n=25$) and for the non-binders 5.87 ($n=25$). Two sample independent *t*-tests revealed that the difference of population means was significantly different (p -value=0.00113). (b) Positional scan of the underlined residues in the peptide KRKQ¹LTLDN²FLKK. These residues correspond to the conserved residues of the archaeal and eukaryal PIP-motif. Green columns correspond to wild-type peptides whereas peptides carrying substitution are indicated in blue. Binding values were normalized using the values obtained for wild-type peptide. (c) Frequency distribution of pI values for 779 human peptides with the predicted PIP-motif. Note that the data shown in Figures 1(c) and 3(a) already provide experimental support, suggesting that some positively charged peptides in this data set could be used to identify novel interaction partners for the human PCNA

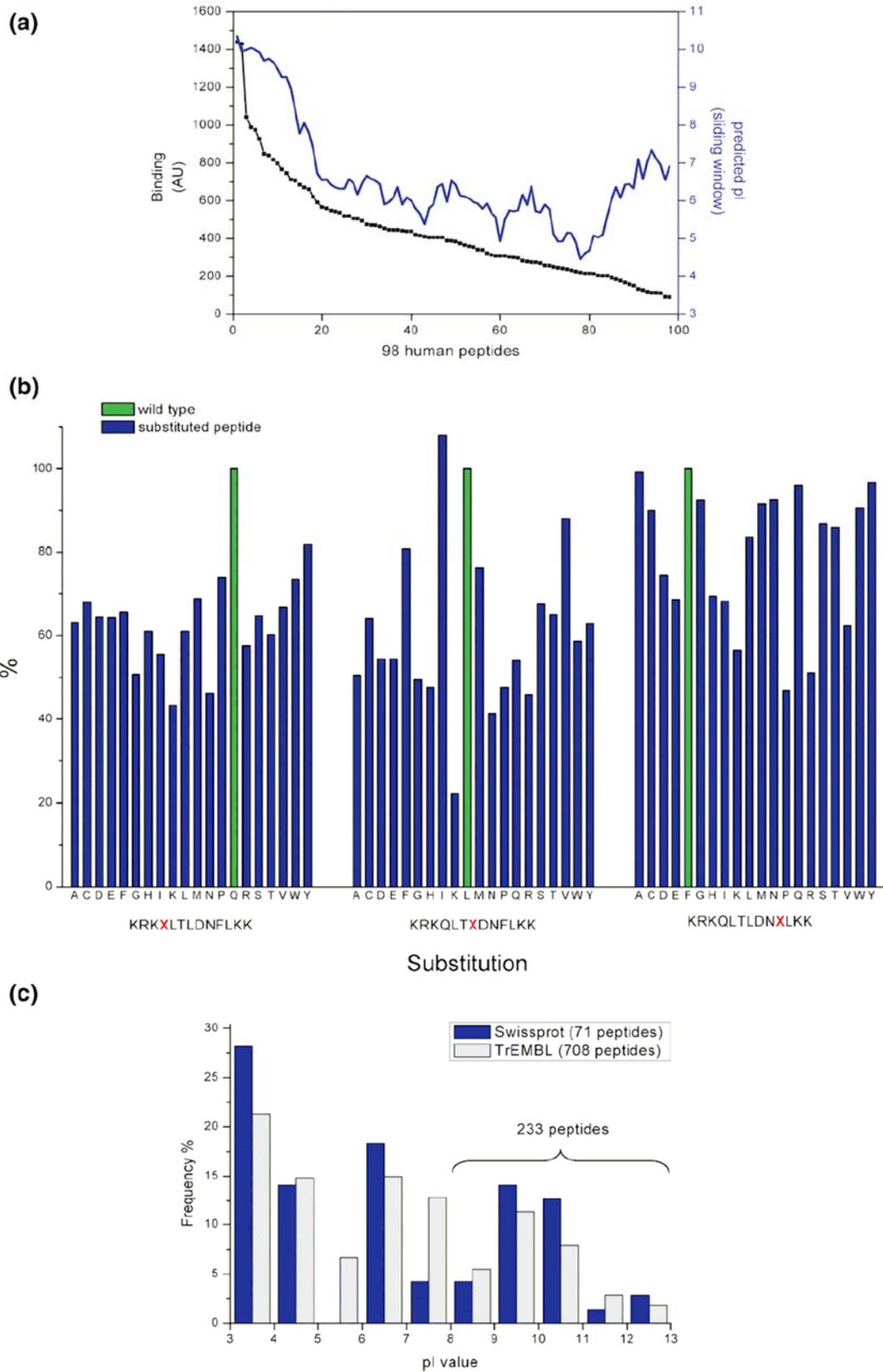


Figure 4 (legend on previous page)

motif, the observed binding values varied between 80 and 1400 (left y axis). These experimental data are consistent with the earlier prediction that variable PIP-box sequences could result in a wide range of PCNA affinities.¹⁸ The sliding average (in a window of ten peptides) of the calculated pIs (Figure 4(a), right y axis), clearly indicates that high affinity peptides tend to be positively charged. Supplementary Data, Figure S1A shows an alignment of the peptides with observed binding values higher than 400. This list not only contains proteins known to interact with PCNA (bolded in Supplementary Data, Figure S1A), but also reveals several potential novel interaction partners for human PCNA. Obviously, more experiments are now required to address whether these proteins interact functionally and/or physically with human PCNA.

Since the interaction surface between the PCNA and its interaction partners is at least partially formed by the beta zipper peptide backbone,^{19,20} formation of the complex interface could be largely independent of the protein sequence. To test this, we synthesized 60 different peptides where the bolded positions in the peptide KRKQLTLDNFLKR were systematically changed to all naturally occurring amino acids (Figure 4(b)) and performed a further functional screening. Our results revealed that not even the "fully conserved" glutamine residue is indispensable for binding in our experiments. On the other hand, the hydrophobic substitutions at positions 7 and 10 were favored. These results indicate that the short PIP-motif sustains a high number of substitutions without having a drastic effect on the binding activity. These observations are in agreement with the notion that the beta zipper formation together with electrostatic and hydrophobic interactions, and not specific protein sequence, is important for binding. Similar results further indicating specific binding independent of peptide sequence were obtained after screening with 80 additional peptides based upon KRKQLTLDNFLKR, where the underlined residues were changed in the same way as described above (Supplementary Data, Figure S1B). With these peptides we also performed a screening with the QLG125 mutant in the domain-connecting loop of human PCNA. This loop participates in specific interactions of PCNA with its interaction partners, and the mutant PCNA has a markedly lower affinity towards Fen1 than the wild-type protein.⁷ The signal observed with the QLG125 mutant was consistently lower than observed for the wild-type protein (Supplementary Data, Figure S1B), further indicating that the binding signal observed in our experiments is highly specific. The fact that binding by this mutant was not completely abolished in our experiments is in agreement with the proposed multi-factorial nature of this interaction.

Conclusions

We have developed an experimental approach that allows the rapid identification of peptide sequences with high affinity towards PCNA. Differently from

the earlier peptide screens to identify PCNA binding peptides, we have benefited from the combined *P. abyssi* and human genome sequence information to synthesize a number of immobilized peptides carrying putative PIP-motifs. Usefulness and efficiency of this approach to identify novel PCNA interaction partners were revealed by our finding that *P. abyssi* PCNA and RNase HII form an inactive complex. While proof of principle was obtained using robust archaeal proteins, our data indicate that this experimental screening method can be extrapolated to human systems. In particular, we have identified new peptide sequences in human proteins with high affinity towards PCNA (Supplementary Data, Figure S1). Moreover, we can predict that some of the 233 positively charged peptides ($pI > 8.0$) of human proteins carrying the modified PIP-motif Q-X(2)-[LIMFV]-X(2)-[DHFYW]-[FYAQ] are likely to correspond to additional candidate PCNA interaction partners (Figure 4(c)). Our results also indicate that the PIP-motif sustains a considerable amount of substitutions, indicating that PCNA binding of the PIP-peptides is mediated through beta zipper formation at the interaction surface¹⁹ and the positively charged residues (Figures 1 and 4). As this interaction is largely independent of the specific sequence, PCNA complexes are likely to persist even when the PIP-motif cannot be recognized anymore using computational approaches. Despite the power of genome comparison in the detection of linear interaction motifs being obvious,^{21,22} proteomics approaches like the one described here are clearly needed to validate the observed predictions.

Materials and Methods

Bacterial strains

Escherichia coli SURE (*e14*– (*McrA*–) Δ (*mcrCB-hsdSMR-mrr*)171 *endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5* (Kan^r) *uvrC* [F' *proAB lacI^qZ* Δ M15 Tn10 (Tet^r)] (Stratagene) was used as a host strain for maintaining plasmids. For protein expression, *cis*-repressed pQE-80L (Qiagen) derivatives were transformed into *E. coli* BL21 CodonPlus (*E. coli* B F– *ompT hsdS*(r_B – m_B –) *dcm*+Tet^r *gal* λ (DE3) *endA Hte* [*argU ileY leuW* Cam^r]) (Stratagene) containing extra copies for arginine, isoleucine and leucine tRNAs.

Expression and purification of recombinant proteins

Genes encoding *P. abyssi* PCNA (PAB1465), RNase HII (PAB0352), Fen1 (PAB1877) and DNA ligase (PAB2002) were amplified using 30 cycles of PCR with CsCl gradient purified chromosomal DNA as template (primer sequences will be made available upon request). The pQE-80L (Qiagen) expression vector containing isopropyl- β -D-thiogalactopyranoside (IPTG)-inducible bacteriophage T5 promoter was used to express all proteins. *P. abyssi* PCNA was cloned using BamHI and SalI restriction sites (all primer sequences are indicated in Table 1); all other proteins were cloned into BamHI and PstI sites of the pQE-80L expression vector. The resulting expression constructs

encode full-length proteins with a six-histidine tag at their amino termini. A construct encoding truncated version of RNase HII lacking the 17 carboxy-terminal residues and carrying PCNA interaction motif was produced in a similar manner.

Expression constructs were transformed into the *E. coli* BL21-CodonPlus-RIL strain (Stratagene) using solid LB medium containing 100 µg/ml of ampicillin. Protein expression was induced in exponential phase cultures by adding 0.5 mM IPTG. After 2 h of induction, the cells were collected by centrifugation. Cell lysis was performed in buffer A (30 mM Hepes (pH 8), 300 mM NaCl) with several freeze-thaw cycles, followed by a brief sonication to decrease viscosity of supernatants. Cellular debris was eliminated by centrifugation. All chromatographic procedures were performed with an ÄKTA FPLC system at 10 °C (Amersham Biosciences). Tagged proteins were purified on immobilized Ni-NTA beads, according to the manufacturer's instructions (Qiagen), followed by purification on an S-200 gel filtration column (Amersham). All protein samples were analyzed for purity and integrity using SDS-13% (w/v) polyacrylamide gel electrophoresis and by MALDI-TOF analyses (Innova Proteomics, France). Recombinant human wild-type PCNA as well as its QLGI25 mutant were purified to over 99% homogeneity as described.^{7,23}

Synthesis of peptides and PEPSCAN screening

The overlapping synthetic peptides were synthesized and screened using credit-card format mini-PEPSCAN cards (455-well plate with 3 µl wells) as described.²⁴ The binding of antibodies to each peptide was tested in a PEPSCAN-based enzyme-linked immunosorbent assay (ELISA). The 455 well credit-card-format polypropylene cards, containing the covalently linked peptides, were incubated with sample, for example 1 µg/ml of PCNA diluted in a phosphate-buffered saline (PBS) solution which contains 5% (v/v) horse serum and 5% (w/v) ovalbumin and 1% (v/v) Tween 80 (4 °C, overnight). After washing, the peptides were incubated with an anti-HIS tag antibody (dilution 1–1000; Novagen) (1 h at 25 °C) and subsequently with a rabbit anti-mouse antibody peroxidase (dilution 1/1000; Dako) (1 h at 25 °C). After washing the peroxidase substrate, 2,2'-azino-di-3-ethylbenzthiazoline sulfonate (ABTS) and 2 µl/ml of 3% (v/v) H₂O₂ were added. After 1 h color development of the ELISA was quantified with a CCD-camera and an image processing system.

Interaction studies

Surface plasmon resonance studies were performed in a BIAcore X apparatus (BIAcore, Uppsala Sweden), the interactions were monitored at 25 °C at a flow rate of 30 µl/min in HBS-P buffer (10 mM Hepes (pH 7.4), 150 mM NaCl and 0.005% (v/v) P20). PCNA was amine-coupled on the surface of CM5 sensor chips (BIAcore) in 10 mM sodium acetate buffer (pH 4.5). The level of immobilisation of PCNA was of about 150 resonance units (RU).

P. abyssi GE5 (strain Orsay) cell-free extracts were prepared from anaerobically grown 2 l cultures using a gas-lift bioreactor, which was prepared as described.²⁵ Cell suspensions were harvested in exponential phase (approximately 2.0 × 10⁸ cells/ml) and centrifuged at 5000 g for 45 min. Cells were suspended in 50 mM Tris-

HCl (pH 8), 0.5 mM dithiothreitol, 150 mM NaCl and lysed by sonication using a Vibracell sonifier (15 pulses of 30 s, 375 W, 40% amplitude). After centrifugation at 10,000 g for 15 min, the supernatant was used as the cell-free extract. Then 20 µg of His-tagged recombinant proteins were incubated with 1.6 mg of Co²⁺ magnetic micro beads (Dynabeads® TALON™; Dynal Biotech) in 1 ml of binding and washing buffer (50 mM sodium phosphate (pH 8), 300 mM NaCl and 0.01% Tween 20) for 20 min at 25 °C. Proteins bound to the beads were washed three times in the same buffer. The beads were then incubated with 12.5 mg of whole-cell extract for 2 h at 4 °C. After five washing steps using the above-mentioned buffer, proteins bound to the target proteins were eluted in SDS sample buffer (1 M Tris (pH 8.8), 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 8% (v/v) sucrose and 0.1% (w/v) bromophenol blue). The eluted samples were separated in an SDS-12% polyacrylamide gel and transferred to a nitrocellulose membrane. After membranes were blocked with 5% (v/v) milk in TBST buffer (10 mM Tris-HCl (pH 7.5), 150 mM NaCl and 0.05% Tween 20), polyclonal anti-PCNA antibody (raised in rabbits by Eurogentec) was incubated for 1 h at 25 °C and then washed in TBST. Immunocomplexes were detected using an enhanced chemiluminescence procedure according to the manufacturer's recommendations (Amersham).

RNase HII assays

Activity assays using RNase HII were performed essentially as described.²⁶ A substrate consisting of oligonucleotides O8 and O9 (Table 2) was obtained by mixing a deoxyoligonucleotide carrying an embedded ribonucleotide in a 1:1 molar ratio with the complementary strand in 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl, followed by an incubation at 75 °C for 15 min. After this incubation step, oligonucleotides were cooled slowly at room temperature and conserved at -20 °C until use.

Where indicated, one of oligonucleotides used in the hybridizations was labeled either on the 5' end with [γ -³²P]ATP by using T4 Polynucleotide kinase (Promega) or, alternatively, the 3' end was marked using the Terminal Desoxynucleotide Transferase (Promega) and [α -³²P] ddATP (Amersham). Free label was removed from oligonucleotide solutions using Sephacryl G25 MicroSpin columns (Amersham). Reaction products and substrates were separated under denaturing conditions. Polyacrylamide gels were dried on Whatman filter paper and reactions were quantified using the Storm system.

Gene deletion in *Haloferax volcanii*

Gene deletion in *H. volcanii* strain H53 (p HV2 Δ pyrE2 Δ trpA) and segregation of mutant alleles were carried out using the optimized pop-in/pop-out system.^{27,28} A non-selective enriched medium Hv-YPC or a selective Hv-Ca medium were used for cell growth. Where appropriate, selective medium was supplemented with uracil and tryptophan (50 µg/ml each) or, alternatively, using a combination of 5-fluoroorotic acid (50 µg/ml) and uracil (10 µg/ml) with or without tryptophan (50 µg/ml). Plasmids used for *H. volcanii* gene inactivation carried approximately 500 bp 5' upstream and 3' downstream the ORF sequences (oligonucleotides used are indicated in Table 2). PCR analyses were used to confirm that the constructed strains corresponded to the expected deletion strains (Figure 3(a)–(c)). Flanking regions were amplified

1146

Identification of PCNA Interacting Peptides

Table 2. Oligonucleotides used in this study

Usage	Oligonucleotides	Sequence 5' to 3'
<i>Recombinant proteins</i>		
PCNA	PCNA BgIII PCNA Sall	ccatagatctATGCCATTCGAGATAGTC ccatgtccgagtegacTTACTCCTCAACCCCTGGG
Fen1	Fen1 BamHI Fen1 PstI	ccatggatccATGGGAGTTCCTATAGGGG ccatctcgagTTAGGGTTTCTTTTGTATGAACCAGC
RnaseIII	Rnase III BamHI Rnase III PstI	ccatggatccATGAAGGTTCGCCGAGCAGATG ccatctcgagTTAATTGCGAAACCGCTTAAG
Ligase	Lig1 BamHI Lig1 PstI	ccatggatccATGAGGTACATAGAGCTGGCC ccatctcgagTTACACCTTCCCTTCATCCTC
<i>Gene deletion</i>		
<i>pcna</i> deletion	Hvo1 PCNA5 Hvo1 PCNA3 Hvo2 PCNA5 Hvo2 PCNA3	gatcgaattcGTCGCGGGCGTCGCCGCGCCA gatcggatccATATGATGTTCCGACCGTCTC gatcggatccGAGCCGTCCCACTCGTTTCTC gatctctagaGGTTCGAAGTCGAGCGGGCC
<i>fen1</i> deletion	OCN39 OCN40 OCN41 OCN42	gctctagagCAGTTCGGCCATGGAGTTCG cgggatcccgAGTGACCGGCTCGAATTTCG cgggatccGGCGGCTCGCATCGAGACA ggaattccAGCAITCCGTAGGTGCGGAG
<i>fen1</i> control	α OCN72 β OCN93 γ OCN144 δ OCN145 ε OCN92	GTCGTAGACGGGTGCGAAG GCGCGAAGGAGTTGAACC GGTGTCTGTTGACGAGGTG GAGCTGTCCGAGCGTGTAG CGAACCAGCCCGTGTAAA
<i>rnhA</i> deletion	OCN138 OCN141 YGH19 OCN140	ggaattcCCTCCTTCCGACA gctctagACTCGAGCGACTCCGGTAA gttttatccacgggatccAGTTGGACTCGACAACCTC acgcacataacggatccAGA ACTCCCGAAGACGTG
<i>rnhA</i> control	α YGH32 β YGH33 γ OCN150 δ OCN151 ε YGH34	GATTCCTTGGCGTTCGAC CCGGTCTCCTCGTGAATC CGGAAAGTCGTCGTACAG CCTCGTTGGCTAGGGAGTC AGCACGTCCTCCGGGAGT
<i>rnhB</i> deletion	OCN43 OCN44 OCN45 OCN46	ccgctegagCGACATGGTCGAGCGCGTG cctcaagcttACCTGGTTCGGGTCTTGGGGTG gctcaagcttGGCGTCGGGAACCGGCGG cgggatccCGGCCTCGACGATTCTCGAAG
<i>rnhB</i> control	α OCN74 β OCN94 γ OCN152 δ OCN153 ε OCN95	CGTGGTCGAAATCATGGAG ACCGCAACCTCTAATTTTCC CGAGACGGACATGAACCTG ACGTACTCCCGAGGAACT GTAATCGTCTCCGCGATT
<i>RNase III Substrate (Bold A refers to a ribonucleotide)</i>		
	O8	CCGTCCC AC CGTGCC
	O9	GGCACGGTGGGACGG

using PCR and cloned into pTA131 containing a *pyrE2* selection marker using EcoRI and XbaI restriction sites, yielding plasmids pCN7 (carrying the $\Delta rnhB$ allele) and pCN6 ($\Delta fen1$). Non-completed genome sequence of *H. volcanii* was used for primer design§. The *trpA* selection marker from pTA298²⁷ was inserted at the BamHI site of pHM306 ($\Delta pcna::trpA$), pYGH40 ($\Delta rnhA::trpA$) and pCN9 ($\Delta rnhB::trpA$) that was created between two inserted fragments. The non-replicative plasmids obtained were re-isolated from *E. coli* SCS110 *9dam⁻/dcm⁻* (Stratagene) and non-methylated DNA was used to transform *H. volcanii* using polyethylene glycol 600 as described.²⁹

Computational methods

Scan prositell with variations of PIP-motifs was used to scan *P. abyssi* or *Homo sapiens* protein sequence(s) from

§ <http://www.tigr.org>
 ¶ <http://www.expasy.org/tools/scanprosite/>

Swiss-Prot and TrEMBL databases. For human proteins, description filter “DNA” was used to limit the number of the hits. The sequence manipulation suite¶ was used to predict isoelectric points for the different peptides. Statistical tests were performed using OriginPro 7.5 software package (OriginLab Corporation).

Acknowledgements

We thank G. Henneke for helpful discussions, P. van Dijken for comments on combinatorial peptide synthesis and U. Liebl for critical comments on the manuscript, We also thank Pascal Trouvé (INSERM

¶ http://www.bioinformatics.vg/sms/protein_iep.html

U613, Molecular Genetics and Genetic Epidemiology, Brest, France) for access to surface plasmon resonance facilities. Our work on archaeal and eukaryotic DNA replication is supported by European Union programs (REPBIOTECH Elsevier Ltd. (contract number QLK3-CT-2002-02071) and Marine Genomics (contract number: PNGOCE-CT-2004-505403)). D.F. also thanks the French National Institute of Marine Genomics for financial support (contract number: FNS-500-370). H. M. acknowledges INSERM AVENIR program and Fondation Bettencourt Schueller for financial support. U.H. and E.F. are supported by the University of Zürich.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2007.06.056](https://doi.org/10.1016/j.jmb.2007.06.056)

References

- Miyachi, K., Fritzler, M. J. & Tan, E. M. (1978). Autoantibody to a nuclear antigen in proliferating cells. *J. Immunol.* **121**, 2228–2234.
- Prelich, G., Tan, C. K., Kostura, M., Mathews, M. B., So, A. G., Downey, K. M. & Stillman, B. (1987). Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. *Nature*, **326**, 517–520.
- Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell*, **87**, 297–306.
- Maga, G. & Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J. Cell Sci.* **116**, 3051–3060.
- Duggin, I. G. & Bell, S. D. (2006). The chromosome replication machinery of the archaeon *Sulfolobus solfataricus*. *J. Biol. Chem.* **281**, 15029–15032.
- Warbrick, E. (1998). PCNA binding through a conserved motif. *Bioessays*, **20**, 195–199.
- Jonsson, Z. O., Hindges, R. & Hubscher, U. (1998). Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *EMBO J.* **17**, 2412–2425.
- Xu, H., Zhang, P., Liu, L. & Lee, M. Y. (2001). A novel PCNA-binding motif identified by the panning of a random peptide display library. *Biochemistry*, **40**, 4512–4520.
- Warbrick, E. (2006). A functional analysis of PCNA-binding peptides derived from protein sequence, interaction screening and rational design. *Oncogene*, **25**, 2850–2859.
- Kiyonari, S., Takayama, K., Nishida, H. & Ishino, Y. (2006). Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J. Biol. Chem.* **281**, 28023–28032.
- Haruki, M., Tsunaka, Y., Morikawa, M. & Kanaya, S. (2002). Cleavage of a DNA-RNA-DNA/DNA chimeric substrate containing a single ribonucleotide at the DNA-RNA junction with prokaryotic RNases HII. *FEBS Letters*, **531**, 204–208.
- Rydberg, B. & Game, J. (2002). Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc. Natl Acad. Sci. USA*, **99**, 16654–16659.
- Liu, Y., Kao, H.-I. & Bambara, R. A. (2004). FLAP ENDONUCLEASE 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* **73**, 589–615.
- Sandler, S. J. (2005). Requirements for replication restart proteins during constitutive stable DNA replication in *Escherichia coli* K-12. *Genetics*, **169**, 1799–1806.
- Yu, K., Roy, D., Huang, F. T. & Lieber, M. R. (2006). Detection and structural analysis of R-loops. *Methods Enzymol.* **409**, 316–329.
- Rossi, M. L., Purohit, V., Brandt, P. D. & Bambara, R. A. (2006). Lagging strand replication proteins in genome stability and DNA repair. *Chem. Rev.* **106**, 453–473.
- Kelman, Z. & White, M. F. (2005). Archaeal DNA replication and repair. *Curr. Opin. Microbiol.* **8**, 669–676.
- Bruning, J. B. & Shamo, Y. (2004). Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-I. *Structure*, **12**, 2209–2219.
- Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B. & Tainer, J. A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, **116**, 39–50.
- Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A. & Tomkinson, A. E. (2007). The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase. *Nucl. Acids Res.* **35**, 1624–1637.
- Neduva, V. & Russell, R. B. (2005). Linear motifs: evolutionary interaction switches. *FEBS Letters*, **579**, 3342–3345.
- Neduva, V. & Russell, R. B. (2006). Peptides mediating interaction networks: new leads at last. *Curr. Opin. Biotechnol.* **17**, 465–471.
- Schurtenberger, P., Egelhaaf, S. U., Hindges, R., Maga, G., Jonsson, Z. O., May, R. P. et al. (1998). The solution structure of functionally active human proliferating cell nuclear antigen determined by small-angle neutron scattering. *J. Mol. Biol.* **275**, 123–132.
- Slootstra, J. W., Puijk, W. C., Ligtoet, G. J., Langeveld, J. P. & Muelen, R. H. (1996). Structural aspects of antibody-antigen interaction revealed through small random peptide libraries. *Mol. Divers.* **1**, 87–96.
- Sharp, R. J. & Raven, N. D. H. (1997). Isolation and growth of hyperthermophiles. In *Applied Microbial Physiology: A Practical Approach* (Rhodes, P. M. & Stanbury, P. F., eds), pp. 23–51, IRL Press, Oxford University Press, Oxford.
- Chapados, B. R., Chai, Q., Hosfield, D. J., Qiu, J., Shen, B. & Tainer, J. A. (2001). Structural biochemistry of a type 2 RNase H: RNA primer recognition and removal during DNA replication. *J. Mol. Biol.* **307**, 541–556.
- Allers, T., Ngo, H.-P., Mevarech, M. & Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* **70**, 943–953.

28. Bitan-Banin, G., Ortenberg, R. & Mevarech, M. (2003). Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. *J. Bacteriol.* **185**, 772–778.
29. Cline, S. W., Schalkwyk, L. C. & Doolittle, W. F. (1989). Transformation of the archaebacterium *Halo-*
- bacterium volcanii* with genomic DNA. *J. Bacteriol.* **171**, 4987–4991.
30. Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y. *et al.* (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science*, **288**, 2212–2215.

Edited by I. B. Holland

(Received 2 April 2007; received in revised form 14 June 2007; accepted 18 June 2007)
Available online 27 June 2007

1.2. *Pab2263* est un nouveau partenaire du PCNA

La protéine *Pab2263* a été identifiée comme nouveau partenaire du PCNA suite aux travaux précédemment exposés. On ne connaît pas la fonction exacte de cette protéine, cependant, le contexte génomique montre une proximité du gène *pab2263* avec ceux codant les deux sous unités de la polymérase D, et de la protéine Cdc6/Orc1. Ce voisinage pourrait indiquer que cette protéine intervient dans des mécanismes de maintenance génomique. L'analyse de sa séquence primaire suggère que cette protéine appartient à la famille RecB (Aravind *et al.*, 2000) et possède une activité nucléase. Cette activité a été confirmée et son substrat préférentiel a été mis en évidence, il s'agit de l'ADN simple brin (figure 30). D'autre part, les travaux de D. Flament ont permis de confirmer l'interaction de cette nouvelle nucléase avec le PCNA (figure 31, A). L'interaction entre ces deux composants semble former un complexe stable, puisque le K_D déterminé à partir des expériences SPR est de 15 nM. Enfin, j'ai pu montrer par *pull-down* que l'on trouve également cette interaction lorsque le PCNA fixé est incubé avec l'extrait cellulaire de *P. abyssi* (figure 31).

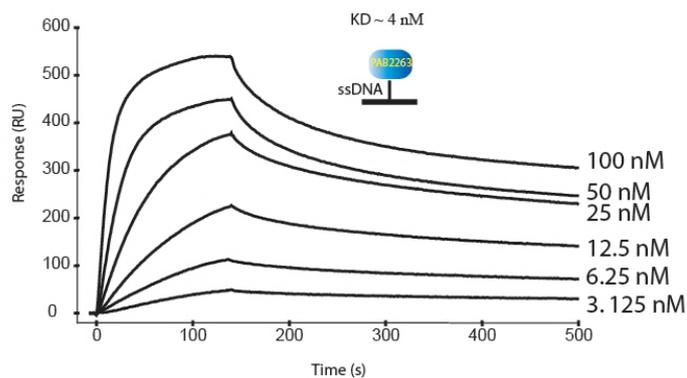


Figure 30 : expérience de SPR avec de l'ADN simple brin fixé sur la puce. Entre 3,125 nM et 100 nM de *Pab2263* ont été injecté sur la puceADN. L'augmentation du signal (unité RU) en fonction du temps indique l'interaction entre *Pab2263* et l'ADN simple brin.

Afin de déterminer la fonction de cette protéine, des approches conjointes de résolution de la structure 3D, de caractérisation biochimique et de recherche de partenaires d'interaction ont été menés en collaboration avec le Dr H. Myllykallio et le Dr R. Ladenstein (Karolinska Institute). Dans ce contexte, j'ai identifié comme partenaires potentiels de la protéine *Pab2263* : le PCNA (*Pab1465*), la grande sous-unité du RFC (*Pab0069*), la petite sous-unité du RFC (*Pab0068*), la sous-unité B de la topoisomérase VI (*Pab0407*), la sous-unité A de la topoisomérase VI (*Pab2411*), la protéine *Pab0588* qui présente une homologie avec des ADN méthylase, et une hélicase dépendante de l'ATP de la famille ERCC4 (*Pab0190*) (tableau 8 et figure 31, B.).

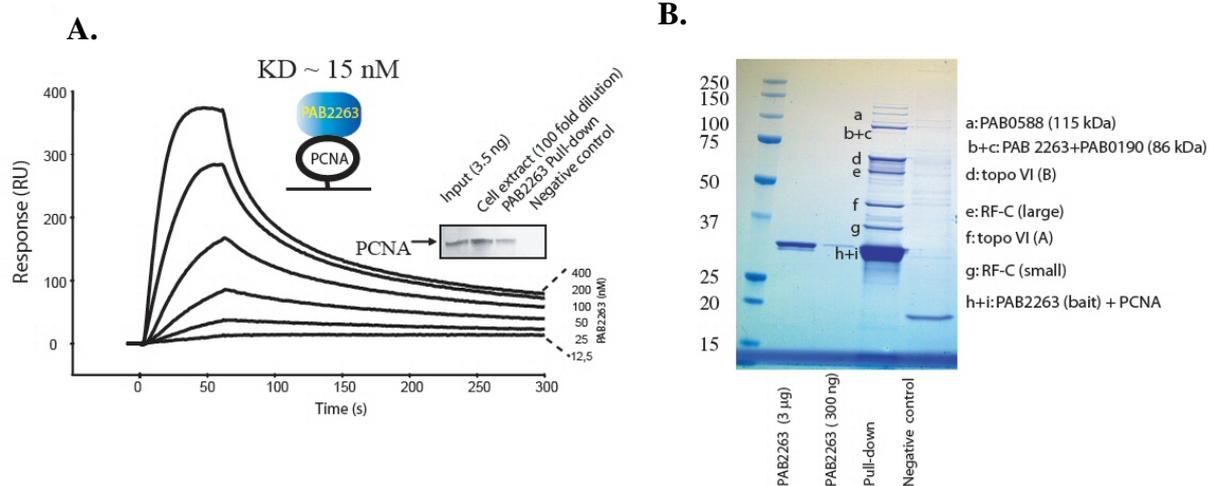


Figure 31 : A. Expérience de SPR et de *pull-down* démontrant l’interaction physique entre le PCNA et *Pab2263* fixé sur la puce. B. Identification des partenaires potentiels de *Pab2263*. Après expérience de *pull-down* dans l’extrait cellulaire, les complexes sont séparés par électrophorèse et révélés au bleu de Coomassie. Les partenaires identifiés par spectrométrie de masse sont indiqués.

Appâts	Partenaires interactants					
	Protéines	Accession number	Masse Moléculaire	Score Total	Peptides matched	% Coverage
<i>Pab2263</i> gi 5457559 (29 kDa)	PCNA	gi 5458797	28 017	102	14/99	57
	Replication factor C large subunit	gi 5457544	55 265	120	24/100	44
	Topoisomerase type VI, subunit B	gi 5458026	63 871	217	15/67	46
	Topoisomerase type VI, subunit A	gi 5457543	37 667	159	22/68	18
	RFC small subunit	gi 14520811	44 200	143		42
	<i>Pab0588</i>	gi 14521085	115 500	219		39,6
	<i>Pab0190</i> (Helicase)	gi 14520497	86 000	93,1		34

Tableau 8 : partenaires potentiels de *Pab2263*, identifiés par spectrométrie de masse.

L’analyse des partenaires identifiés semble indiquer un lien entre *Pab2263* et la maintenance de l’ADN. L’ADN topoisomérase interviendrait dans la dynamique de la structure de l’ADN, facilitant ainsi l’accès des protéines de la réparation ou de la recombinaison. L’hélicase de la famille ERCC4 pourrait intervenir dans l’ouverture de l’ADN au niveau d’une lésion. Les hélicases sont également connues pour intervenir dans le traitement de molécule d’ADN branché lors des processus de recombinaison. Le PCNA interviendrait dans la coordination des actions des différentes protéines.

Les protéines eucaryotes XPF en complexe avec la protéine ERCC1, présentent une activité nucléase et interviennent dans la réparation par excision de nucléotide (NER) et la recombinaison. La protéine XPF, dénommée à l’origine ERCC4, contient un domaine

nucléase ERCC4-like. Le complexe ERCC1-XPF clive de façon spécifique près d'une jonction entre de l'ADN simple brin et de l'ADN double brin (de Laat *et al.*, 1998).

La caractérisation biochimique de *Pab2263*, réalisé en collaboration avec l'équipe de H. Myllykallio, a montré que *Pab2263* interagit physiquement avec l'ADN simple brin (figure 30) et qu'elle est capable de couper des *flap* 5' ou 3' ADN au voisinage des jonctions ADN simple brin/ADN double brin. Le rôle physiologique de *Pab2263* reste encore à éclaircir, mais l'analyse de l'ensemble de ces données préliminaires nous indique que *Pab2263* est une nucléase qui interviendrait lors d'étapes intermédiaires de processus de réparation ou de recombinaison, au niveau des jonctions ADN simple brin et double brin, de manière similaire au complexe eucaryote ERCC1-XPF.

* *
*

1.3. Le complexe Mre11 et Rad50 est physiquement associé au PCNA

Parmi les partenaires potentiels du PCNA, les protéines Mre11 et Rad50 ont particulièrement suscité mon intérêt. En effet, cette interaction n'avait jamais été rapportée jusqu'à présent. Ce résultat a été reproduit en utilisant le PCNA comme appât dans trois expériences indépendantes. La figure 32 montre un exemple de spectre correspondant à l'empreinte peptidique de Mre11. Sur les 100 pics obtenus, 24 ont été identifiés comme correspondant à des peptides appartenant à la séquence de la protéine Mre11 de *P. abyssi*. Le pourcentage de recouvrement était de 57 % (figure 33).

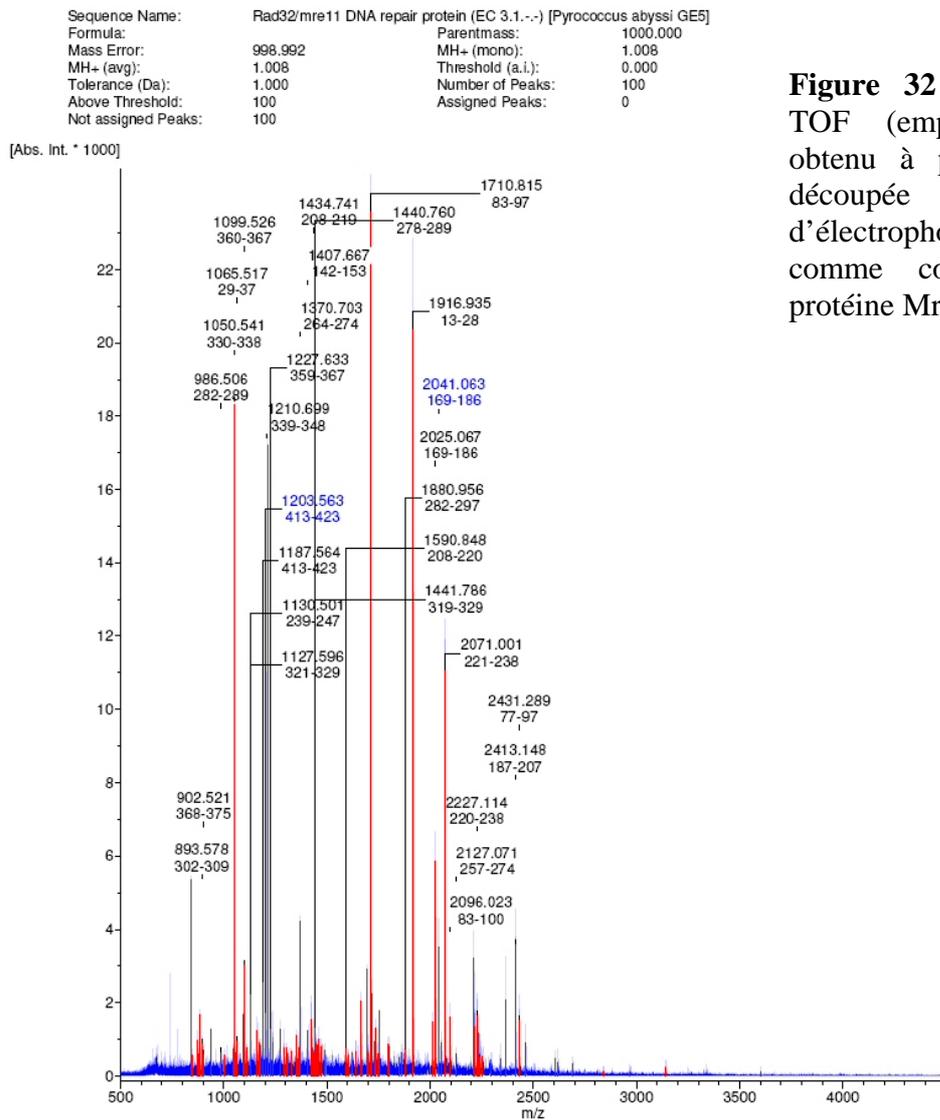


Figure 32 : spectre MALDI-TOF (empreinte peptidique) obtenu à partir de la bande découpée sur gel d'électrophorèse et identifiée comme correspondant à la protéine Mre11.

Sequence data:

Rad32/mre11 DNA repair protein (EC 3.1.-.-) [Pyrococcus abyssi GE5]

Intensity Coverage: 63.0 % (128012 cnts) Sequence Coverage MS: 57.7%
 Sequence Coverage MS/MS: 0.0% pl (isoelectric point): 5.9

10	20	30	40	50	60	70
MMTSEVFMVG	IKFAHLADVH	LGVEQFNRSQ	RAEFPKAFE	DAIKICVDEK	VDFIVIAAGDL	FNSSRPPSPGT
80	90	100	110	120	130	140
IKTAVKILQI	PRDNNIPVFA	IEGNHRTQR	GPSILHLED	LGLLYVLGVR	DEKVENEYLT	SEKTKAGWLIV
150	160	170	180	190	200	210
KGMYKDVEIH	GKMYMSAAF	EGNIELFKSM	FRPEGDAILV	LHOGVREITE	NNYPNYSSEL	SLSDLPKGYL
220	230	240	250	260	270	280
YYALGHIHR	FELTYDDAPV	VYPGSLERWD	FGDYSLKLTW	NGFQFKEEVG	VDKGFYIVED	FKPRFIRINA
290	300	310	320	330	340	350
RDFIDVHIKQ	HSENEIKKAV	KLAVPKIPRN	SYVRFNIRWK	RPVDVDWIKS	IVNAEYVRVN	PIIKEERGP
360	370	380	390	400	410	420
DGKDLVDKVF	FTELEWKIIE	LASEKEYEY	LNQIIDLLAS	EEGKVEAKID	AKREEKKFVR	PKNPGDIMAW
430						
VKG						

Sequence data:

Rad50 purine ntpase [Pyrococcus abyssi GE5]

Intensity Coverage: 68.5 % (49518 cnts) Sequence Coverage MS: 27.5%
 Sequence Coverage MS/MS: 0.0% pl (isoelectric point): 6.7

10	20	30	40	50	60	70	80
MKIEEVKVMN	FRSHEETVVR	FRKGINLITG	QNGSGKSL	DAILVGLVYS	KKLRLRGLKK	DEFRRIGGK	GTRIEIKFEN
90	100	110	120	130	140	150	160
DDSKYVLPFD	FSENVAYLKV	QENGRWPHAS	EPSMESVSSY	IERTLPYNVF	LNATVIRGGQ	IDAILESDET	RDVVREILN
170	180	190	200	210	220	230	240
LDRLESAYEN	LKRIKTNINL	LIESKKSPIA	RTENIEELIK	ANEDELTKKL	SEINEISSKL	PPIRGELEKV	RENVKELESI
250	260	270	280	290	300	310	320
RKISELKIQQ	VEKLRGRKMG	LEERIVQIER	SIEEKKAKIS	ELEEIVKDFE	KLQEKKEEYR	KLKGFDRDEY	SKLRRLKEL
330	340	350	360	370	380	390	400
SKWESELKAI	EEVIKEGEKK	KERAEIIEK	LSEIEKRLLE	LKPYVEELED	AKOVQKOIER	LKARLKGKLS	GEVIEKLESL
410	420	430	440	450	460	470	480
EKERTIEEA	IKEITTRIGQ	MEQEKNERMK	AIEELRKAG	KCPVCGRELT	EEHKKELMER	YTLIEKIEE	ELKRTTEER
490	500	510	520	530	540	550	560
KLRVNLKLE	IKLREFSVMR	DIAEQIKELE	SKLKFNLEE	LEQKREFEG	LNEEFNKLKG	ELGLEEDLK	RIKALEGRK
570	580	590	600	610	620	630	640
LIEEKVRKAK	EELNLRHQL	RELGFESVEE	LNLRIOELE	FHDKYVEAK	SESELRELKN	KLEKEKTELD	QAFEMLADVE
650	660	670	680	690	700	710	720
NEIEEKAEL	KDLESKFNEE	EYEEKRERLV	KLEREVSSLT	ARLEELKRSV	EQIKATLRLK	KEEKEREKA	KLEIKKLEKA
730	740	750	760	770	780	790	800
LSKVEDLRKK	IKDYKTLAKE	QALNRISEIA	SEIFSEFDG	KYSNVIVRAE	ENKTKLFFVY	EGKEVPLTFL	SGGERIALGL
810	820	830	840	850	860	870	880
APRLALSYL	VGRIDLLILD	EPTPFLDEER	RRELLDIMER	HLRRISQVIN	VSHDEELKDA	ADYVIELRL	GGKSKVEVVS
890							

Figure 33 : Alignement des séquences peptidiques identifiées en spectrométrie de masse avec les séquences des protéines Mre11 et Rad50. La coloration rouge indique les séquences peptidiques expérimentales correspondant à la protéine.

Nous avons choisi de produire et purifier les protéines Mre11 et Rad50 pour, d’une part valider l’interaction PCNA-Mre11/Rad50 *in vitro* avec les protéines recombinantes, et d’autre part identifier de nouveaux partenaires du complexe Mre11/Rad50. Ces analyses nous permettraient de mieux appréhender le rôle de Mre11 et Rad50 dans la cellule, mais aussi d’étudier l’interaction fonctionnelle avec le PCNA.

La partie suivante présente donc l’analyse des gènes *mre11* et *rad50* au niveau du génome de *P. abyssi*, puis les stratégies mises en œuvre pour le clonage, l’expression et la purification de ces deux entités.

1.3.1. Analyse des séquences primaires de Mer11 et Rad50

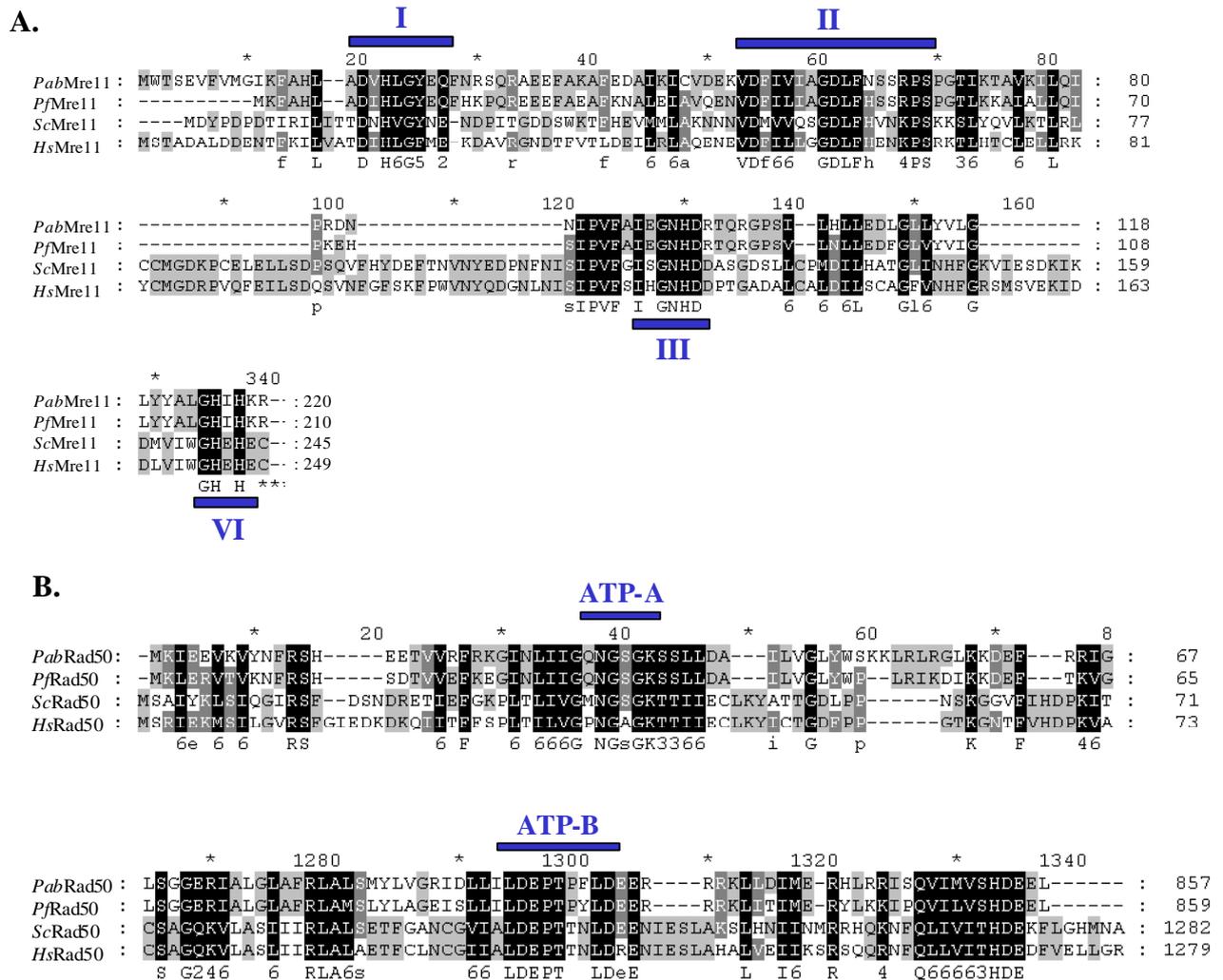


Figure 34 : alignement des séquences de *PabMre11* et *PabRad50* avec celles d’autres organismes. Les séquences ont été alignées avec le programme ClustalW et éditées avec GeneDoc (éditeur d’alignements multiples). (A) Alignement de *PabMre11* avec des homologues provenant de *Pyrococcus furiosus* (*PfMre11*), *Sacharomyces cerevisiae* (*ScMre11*), *Homo sapiens* (*HsMre11*). Les quatre domaines nécessaires à l’activité nucléase sont indiqués (I, II, III et IV). (B) Alignement de *PabRad50* avec des homologues provenant de *Pyrococcus furiosus* (*PfRad50*), *Sacharomyces cerevisiae* (*ScRad50*), *Homo sapiens* (*HsRad50*). Les deux domaines de liaison à l’ATP sont indiqués.

Le gène *pabmre11* de 1272 pb code une protéine de 423 acides aminés, de masse moléculaire 49 kDa. Le gène *pabrad50* de 2643 pb code une protéine de 880 acides aminés, de masse moléculaire 104 kDa. Les protéines homologues à Mre11 et Rad50 sont identifiés dans les trois domaines du vivant. La figure 34 présente l’alignement des séquences archées et les eucaryotes des protéines Mre11 et Rad50. Les quatre domaines proposés comme critique pour l’activité nucléase de Mre11 sont retrouvés chez *P. abyssi*. Les motifs *walker A* et *B* aux extrémités N- et C-terminales et permettant la liaison aux nucléotides sont également

conservés entre *PabRad50* et les homologues Archées et Eucaryotes (Sharples and Leach, 1995). Les activités biochimiques des protéines Mre11 et Rad50 de *P. furiosus* ont été caractérisées, elles sont similaires aux propriétés de leurs homologues dans les autres domaines du vivant (Hopfner *et al.*, 2000a). Elles présentent, respectivement, une activité endonucléase simple brin, et une activité exonucléase double brin, dépendante de l'ATP. Chez *P. abyssi*, les gènes codant Mre11 et Rad50 sont organisés en opéron et sont co-transcrits avec une nucléase NurA et une hélicase HerA (Constantinesco *et al.*, 2004). La figure 35 présente l'organisation génomique de ces quatre gènes dans le génome de *P. abyssi*.



Figure 35 : Organisation des gènes de *herA* (1647 pb) *mre11* (1272), *rad50* (2643), et *nurA* (1332 pb) dans le génome de *Pyrococcus abyssi*.

1.3.2. Expression et purification des protéines *PabMre11* et *PabRad50*

Afin de produire les protéines Mre11 et Rad50, j'ai choisi la stratégie de co-expression dans *E. coli* en utilisant le kit de clonage LIC DuetTM Adaptator (Novagen), comme indiqué dans le chapitre Matériels et Méthodes. Différentes souches et températures d'expression ont été testés à l'échelle pilote. Les conditions retenues étaient une induction à 15°C, avec 1 mM d'IPTG final, à DO₆₀₀ de 0,6. Les protéines ainsi produites ont été co-purifiées par chromatographie d'affinité sur colonne de nickel. En effet, Mre11 possède une étiquette histidine qui permet la purification du complexe comme indiqué sur la figure 36.

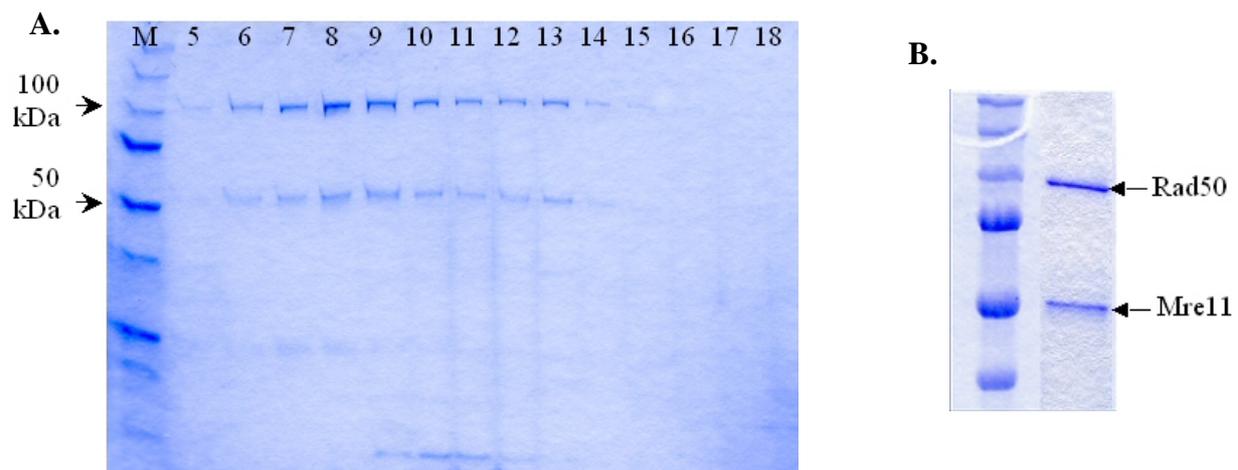


Figure 36 : Purification de *PabMre11* et *PabRad50*. A. Analyse par SDS-PAGE 12% d'une partie des fractions recueillies après purification par FPLC. B. Migration SDS-PAGE des fractions regroupées, contenant *PabMre11* et *PabRad50*, après les différentes étapes de dialyses.

1.3.3. Validation de l'interaction *in vitro* entre les protéines recombinantes PCNA et Mre11/Rad50

Afin de valider l'interaction entre le *Pab*PCNA et le complexe *Pab*Mre11/Rad50, des expériences de *pull-down in vitro* ont été réalisées à partir des protéines recombinantes. Le complexe Mre11/Rad50 était fixé sur les billes magnétiques couplées au cobalt (Dynabeads Talon, Invitrogen) par l'intermédiaire de la queue histidine de Mre11, et incubé avec un PCNA recombinant, non étiqueté. Le contrôle négatif correspondait à l'addition de ce PCNA sur les billes, en absence du complexe Mre11/Rad50. Le western blot, révélé grâce à un anticorps anti-PCNA, permet de mettre en évidence la co-précipitation du PCNA avec le complexe Mre11/Rad50 (figure 37). Cette étude ne permet cependant pas de discriminer laquelle des deux protéines Mre11 ou Rad50 interagit physiquement avec le PCNA.

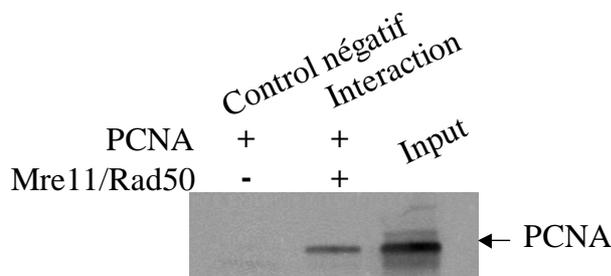


Figure 37 : visualisation par western blot de l'interaction entre le PCNA et le complexe Mre11/Rad50 de *P. abyssi*. 20 pmoles du complexe Mre11/Rad50 sont fixés sur des billes magnétiques par l'intermédiaire de la queue histidine de Mre11. Ce complexe billes/appât est ensuite incubé avec 30 pmoles de PCNA dans 1 mL de PBS. Après lavage, les protéines qui co-précipitent sont séparées, transférées sur membrane et révélées par un anticorps anti PCNA.

Ce travail constitue, à ma connaissance, la première démonstration d'une association physique entre le PCNA et le complexe Mre11/Rad50 chez les archées. Si aucune interaction n'est démontrée à ce jour, ni chez les archées, ni chez les eucaryotes, l'analyse bibliographique nous indique que le complexe Mre11/Rad50 a été co-localisé chez les cellules de mammifères, lors de la phase S du cycle cellulaire qui correspond à la phase de synthèse de l'ADN (Maser *et al.*, 2001). Chez les eucaryotes, le complexe Mre est composé des trois protéines : Mre11, Rad50 et Nbs1. Maser et al. montrent que le complexe Mre est physiquement associé aux protéines de la famille E2F via la région N-terminal de Nbs1, et que cette interaction a lieu près de l'origine de réplication. E2F est un régulateur de la transcription qui régule les gènes intervenant dans la réplication de l'ADN. Les auteurs énoncent que cette interaction interviendrait dans le point de contrôle de la phase S. La colocalisation du PCNA avec le complexe Mre lors de la phase S semble indiquer que le complexe a également un rôle au niveau des fourches de réplication de l'ADN. Comme la recombinaison et la réplication sont intimement liées, Maser et al. proposent que le complexe Mre pourrait intervenir dans la résolution des fourches de réplication bloquées ou endommagées. Par ailleurs, chez les cellules de mammifères synchronisées, une étude indique

que le complexe Mre est chargé sur la chromatine, lors de la phase S du cycle cellulaire (Mirzoeva and Petrini, 2003), ce qui suggère que la fourche de réplication serait le site de ce chargement. Le complexe Mre pourrait promouvoir l'association des chromatides sœurs et faciliter la recombinaison. Des analyses cytologiques ont également démontré que le complexe Mre11 était capable de se lier au cassure double brin de l'ADN après leur formation, impliquant ainsi ce complexe dans la détection de l'ADN endommagé (Limoli *et al.*, 2002; Mirzoeva and Petrini, 2001). Le complexe Mre paraît ainsi avoir plusieurs rôles physiologiques en relation avec la fourche de réplication, comme la recombinaison et la détection des lésions de l'ADN. Chez les eucaryotes, le complexe Mre associés aux protéines ATM et ATR, participent à la stabilité du génome. Il a été démontré que, lors de la réplication de l'ADN, le redémarrage des fourches bloquées est impossible en absence de Mre11 (Trenz *et al.*, 2006). Ainsi nos données pourraient suggérer que chez les archées, le PCNA ait un rôle dans le recrutement et la dynamique du complexe Mre au niveau de la fourche de réplication, et ceci afin de pouvoir intervenir au niveau de cassures double brin engendrés lorsque le réplisome rencontre une coupure sur l'ADN (figure 38).

Afin d'imaginer de nouveaux rôles du complexe Mre chez les archées, la recherche des partenaires de ce complexe a été entreprise.

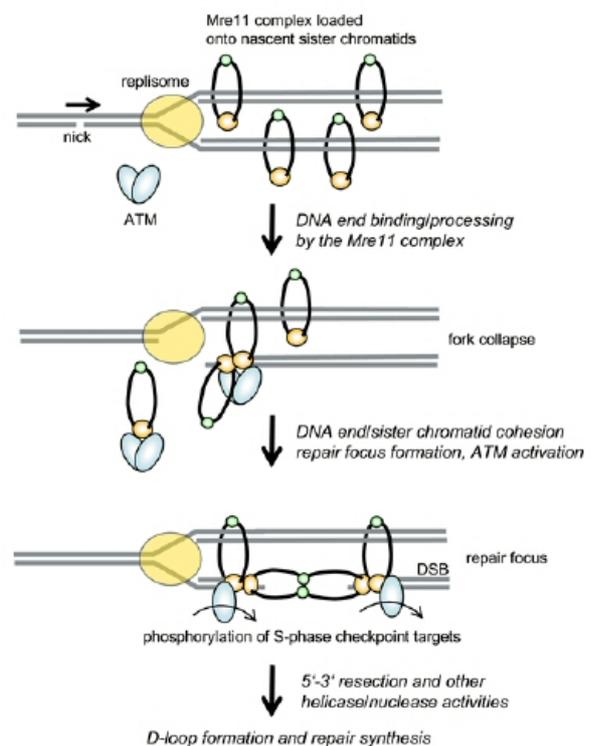


Figure 38 : model hypothétique montrant le lien entre le complexe Mre et la réplication de l'ADN, lorsque le réplisome rencontre une coupure simple brin sur l'ADN. D'après (Assenmacher and Hopfner, 2004).

1.3.4. Recherche de partenaires du complexe Mre11/Rad50

La recherche des partenaires du complexe Mre11/Rad50 dans l'extrait cellulaire total de *P. abyssi* a été entreprise par la technique de *pull-down*. 7,5 µg (25 µL) de Mre11/Rad50 recombinants ont été fixés sur 25 µL de billes magnétiques (Dynabeads Talon), puis mis en présence de 4 mL d'extrait cellulaire. Après incubation et lavage extensif, comme décrit préalablement (chapitre 3.7 de Matériels et méthodes), les complexes élués sont séparés et visualisés par électrophorèse SDS-PAGE 12%, coloré au bleu de Coomassie (figure 39).

Les protéines ainsi capturées et identifiées en spectrométrie de masse sont présentées dans le tableau 9. Les protéines identifiées comme étant potentiellement des partenaires du complexe Mre11/Rad50 étaient : le PCNA (*Pab1465*), l'hélicase HerA (*Pab0810*), la nucléase NurA (*Pab0813*), une ATPase de la famille PiIT (*Pab1633*), une protéine hypothétique (*Pab1110*).



Figure 39 : Identification des partenaires potentiels du complexe Mre11/Rad50 par *pull-down*. La piste 1 correspond au contrôle négatif, la piste 2 correspond à l'expérience de *pull-down* avec le complexe Mre11/Rad50 comme appât (les protéines Mre11 et Rad50 ont été identifiées en spectrométrie de masse et sont indiquées par un point rouge). Les flèches noires correspondent aux partenaires potentiels. La piste 3 correspond aux protéines Mre11 et Rad50 recombinantes.

Appât	Partenaires interactants					
	N° spot	Protéines	Accession number	Molécule r Mass	Total Score	%Coverage
Mre11/Rad50	1	Twitching mobility (pilt) related protein	gi 514521315	68 700	150	47
	2	ATPase her A Pab0810	gi 14521423	61 400	178	55,5
	3	NurA Pab0813	gi 14521426	37 667	131	46,3
	4	PremRNA splicing protein Nop58plike	gi 14520278	47 200	121	50,7
	5	Pab1110 homol Asnc family	gi 14521893	36 600	119	50,2
	6	PCNA	gi 14521578	28 017	132	68,3

Tableau 9 : partenaires potentiels de Mre11/Rad50 identifiés par spectrométrie de masse.

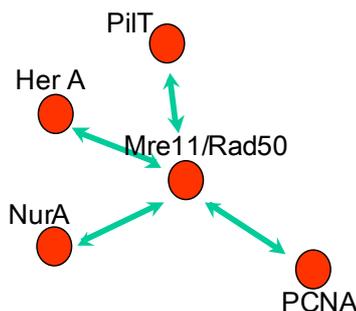


Figure 40 : Réseau d'interaction des partenaires potentiels du complexe Mre11/Rad50 mis en évidence par *pull-down*.

1.3.4.1. Le PCNA est associé au complexe Mre11/Rad50

Nous retrouvons ici le PCNA, capturé dans l'extrait cellulaire, en utilisant cette fois le complexe Mre11/Rad50 comme appât. J'ai pu à nouveau confirmer, par cette approche inverse, l'association du PCNA avec Mre11/Rad50.

1.3.4.2. Les protéines NurA et HerA, partenaires du complexe Mre11/Rad50

Les protéines NurA et HerA ont été caractérisées chez *Sulfolobus acidocaldarius* (Constantinesco *et al.*, 2002; Constantinesco *et al.*, 2004). HerA a également été caractérisé, sous le nom de MlaA par (Manzan *et al.*, 2004). HerA est une ATPase de la famille PilT/VirD4. Les protéines de cette famille d'ATPase sont impliquées dans divers processus de transport et notamment celui de l'ADN lors de la division cellulaire. HerA est une ADN hélicase capable de fonctionner dans les deux sens, 5' vers 3', et 3' vers 5'. NurA présente une activité endonucléase sur l'ADN circulaire et une activité exonucléase 5' vers 3' sur l'ADN linéaire simple brin et double brin. Les gènes *nurA* et *herA* sont détectés dans les génomes de toutes les Archées hyperthermophiles séquencés et sont organisés en opéron avec les gènes *mre11* et *rad50*. Constantinesco *et al.* ont également montré que les gènes codant les quatre protéines, Mre11, Rad50, NurA et HerA était co-transcrites. Cette organisation génomique suggère que ces protéines sont associées physiquement et interviennent dans une voie métabolique commune. Nos travaux semblent confirmer que NurA et HerA sont des partenaires du complexe Mre11/Rad50. Constantinesco *et al.* proposent que ces quatre protéines sont des analogues fonctionnels du complexe enzymatique multifonctionnel de *E. coli* RecBCD. En effet, ce complexe présente à la fois des activités hélicases et exonucléases, impliquées dans l'initiation de la recombinaison. Ainsi NurA et HerA pourrait participer avec le complexe Mre11/Rad50 à l'initiation de la recombinaison homologue dans le processus de réparation des cassures double brin.

1.3.4.3. La protéine Pab1633 de la famille PilT

La protéine Pab1633 est annotée comme protéine putative non caractérisée dans le génome de *P. abyssi*. Après analyse de sa séquence et recherche d'homologie dans les bases de données, la protéine Pab1633 apparaît être une ATPase de la famille PilT. Elle possède un domaine ATPase de la famille AAA, un domaine PINc (PilT N terminus) et un domaine KH (K homology), tous deux susceptibles de permettre la liaison à des acides nucléiques (figure 41).

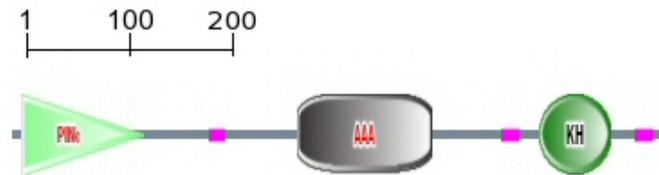


Figure 41 : les domaines de la protéine *Pab1633* ont été recherchés grâce à l’outil informatique SMART (*Simple Modular Architecture Research Tool*) (Schultz *et al.*, 1998). Les domaines PINc, AAA+ et KH sont indiqués. L’échelle indique le nombre d’acide aminé, *Pab1633* comporte 608 acide aminé.

Le domaine PIN est associé à une fonction exonucléasique. Chez les eucaryotes, il est associé à des ribonucléases intervenant dans la dégradation des ARN (Clissold and Ponting, 2000). Chez les archées et les bactéries thermophiles, il a été proposé que le domaine PIN avait un rôle dans la réparation de l’ADN (Makarova *et al.*, 2002). Un domaine PIN a été caractérisé chez l’archée *Pyrobaculum*, il a été démontré que ce domaine présente une activité exonucléase sur l’ADN (Arcus *et al.*, 2004). Il est ainsi possible de suggérer une implication de la protéine *Pab1633* dans des mécanismes de réparation.

Les domaines KH présentent habituellement un repliement permettant la liaison aux acides nucléiques. Ils sont retrouvés généralement dans les RNases constituant l’exosome, dans des PNPase (polynucléotide phosphorylase) et dans des protéines ribosomales.

Les activités ATPases et exonucléases potentielles de *Pab1633* pourraient être associés aux mécanismes de recombinaison et réparation impliquant les autres protéines Mre11, Rad50, NurA et HerA. D’autre part, les protéines de la famille PilT peuvent être impliquées dans des phénomènes de transfert d’ADN d’une cellule à l’autre. Il est ainsi possible d’imaginer un lien entre les protéines de la recombinaison et le mécanisme de conjugaison ce qui permettrait d’échanger le matériel génétique entre les cellules pour réparer l’ADN endommagé.

1.3.4.4. La protéine *Pab1110*

Après recherche d’alignement avec les bases de données, La protéine *Pab1110* semble être conservée uniquement chez les autres espèces de thermococcales. Elle présente toutefois une faible homologie avec des régulateurs de la transcription de la famille Asnc intervenant chez les bactéries dans le contrôle de l’expression des gènes. Elle possède un motif HTH (*Helix-Turn-Helix*) est responsable de la liaison spécifique à l’ADN. On pourrait envisager que cette protéine intervient dans la régulation de l’expression du complexe Mre11/Rad50. Il ne semble cependant, à priori, pas avoir de rapport entre cette famille Asnc de régulateur de la transcription, et le régulateur E2F qui régule l’expression des gènes intervenant dans la réplication chez les eucaryotes (voir chapitre 1.3.3.).

Nous avons mis en évidence que le PCNA interagit avec le complexe ubiquitaire Mre11/Rad50. Le complexe Mre11/Rad50 est connu, chez les eucaryotes, pour intervenir dans la reconnaissance des lésions de l'ADN et dans des processus de réparation et de recombinaison de l'ADN (D'Amours and Jackson, 2002). Lors de ce travail, nous avons également identifié des partenaires potentiels du complexe Mre11/Rad50 au sein de l'extrait cellulaire. Parmi ceux-ci, nous retrouvons le PCNA et deux protéines, l'endonucléase NurA et l'hélicase HerA. Ces gènes, codant ces deux protéines, sont organisés en opéron avec ceux codant les protéines Mre11 et Rad50 (Constantinesco *et al.*, 2004). Cette organisation suppose l'implication de ces quatre protéines dans des mécanismes communs. Le PCNA, identifié en complexe avec ces quatre protéines, pourrait également intervenir dans ces mécanismes. Nos résultats sont encore préliminaires, mais ils nous permettent d'imaginer un rôle du complexe PCNA-Mre11-Rad50, associé ou non avec les protéines NurA et HerA, dans le maintien de l'intégrité génomique durant le processus de réplication de l'ADN. Nos données suggèrent que le PCNA pourrait par exemple recruter le complexe Mre11/Rad50 sur des sites d'ADN endommagés. Le PCNA permettrait le couplage entre la réplication et la réparation des cassures double brin.

* *
*

II. Les protéines Fen1 et l'ADN primase interviennent dans des mécanismes de réparation de l'ADN

Avec le PCNA, la protéine Fen1 est un acteur majeur du métabolisme de l'ADN (Liu *et al.*, 2004). Chez les eucaryotes, Fen1 est nécessaire à la stabilité du génome et peut être considéré comme un suppresseur de tumeur (Henneke *et al.*, 2003). Les différentes fonctions de Fen1 seraient régulées par la formation de complexe avec différents partenaires protéiques (Shen *et al.*, 2005). Au cours de cette étude, j'ai utilisé la protéine Fen1 recombinante comme appât pour capturer des partenaires dans le lysat cellulaire de *P. abyssi*.

Fen1 est une endonucléase structure spécifique capable de dégrader un brin d'ADN simple brin flottant (structure *flap*) résultant de l'activité de déplacement de brin d'une ADN polymérase. Cette structure est un intermédiaire important lors du métabolisme de l'ADN. Fen1 est capable de reconnaître et de cliver cette structure, lors de la maturation des fragments d'Okazaki durant la réplication de l'ADN, mais également lors du mécanisme de réparation par excision de base (BER), chez les eucaryotes. Les travaux présentés précédemment mettent également en évidence la participation de Fen1 dans des mécanismes de réparation chez les archées (Matsui *et al.*, 1999; Meslet-Cladiere *et al.*, 2007).

Trois partenaires potentiels ont été identifiés en spectrométrie de masse, en utilisant Fen1 comme appât (figure 42) :

- Le PCNA (*Pab1465*), cette interaction a déjà été évoqué au chapitre précédent.
- La sous unité catalytique de la primase (41 kDa) (*Pab2236*). Cette interaction a fait l'objet d'une caractérisation détaillée que j'exposerais au chapitre suivant. Cette interaction apporte de nouvelles perspectives sur les mécanismes de réparation chez les *Archaea*.
- La sous-unité B de la topoisomérase VI (*Pab0407*).

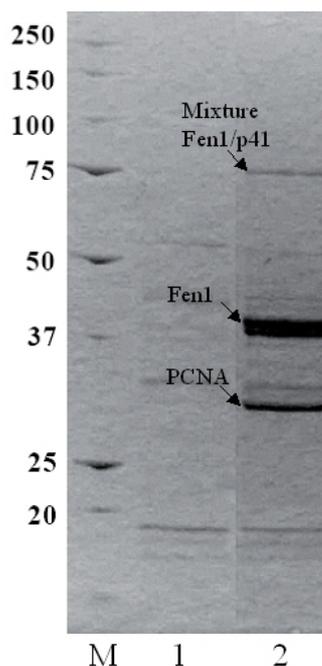


Figure 42 : Identification des partenaires potentiels de Fen1. La piste 1 correspond au contrôle négatif, la piste 2 correspond à l'expérience de *pull-down* avec Fen1 comme appât. Les flèches noires correspondent aux partenaires potentiels.

2.1. Association physique entre *PabFen1* et l'ADN primase de l'Archaea hyperthermophile *Pyrococcus abyssi* – implication pour des mécanismes de réparation

Les ADN primases sont connues pour intervenir dans l'initiation de la réplication de l'ADN en synthétisant une amorce indispensable aux ADN polymérases qui ne peuvent pas démarrer la polymérisation de l'ADN *de novo*. Les ADN primases d'archées ont la faculté de synthétiser de l'ADN et de se comporter de façon similaire à des ADN polymérases sur certains substrats (Liu *et al.*, 2001). Une étude précédente a démontré la capacité de l'ADN primase de *P. abyssi* à combler une brèche et à déplacer un brin d'ADN (Le Breton *et al.*, 2007). D'autre part, l'ADN primase de *Sulfolobus* présente de plus une activité nucleotidyl-transférase (De Falco *et al.*, 2004). Ces observations suggèrent une implication des ADN primases d'archées dans des processus autres que l'initiation de la réplication. Par ailleurs, les analyses structurales révèlent un lien entre l'architecture de la sous-unité catalytique des primases d'archées et les ADN polymérases de la famille X des eucaryotes. L'ADN polymérase β , membre représentatif de la famille X des ADN polymérases est impliqué dans des processus de réplication, réparation et recombinaison de l'ADN. Son rôle principal intervient dans la réparation par excision de base (BER) en collaboration avec Fen1 (Prasad *et al.*, 2000). Aucune polymérase de la famille X n'a été identifiée dans les génomes de *Pyrococcus*. Par analogie, il est ainsi tentant de suggérer une association entre l'ADN primase et Fen1 chez les archées, intervenant également dans des mécanismes liés au BER.

Ces hypothèses nous ont conduits à tester l'influence de Fen1 sur les capacités de synthèse d'ADN de p41 seul ou en complexe avec p46, la sous-unité régulatrice, en présence de différents substrats. Ces substrats correspondaient à une matrice ADN amorcée à élonger ou une brèche à remplir. Nous avons observé que Fen1 était capable de stimuler l'activité de synthèse d'ADN de p41 sur une matrice amorcée, mais également de stimuler l'activité de colmatage de brèche et de déplacement du brin aval. Par ailleurs l'interaction physique entre p41 et Fen1 a été démontrée grâce à la résonance plasmonique de surface et à la co-immunoprécipitation.

Afin de donner un sens fonctionnel à l'interaction, la collaboration des protéines p41, Fen1 et Lig1 a été étudié pour une étape d'un mécanisme de réparation de l'ADN.

**Physical association between *PabFen1* and DNA primase from hyperthermophile
Archaea Pyrococcus abyssi – implication for DNA repair mechanisms.**

Julien Briffotiaux^{1,2}, Magali Le Breton^{1,2}, Laurence Meslet-cladière³, Anne Godfroy^{1,2}, Audrey Le Goaziou^{1,2}, Ghislaine Henneke^{1,2}, Régis Lavigne⁴, Charles Pineau^{4,5}, Hannu Myllykallio³, Joël Querellou^{1,2}, Jean-Paul Raffin^{1,2} and Didier Flament^{1,2,*}.

Submitted at *Journal of Molecular Biology*

1. Ifremer, UMR6197, Laboratoire de Microbiologie des Environnements Extrêmes, BP 70, F-29280 Plouzané, France.
2. Université de Bretagne Occidentale, UMR6197, Laboratoire de Microbiologie des Environnements Extrêmes, 29280 Plouzané, France.
3. Institut de Génétique et Microbiologie, Université Paris-Sud, UMR8621, 91405 Orsay, France
4. High-Throughput Proteomics Platform OUEST-genopole, 263 Avenue du Général Leclerc, Campus de Beaulieu, Rennes, F-35042, France.
5. Inserm, U625; IFR-140; Université de Rennes I, Campus de Beaulieu, Rennes cedex, F-35042 France.

Running title : interaction between Archaeal Fen1 and DNA primase.

Corresponding author: Didier Flament, Ifremer, Laboratoire de Microbiologie des Environnements Extrêmes, BP 70, 29280 Plouzané, France.

Phone: 33 (0)2 98 22 45 27

Fax : 33 (0)2 98 22 47 57

e-mail : dflament@ifremer.fr

Keywords: Hyperthermophilic archaea; DNA primase; DNA repair; Fen1.

Abstract

The Hyperthermophilic archaeon *Pyrococcus abyssi* possesses three DNA polymerases belonging to the Pol B, Pol D and Archaeo-Eukaryotic Primase families. We previously reported that the two members from family B and D are likely to be replicative polymerases. We have also demonstrated that the DNA primase from *P. abyssi* display unique features and exhibits strand displacement and gap filling activity, raising the possibility that this polymerase may have dual functionality in both primer initiator synthesis and DNA repair mechanisms. Here we report evidence for the physical interaction between *Pabp41* and *PabFen1* by pull-down/MS, SPR and co-immunoprecipitation assays. *PabFen1* enhance the primer elongation activity of *Pabp41*, alone or in complex with its regulatory subunit *Pabp46*, on nick and various gapped substrates. The complex *PabFen1/p41* exhibit the ability to resolve BER intermediates via nick translation or strand displacement and subsequent flap cleavage. The implications of these results for DNA repair in Archaea are further discussed.

Introduction

Living organisms use specific macromolecular complexes to respond to DNA damage caused by intra- and extracellular sources. The large variety of lesions that can arise in DNA and the biological importance of coping with it, reflect the diversity of biochemical strategies for DNA repair that have been found in well studied model micro-organisms. In this context, hyperthermophilic Archaea are of special interest given that DNA in hyperthermophiles is exposed to temperatures that increase the rate of DNA decomposition reactions (depurination and deamination, (Lindahl, 1993)). However, the DNA repair mechanisms in Archaea remains poorly described. Genomic sequences analyses suggest that Archaea have repair proteins that are generally more similar to their eucaryotic than bacterial equivalents and point to the lack of the otherwise well conserved mismatch repair pathway (Grogan, 2004;

Kasiviswanathan *et al.*, 2005). In addition, although Archaea have homologs of many of the eukaryal proteins involved in Nucleotide Excision Repair (NER), damage recognition proteins are still missing (Kasiviswanathan *et al.*, 2005) and up to now, the only biochemical evidence of NER in Archaea has been found in *Methanobacterium thermoautotrophicum* which has a complete set of laterally transferred UvrABC genes (Ogrunc *et al.*, 1998).

Another conspicuous feature highlighted by genomic analysis is the relative paucity of DNA polymerases dedicated to repair functions encoded in archaeal genomes. Members of the Crenarchaea phylum possess at least two and up to three representatives of the ubiquitous family B (Cann and Ishino, 1999; Iwai *et al.*, 2000; Uemori *et al.*, 1995) and in some strains, also have one family Y DNA polymerase (Kulaeva *et al.*, 1996). In contrast, Euryarchaea possess only one family B and one heterodimeric family D DNA polymerase (Cann *et al.*, 1998; Gueguen *et al.*, 2001). Based on the biochemical behaviour of *Pyrococcus abyssi* D and B polymerases, we previously provided evidence that both polymerases are replicative and proposed a model in which Pol B synthesizes the leading strand and Pol D replicates the lagging strand (Henneke *et al.*, 2005). In crenarchaea *Pyrobaculum aerophilum*, based on biochemical properties and sequence analysis, it has been proposed that in PolB2, a functional homologue of mammalian Pol β , participates in BER process, whereas PolB1 and PolB3 might be the replicative polymerases (Böhlke *et al.*, 2002; Sartori and Jiricny, 2003). With the exception of *Methanobacterium thermoautotrophicum*, Archaea do not possess Pol X family proteins which representative members are known to have implications in DNA metabolism, including replication but also repair and recombination (Ramadan *et al.*, 2004). The lack of specialized DNA polymerases also contrast with the life style of these micro-organisms at high temperature. Recent structural studies have proposed suggested an architectural conservation, via convergent evolution, between Pol X family members and the small subunit of Archaeal and Eukaryotic primases (Kirk and Kuchta, 1999; Lao-Sirieix *et al.*, 2005b).

Furthermore biochemical studies have revealed the unique properties of archaeal primases, such as DNA polymerase and 3'-nucleotidyl terminal transferase activities (Bocquier *et al.*, 2001; De Falco *et al.*, 2004; Lao-Sirieix and Bell, 2004; Le Breton *et al.*, 2007; Liu *et al.*, 2001)), leading to the proposition that archaeal DNA primase might play a role in DNA repair (Lao-Sirieix *et al.*, 2005b).

Here, using Pull-down/MS, Surface Plasmon Resonance and co-immunoprecipitation experiments, we report evidence that the small catalytic subunit of the DNA primase physically interacts with the endonuclease Fen1 from *P. abyssi*. We also demonstrate a cooperative interaction between *Pab*Fen1 and *Pab* DNA primase and propose the involvement of this new complex in Base Excision Repair mechanisms. These results give additional insight into genome stability processes in hyperthermophilic Archaea.

Experimental Procedures

Recombinant proteins

Fen1 (*PAB1877*), DNA ligase (*PAB2002*), DNA primase small subunit p41 (*PAB2236*) and DNA primase large subunit p46 (*PAB2235*) were expressed and purified as described previously (Le Breton *et al.*, 2007; Meslet-Cladiere *et al.*, 2007).

DNA substrates

The oligonucleotides used for substrates (Table I) were 5'-FAM or 3'-Cy5 end-labelled and assembled into 25-nt gap, 5-nt gap and nick substrates onto M13mp18 DNA template as described previously (Henneke *et al.*, 2005).

Enzyme assays

Enzymatic assays (10 μ l) were performed in 50 mM Tris-HCl pH 8, 1 mM β -mercaptoethanol, 5 mM MgCl₂, 0.1 mM dNTPs containing 0.04 pmol labelled substrate and 0.1 to 10 pmol of each purified recombinant proteins. When *Pab* DNA ligase was present, the reaction

mix was supplemented with 0.01 mM ATP. After incubation at 60°C for 20 min, the reaction was stopped with 10 µl of stop buffer (98% formamide, 10 mM EDTA). The reaction products were denatured for 10 min at 95°C, and separated on denaturing 15% polyacrylamide gel containing 7 M urea and 3.075 M formamide. The bands were visualized on a laser-based imaging system (Typhoon 9400, GE Healthcare).

Surface Plasmon resonance analysis

Surface Plasmon resonance (SPR) was performed at 25 °C in a BIAcore X apparatus (BIAcore, Uppsala Sweden). *PabFen1* was amine-coupled on the surface of CM5 sensor chips (BIAcore) in 10 mM sodium acetate buffer (pH 4.5) using the amine coupling kit and according to the manufacturer instructions. The level of immobilisation of *PabFen1* was of about 2000 RU. Various concentrations of *Pabp41* were injected at a flow rate of 30 µl/min in HBS-P buffer (10 mM Hepes pH 7.4, 150 mM NaCl, 0.005% (v/v) P20). The chip was regenerated between successive interaction assays, with a pulse of 50 mM NaOH (8 µl at 10 µl/min). The data were fitted to a 1:1 interaction model with drifting baseline using the BIAevaluation software (V3.2).

Preparation of cell extracts

P. abyssi GE5 (strain orsay) was anaerobically grown at 90°C in a 2 liters gaz-lift bioreactor in SME medium pH 6,8 (Godfroy *et al.*, 2006). During the continuous culture, cell suspensions in exponential phase ($2 \cdot 10^8$ cells/mL) were harvested and centrifuged at 5000 g for 45 min. Cells were resuspended in 50 mM Tris-HCl buffer pH 8.0 containing 0.5 mM dithiothreitol and 150 mM NaCl. The cells were lysed by sonication using a vibracell sonifier (15 pulses of 30 seconds, 375 W, 40% amplitude). Cellular debris were removed by centrifugation at 10 000 g for 20 min. The supernatant was used as the whole-cell extract. Protein concentration was determined using the Bio-Rad DC protein assay, with BSA as the standard.

Pull-down assays

20 µg of HIS-tagged recombinant proteins were incubated with 1 mg of magnetic micro beads (Dynabeads® TALON™, Dynal Biotech) in 1 ml of binding and washing buffer (WB buffer: 50 mM Na-phosphate pH 8, 300 mM NaCl, 0.01% tween 20) for 20 min at 25 °C. HIS-proteins bound to the beads were washed three times in the same buffer. The beads were then incubated with 12.5 mg of whole-cell extract for 2 h at 4°C. After five washing steps with 4 ml of WB buffer, proteins bound to the target proteins were resuspended in 10 µl of WB buffer.

Co-immunoprecipitation

20 µL of Dynabeads Protein A (Dynal®Biotech) was conjugated with anti-*PabFen1* antibodies in 1 mL of PBS buffer (140 mM NaCl, 2.7 mM KCl, 4 mM Na₂HPO₄, 1.5 mM K₂HPO₄) for 20 minutes. After three washing steps with 1 mL of PBS, 0.5 µg of *PabFen1* and/or *Pabp41* was added to the Dynabeads protein A conjugated antibody and incubated for 2 hours at 4°C. The beads of each mixture was then washed with 4 mL of PBS buffer and re-suspended in 30 µL of elution buffer (0.1 M citrate, pH 3). Proteins were eluted from the beads in loading buffer (XT sample buffer, Biorad) and examined by SDS-PAGE and western blotting.

Western blot analysis

After electrophoresis on Criterion XT 12% polyacrylamide precast gel (Biorad), proteins were transferred to nitrocellulose membrane. The membrane was incubated with blocking buffer containing 5% (wt/vol) milk powder in TBS-T buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% (vol/vol) tween-20) for 30 minutes. The membrane was then incubated with the primary polyclonal antibody (anti-*PabFen1* and anti-*Pabp41*) at 1/10 000 dilution, in TBS-T for 1 h at room temperature. After three washing steps in TBS-T, the membrane was incubated with secondary antibody (ECL anti-rabbit IgG, horseradish peroxydase-linked,

Amersham) diluted 1/20 000 overnight at 4°C with gentle agitation. The membrane was washed three times and the detection was carried out by the enhanced chemiluminescence system with ECL Plus western blotting detection system (GE Healthcare) according to the manufacturer's instructions.

SDS-PAGE and gel staining

Immunoprecipitated proteins were resuspended in Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 5% dithiothreitol, and 0.002% bromophenol blue) and boiled for 5 min before analysis on NuPAGE 12% polyacrylamide precast gels (Invitrogen). For colloidal Coomassie staining, polyacrylamide gels were fixed overnight in solution containing 50% ethanol and 1.4% orthophosphoric acid. After three washes for 30 min in MilliQ water, gels were incubated in impregnation solution (1.3 M ammonium sulfate, 34% methanol, and 1.4% orthophosphoric acid) for 1 h and placed in staining solution (1.3 M ammonium sulfate, 34% methanol, 1.4% orthophosphoric acid, and 0.07% Coomassie Brilliant Blue G250) for 24 h. Finally gels were destained with several washes of MilliQ water until the background was clear.

In-gel Trypsin Digestion

Coomassie Blue-stained protein bands were excised from the SDS-PAGE gel and processed for trypsin digestion as described previously (Com *et al.*, 2003) with minor modifications. In brief, gel bands were destained twice for 10 min in a solution containing 100 mM NH_4HCO_3 and 50% acetonitrile, dehydrated in acetonitrile, and dried in a vacuum centrifuge. Gel pieces were then rehydrated at 4 °C for 45 min in a digestion buffer (50 mM NH_4HCO_3 and 12.5 ng/l trypsin). The supernatant was replaced by 40 μl of 50 mM NH_4HCO_3 , and the samples were incubated overnight at 37 °C. The tryptic peptides were recovered by 10-min incubations, once in 25 mM NH_4HCO_3 and 50% acetonitrile and twice in 25 mM NH_4HCO_3 , 50%

acetonitrile, and 5% formic acid. All supernatants were pooled and dried in a vacuum centrifuge.

Protein identification by Mass Spectrometry

For MALDI-TOF analyses, the digested peptides contained in the dried supernatant were purified using a home-made micropurification column containing 0.1 μ L of a slurry consisting of 20R2 reversed-phased material (PerSeptive Biosystems, Framingham, MA) packed in a GeLoader tip (Eppendorf, Hamburg, Germany) and equilibrated with 1% trifluoroacetic acid as described previously (Com *et al.*, 2003). Ten microliters of the mixture were loaded onto the column, and after washing with 1% trifluoroacetic acid, adsorbed peptides were eluted directly onto a 384 Scout MPT stainless steel 600 μ m AnchorChip MALDI target (Bruker Daltonik, GmbH, Bremen, Germany) with 0.8 μ L of 70% acetonitrile, 0.1% trifluoroacetic acid containing saturated alpha-cyano-4-hydroxycinnamic acid. MALDI-TOF mass spectra were acquired using an UltraflexTM MALDI-Tof/ToF instrument (Bruker Daltonik, GmbH, Bremen, Germany). For mass fingerprint analysis, each raw spectrum was opened in Flex Analysis (version 2.2; Bruker Daltonik, GmbH) and treated using ScoreBoosterTM calibration in ProteinScape (version 1.3; Bruker Daltonik, GmbH), that include calibration with two peptides resulting from trypsin autolysis (m/z 842.5100 and 2211.1046). The filtered peak list for monoisotopic masses was used to query NCBI nr sequence databases (version 20060609, 3682060 sequences), using the Mascot search algorithm (Mascot server version 2.1.04; <http://www.matrixscience.com>) using the following parameters: trypsin as enzyme, two possible missed cleavages, oxidized methionine as variable modification, MH⁺ monoisotopic masses, and peptide tolerance of 70 ppm. Results were scored using the probability-based Mowse score (The protein score is $-10 \times \log(p)$ where p is the probability that the observed match is a random event. Protein scores greater than 66 are significant ($p < 0.05$)).

Table 1: Oligonucleotides used in this study

Oligonucléotides	taille	Séquences (5' vers 3')
L87	87-mer	CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCC GGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCA
L57	57-mer	TCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATG CAAGCTTGGCA
M13 CutFAM	32-mer	TGCCAAGCTTGCATGCCTGCAGGTCGACTCTA
M13 prim	30-mer	ATTCGTAATCATGGTCATAGCTGTTTCCTG
Nick-55	55-mer	GAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATAG CTGTTTCCTG
THFnick-cy5	55-mer	THF5'GAG-GAT-CCC-CGG-GTA-CCG-AGC-TCG-AAT-TCG-TAA-TCA- TGG-TCA-TAG-CTG-TTT-CCT-G 3'CY5
5nt-cy5	50-mer	5'TCC-CCG-GGT-ACC-GAG-CTC-GAA-TTC-GTA-ATC-ATG-GTC- ATA-GCT-GTT-TCC-TG 3'CY5
THF5nt-cy5	50-mer	THF5'TCC-CCG-GGT-ACC-GAG-CTC-GAA-TTC-GTA-ATC-ATG-GTC- ATA-GCT-GTT-TCC-TG 3'CY5

Results and Discussion

***PabFen1* stimulate the primer elongation activity of *Pabp41* and *Pabp41/ Pabp46*.**

All the genomes of Archaea sequenced to date contain one representative of the archaeo-eukaryotic primases superfamily. This family encompasses diverse members from Archaea, Eukarya, viruses and Bacteria, the latter representatives being generally fused to a DNA ligase domain (Iyer *et al.*, 2005; Weller *et al.*, 2002). Recent studies have demonstrated that archaeal primases act on a wider range of substrates, compared to eukaryotic counterparts, and display an architectural connection with the active site of pol X family members (Augustin *et al.*, 2001; Lao-Sirieix *et al.*, 2005b), raising the paradigm of archaeal primases potentially involved in DNA repair processes. Consistent with this hypothesis, we previously reported that the DNA primase from *P. abyssi* displays gap-filling and strand displacement activities, generating 5' single stranded flaps (Le Breton *et al.*, 2007). Because of *PabFen1*'s established endonuclease activity on these products (Kaiser *et al.*, 1999), we tested the functional association, *in vitro*, of *PabFen1* with *Pabp41* and *Pabp41/p46* complex.

We initially determined the minimum amount of *Pabp41* and *Pabp41/p46* complex to use with gapped and nicked substrates in order to observe a gap filling activity. For both

entities the enzymatic assays were set up with 5 pmoles of proteins and performed at 60°C, below the melting temperature (67°C) of the annealed substrates (data not shown). This temperature reaction is below the optimum temperature for these hyperthermostable enzymes and could explain the low rate of activity observed. Then we examined the effect of *PabFen1* on the elongation primer activity of the catalytic subunit *Pabp41* alone or in complex with its regulatory subunit *Pabp46*. For this purpose, the circular DNA M13mp18 was used as substrate, hybridized with a 32 mer labelled with fluoresceine at the 5' extremity and with various downstream primers mimicking nicked, 5-nt and 25-nt gapped DNA substrates (Fig. 1). As previously described, *Pabp41* and *Pabp41/p46* complex exhibited primer elongation activity, although *Pabp41/p46* complex display a reduced ability to incorporate dNTPs and generate shorter products (Fig.1A, lanes 2 and 4). Surprisingly, in the presence of *PabFen1*, the elongation products of *Pabp41* were longer than with *Pabp41* alone, indicating that *PabFen1* can stimulate the polymerisation activity of this catalytic subunit (Fig. 1A, compare lanes 2 and 3). In contrast, the ability of the *Pabp41/Pabp46* complex to elongate DNA primer was not enhanced by *PabFen1* and the DNA fragments generated were even smaller when *PabFen1* was present (Fig. 1A, compare lanes 4 and 5). Then, we tested the influence of *PabFen1* on nicked, 5-nt and 25-nt gapped circular substrates. From the data presented in Fig. 1B, we estimated that *Pabp41* alone or in complex with *Pabp46* can displace one or two bases on nicked M13 substrate (Fig. 1B, lanes2 and 4). When *PabFen1* was present with *Pabp41* or *Pabp41/p46*, we observed a significant increase in the length of the elongated products, up to 20 bases (Fig. 1B, lanes 3 and 5). The Figure 1C presents the data obtained with 5nt gapped substrates. *Pabp41* alone filled in the gap and displayed a weak strand opening activity. When *PabFen1* was added, longer DNA fragments were generated, up to 87 bases that corresponded to a product elongated up to the 3' end of the downstream primer. In contrast, we could not observe a clear effect of *PabFen1* on the *Pabp41/p46* complex on this substrate (Fig. 1C,

lanes 4 and 5). Also, the heterodimeric enzyme seemed to display an enhanced primer elongation activity on this substrate, when compared with *Pabp41* alone (Fig.1C, compare lanes 2 and 4). This different behaviour was also observed, albeit to a lesser extent, on 25-nt gapped substrates (Fig. 1D, compare lanes 2 and 4). Though, in that case, a net increase of the length of the fragments synthesized by *Pabp41* and by the complex *Pabp41/p46* can be observed in the presence of *PabFen1* (Fig. 1D, lanes 3 and 5). It is important to note that these reactions were also performed with BSA instead of *PabFen1*, no stimulation of *Pabp41* or *Pabp41/Pabp46* was observed (data not shown), indicating that the influence of *PabFen1* is specific.

In brief, *Pabp41* as well as the complex *Pabp41/p46* exhibit a weak strand opening activity on nicked substrates and the complex *Pabp41/p46* appears to be more active than *Pabp41* alone on gapped substrates (Fig.1C, lane 4 ; Fig.1D, lane 4). With the exception of the complex *Pabp41/p46* onto primed and 5-nt gapped substrates, upon association with *PabFen1*, the size of the DNA fragments synthesized by *Pabp41* and the holoenzyme is increased and exceeds the size of the gap. We estimated the length of the major product synthesized in the presence of *PabFen1* to be of about 20 nucleotides beyond the nick or the gap, independently of the substrate considered. Altogether, these results indicate that *PabFen1* has an effect on both *Pabp41* and the complex *Pabp41/p46*, potentially stimulating their strand displacement activity. This observation raises two fundamental questions, does this functional association reflect a physical association between these components and what is the molecular basis of this stimulation ?

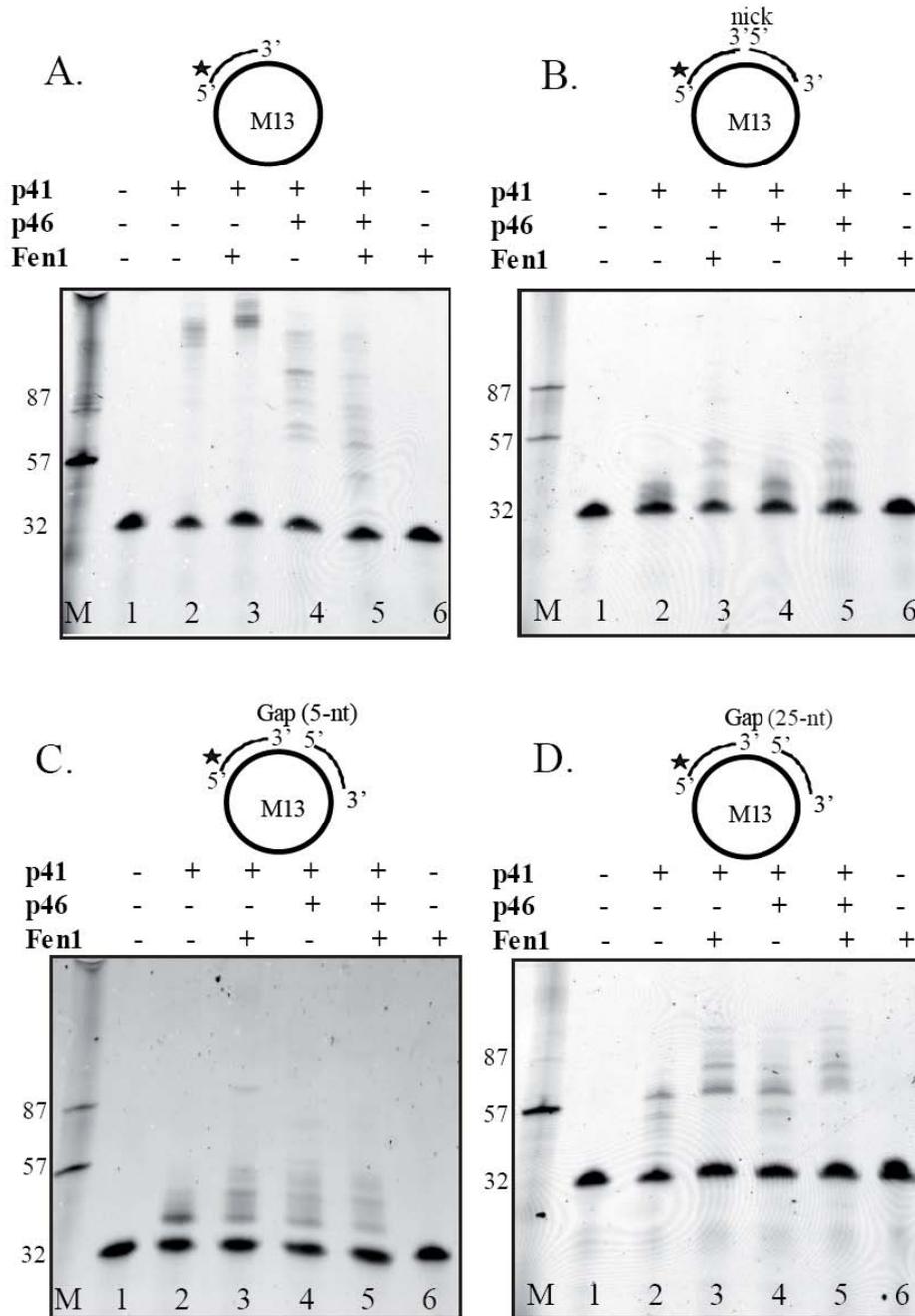


Figure 1 : *PabFen1* stimulate the primer elongation activity of *Pabp41* or *Pabp41/p46* complex. The DNA templates used were the circular M13mp18 DNA with fluoresceine-5' labeled primers assembled to create primed (A), nicked (B), 5-nt (C) and 25-nt (D) gapped templates. The assays were performed at 60°C for 20 minutes, in the presence of all four dNTPs and with 5 pmol of *Pabp41*, *Pabp46* and *PabFen1* or without any enzymes as control. The products were separated on a denaturing 15% polyacrylamide gel and visualized using Typhoon 9400 phosphorimager.

The catalytic subunit of the DNA primase physically associates with the endonuclease *PabFen1* *in vitro* and in the cell-free extract.

To address the former issue, we have looked for interacting partners of *PabFen1*, *Pabp41* and *Pabp41/Pabp46* in cell-free extracts of *P. abyssi*, using a Pull-down/MS approach. We have used the recombinant forms of *PabFen1*, *Pabp41* and *Pabp46* from *P. abyssi* as baits immobilized on magnetic beads. These baits were further incubated with the total protein extract of *P. abyssi* and the interacting proteins that co-sedimented with the baits were separated and identified by MALDI-TOF analyses. Compared with a negative control, that consists of the magnetic beads alone, *PabFen1* specifically co-precipitated with the *PabPCNA* (44 % of sequence coverage, Fig. 2A). This interaction has already been extensively described in Eukaryotes and Archaea (Chapados *et al.*, 2004; Matsumoto, 2001; Sakurai *et al.*, 2005), therefore this finding can be viewed as a proof of principle of this methodological approach. In addition, a mixture of *PabFen1* and *Pabp41* was also identified in a band migrating at about 75 kDa during two independent pull-down/MS assays (47 and 55 % sequence coverage, respectively. Fig. 2A). This value approximately corresponds to the sum of the two components (41kDa + 39kDa), suggesting that the complex co-migrate in the same band. In support to this hypothesis, both the precast gel and the loading buffer are SDS free, although they are formulated for denaturing gel electrophoresis applications. These conditions could nevertheless impair a full denaturation of the protein extract and lead to the co-migration of a pre-assembled complex. It should also be noted that the heat denaturation step in a loading buffer devoided of SDS might not be efficient enough to promote a full denaturation of the hyperthermostable proteins extract.

The reverse experiment with *Pabp41* or *Pabp41/Pabp46* complex as the cellular baits were also performed, but in that case, *PabFen1* could not be detected as a potential interacting partner (data not shown). It is possible that the surface area involved in the physical

association between *Pab41* and *PabFen1* might not be accessible when *Pabp41* was bound to the beads via the histidine tag. To discard the possibility that *Pabp41* was co-precipitated with *PabPCNA*, we performed a Surface Plasmon Resonance (SPR) analysis with a *PabPCNA* chip and confirmed that *Pabp41* did not interact with *PabPCNA* (data not shown). This result was consistent with the fact that *Pabp41* does not display the canonical motif of PCNA interacting proteins (PIP motif, (Warbrick *et al.*, 1998)). Altogether, these data suggest that a *Pabp41*/*Fen1* strong complex or alternatively a higher order complex including *Pabp41*, *PabFen1* and *PabPCNA* might form in the cellular extract of *P. abyssi*.

The regulatory subunit *Pabp46* was not found as a potential partner of *PabFen1*, suggesting that *Pabp46* is not part this new complex. However, we can not rule out the possibility that *Pabp46* might be degraded or present at concentration below the sensitivity threshold of coomassie brilliant blue staining. Yet, we have further characterized the interaction between *Pabp41* and *PabFen1*, using co-immunoprecipitation and SPR experiments. As shown in Fig. 2B, anti-*Pabp41* and anti-*PabFen1* antibodies recognized specifically the purified recombinant proteins, respectively (Fig. 2B, lanes 6 and 7). Moreover, the recombinant *Pabp41* was co-immunoprecipitated with anti-*PabFen1* antibodies *in vitro* (Fig. 2B, lane 5). The controls performed confirmed the specificity of the interaction (Fig. 2B, lanes 1 and 3). These results show that recombinant *Pabp41* and *PabFen1* form a stable complex. This finding was corroborated by SPR experiments, and the *Pabp41* titration experiment over a *PabFen1* coated chip allowed us to determine that the association and dissociation rates were of $2,7 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$ and $3,2 \cdot 10^{-3} \text{ s}^{-1}$, respectively. These values were consistent with the binding affinity value of 10 nM deduced from the simultaneous fit of the association and dissociation phases (although the statistical chi2 value of 15 is slightly out of significative range, Fig. 2C). This value indicated a strong affinity between *Pabp41* and *PabFen1*.

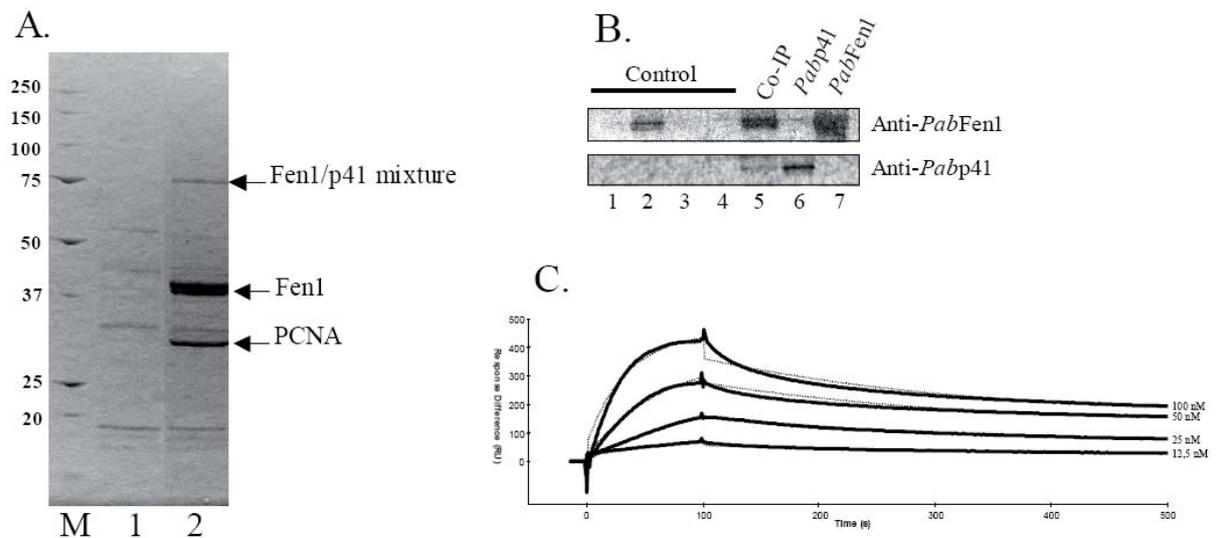


Figure 2 : Physical interaction between *PabFen1* and *Pabp41*. (A) Pull-down of interacting partners of *PabFen1* in the cell extract. Proteins bound to Fen1 were analysed on SDS-gel electrophoresis (lane 2). The specific interaction with Fen1 can be compared with a negative control that consists of the magnetic beads alone (lane 1). (B) Co-immunoprecipitation assay in vitro. (C) SPR assay. Fen1 was immobilized on a CM5 sensor chip (Biacore) and 30 μ l of various concentrations of p41, ranging from 12,5 nM to 100 nM, were injected over the chip, as described in Materials and methods.

***PabFen1* and *Pabp41* cooperate via nick translation or strand displacement and subsequent flap endonuclease activities on nick and gap substrates.**

We have further focussed our work on the functional basis of the interaction between *Pabp41* and *PabFen1*. In order to determine the influence of *PabFen1* on *Pabp41* activity, we have carried out experiments using increasing concentrations of *PabFen1* with a constant amount of *Pabp41*. These assays were performed with nicked and 5-nt gapped templates. For all the substrates tested, *Pabp41* had the ability to fill in the gapped substrates but no stimulation was observed when *PabFen1* was added at lower concentrations than *Pabp41* (figure 3A, lanes 1-2 and 5-6). In contrast, when *PabFen1* and *Pabp41* were present at equimolecular or suprastochiometric ratio, longer DNA fragments were generated (Fig. 3A, lanes 3-4 and 7-8). For both substrates, the upstream primers were elongated of around 20 bp beyond the nick or the gap, resulting in maximum length product of about 52 and 57 bases,

respectively. The fate of the downstream primer, labelled with a distinct fluorophore, was simultaneously studied. As shown in fig.3B, *PabFen1* displayed a nucleolytic activity towards the downstream primers, when present at equimolecular or suprachrometric ratio, for both substrates (Fig. 3B, lanes 3-4 and 7-8). It should be noted that at these concentrations and in the absence of DNA polymerase, *PabFen1* exhibited no cleavage activity on the downstream primer (data not shown). The cleavage products ranged from 50 and 23 nucleotide in size and were distributed according to a processive pattern. The main cleavage products ranged between 50 and 55 nucleotide, indicating that *PabFen1* preferentially generate fragments of 5 to 10 nucleotides. The discrepancy between the size of the elongated upstream primers and the size of the cleavage products of the downstream primers could be explained by the difference in sensitivity of the distinct fluorophore used.

Taken together, these data may be explained according to two distinct hypotheses. One proposed model could be that upon association with *PabFen1*, *Pabp41* gains the ability to displace the downstream primer into a flap, *PabFen1* can track from the 5' end to the elbow of the flap for cleavage and generate products between 5 and 10 nucleotides in length. The preferential substrate for archael Fen1 has been demonstrated to be similar to that from eukaryotic counterparts and consists of a "double flap structure" (Chapados *et al.*, 2004; Friedrich-Heineken and Hubscher, 2004; Hosfield *et al.*, 1998). The physical association of *PabFen1* with *Pabp41* at equimolecular ratio could stimulate the primer elongation and strand displacement activity of this latter DNA polymerase, thereby promoting the formation of the double flap structure. Subsequently, this DNA structure would stimulate the endonucleolytic activity of *PabFen1* that potentially generate a ligatable nick. On the other hand, it is conceivable that the cleavage products of *PabFen1* might be generated by a 5'->3' exonuclease activity on fully annealed double strand DNA. This activity has been shown for mammalian Fen1 (Harrington and Lieber, 1994; Murante *et al.*, 1995) but up to now this has

not been demonstrated for the Archaeal homologues. Yet, we can not rule out the possibility that *PabFen1* might exhibit such an activity that could be stimulated by *Pabp41*. In that case, the potential ligatable nick could be maintained by nick translation, with the coordinated action of the 5'→3' exonuclease activity of *PabFen1* and the gap filling activity of *Pabp41*. Although our results can not help us to discriminate between these two hypothesis, the length of the cleaved products was not consistent with the nick translation hypothesis. In addition, the preference of *PabFen1* for flap or double flap structures also argue for the strand displacement and subsequent flap endonuclease hypothesis.

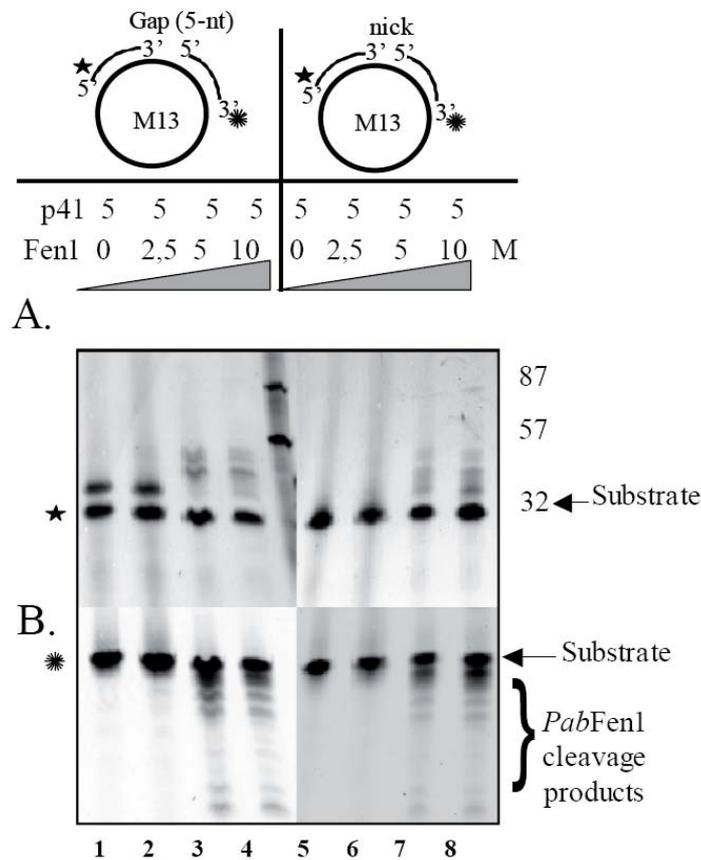


Figure 3 : Pabp41 and PabFen1 cooperate for the degradation and replacement of primers downstream nicked or gapped substrates. The DNA templates used were the circular M13mp18 DNA with fluoresceine-5' labeled primer assembled to create nicked and 5-nt gapped templates. Primer extension assays were performed at 60°C for 20 minutes with 5 pmol of *Pabp41* without or with various amount of *PabFen1* (2.5; 5 and 10 pmol). The product were separated on a denaturing 15% polyacrylamide gel and visualized using Typhoon 9400 phosphofluorimager. (A) Visualization of DNA primer fluoresceine-5' end labeled and DNA synthesis by *Pabp41* with a marker between Lane 4 and 5. (B) Visualization of CY5-3'-end labeled downstream DNA and *PabFen1* cleavage products.

Reconstitution *in vitro* of Base Excision Repair using purified recombinant *P. abyssi* proteins.

The experiments carried out up to this point did not allow us to determine the biological rationale of the functional association between *PabFen1* and *Pabp41*. However, as a structural connection between the active sites of Eukariotic Pol β and archaeal DNA primase catalytic subunit (Augustin *et al.*, 2001; Lao-Sirieix *et al.*, 2005b) has been demonstrated and since Pol β has been shown to collaborate with Fen1 in Mammalian Long Patch Base Excision Repair (LP-BER), we have focused on the study of the resolution of LP-BER intermediates by this new complex. For this purpose, we have set up a functional assay using THF substrates, mimicking the initial step of LP-BER and resistant to β -elimination, that can only be “repaired” by displacement or degradation of the downstream primer. We have used a nicked substrate with a THF residue at the 5' moiety of the downstream primer, which has been shown to stimulate a naturally reduced abasic site. Such a site would be a typical intermediate of LP-BER *in vivo*. Fig. 4 compares the synthesis activity of *Pabp41* on nick and nick-THF substrates. The pattern of elongated DNA fragments is similar for both substrates in presence or absence of *PabFen1*, indicating that this LP-BER intermediate is not a barrier for strand displacement activity of *Pabp41* or for putative 5'->3' exonucleolytic activity of *PabFen1*, arguing for the potential implication of this complex in BER mechanisms. To further demonstrate the capability of the complex *Pabp41/Fen1* to resolve BER intermediates, we have produced and purified the DNA ligase I of *P. abyssi* in order to assess whether the coordinated activities of *Pabp41* and *PabFen1* can maintain a ligatable nick. Whereas DNA ligase I was able to seal the upstream primer with the 5'-phosphate downstream primer (data not shown), THF residue at the 5' terminal position prevent ligation between these two primers (Fig. 5, lane 3). This result is consistent with the fact that ligation

of a THF-terminal downstream primer during repair would be a harmful process (Ranalli *et al.*, 2002). DNA ligase I had no effect on the activity of *Pabp41* alone (Fig. 5, lane 6) but when *Pabp41*, *PabFen1* and DNA ligase were present in the same mixture, a product of 87 pb that carry both fluorescent dyes was generated and corresponded to the sealing of upstream and downstream primers (Fig. 5, lane 7). Altogether, these data indicate that the coordinated activities of *Pabp41* and *PabFen1* generate the formation of a nick that can be sealed by the DNA ligase I. This result argues for the involvement of this complex at DNA structures where a non ligatable 5' end has to be degraded before restoring by ligation the integrity of double stranded DNA. These structures occur both during Okazaki fragment maturation and in LP-BER mechanisms where a RNA initiator or a 5' moiety with modified sugar group, respectively, have to be removed and replaced.

Base Excision Repair has only been demonstrated in the hyperthermophilic archaeon *Pyrobaculum aerophilum* (Sartori and Jiricny, 2003). This Crenarchaeal representative encompasses the minimum set of proteins required to repair a modified DNA base. Sartori and Jiricny reported the biochemical evidence that a combination of four enzymes consisting of Uracil DNA Glycosylase, AP endonuclease IV, DNA polymerase B2 (Pol B2) and DNA ligase can efficiently repair a G:U mispair, emphasizing the importance of BER process for hyperthermophile archaea. Whereas homologs of Uracil DNA glycosylases, AP endonuclease IV and DNA ligase are present in all the Euryarchaea sequenced to date, the DNA polymerases content differs between Euryarchaea and Crenarchaea. In particular, Pol B2, has no homologues in Euryarchaea (Böhlke *et al.*, 2002). This shorter representative of family B DNA polymerase, devoided of 3'->5' exonuclease, has been shown in *P. aerophilum* to be a functional homologue of mammalian pol β in BER process (Sartori and Jiricny, 2003).

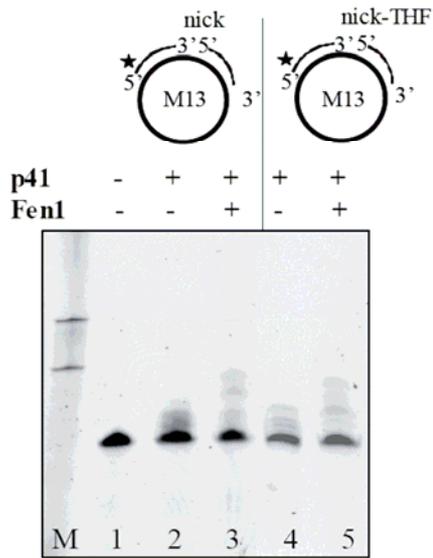


Figure 4: Nick-THF substrate is not a barrier for the complex *Pabp41/Fen1*. The Nick substrate and the Nick-THF substrate was incubated for 20 min at 60°C in the presence or absence of *Pabp41* and *PabFen1* (5 pmol of each protein). The product were separated on a denaturing 15% polyacrylamide gel and visualized using Typhoon 9400 phosphorimager.

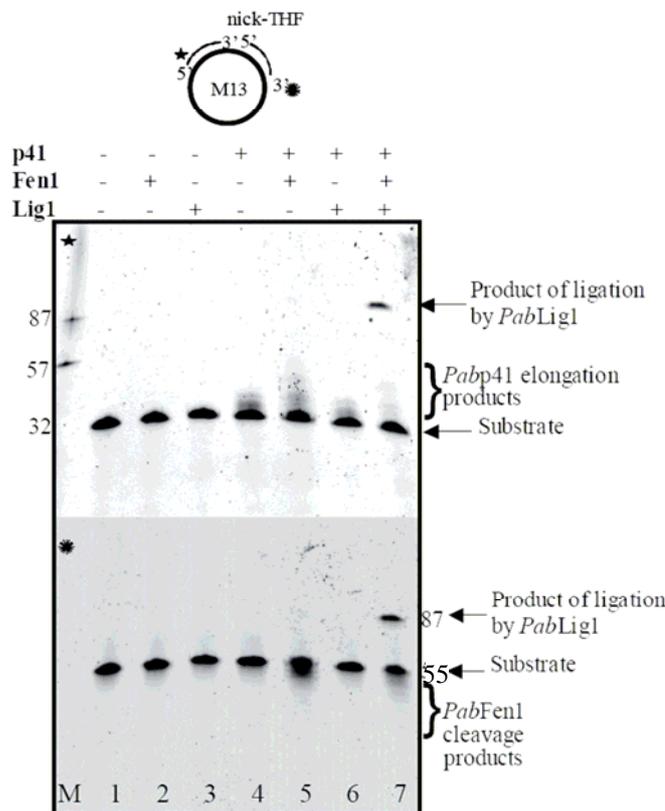


Figure 5 : *in vitro* resolution of LP-BER intermediate using recombinant *Pyrococcus abyssi* proteins. The Nick-THF substrate was incubated for 25 min à 60°C in the presence or absence of *PabFen1*, *PabLig1* and *Pabp41* (lanes 2-7) (5 pmol of each protein).. A, visualization of DNA primer fluoresceine-5' end labeled and DNA synthesis by *Pabp41*. B, visualization of CY5-3'-end labeled downstream DNA and *PabFen1* cleavage products. The position of the substrates and products are indicated. The product of ligation by *PabLig1* can be visualized on both pictures.

The present data provide additional evidence for the involvement of archaeal DNA primase in another role that primer synthesis. Given the lack of genuine Pol X proteins in *P. abyssi* and based on our results, we propose that *Pabp41* alone or associated with *Pabp46*

might play a role, similar to that of Pol B2, in BER processes. In addition, we suggest that the coordinated activities of *PabFen1* and of the Archaeal DNA primase can resolve nicked or gapped DNA structures and generate contiguous DNA extremities that can be sealed by DNA ligase I. While not disregarding the putative implication in Okazaki fragment resolution, we suggest that *Pabp41* or the complex *Pabp41/Pabp46* can physically associate with *PabFen1* and resolve DNA intermediates of BER pathway.

It will be interesting, in the future directions of this work, to determine whether the Archaeal DNA primase, as Pol B2 and Pol β , also display d-Rp lyase activity. In addition, the coordination with other factor and enzymes of the BER (DNA glycosylase, AP endonuclease, PCNA) will be studied. These data will benefit to the understanding of BER process and more generally to the global knowledge of genomic stability in Archaea.

Acknowledgment: This work was supported by Ifremer, Marine Genomics Europe (contract number: PNGOCE-CT-2004-505403) and French program “GIS Institut de la Génomique Marine” (contract Number : FNS-500-370). We also thank Pascal Trouvé (INSERM U613, Molecular Genetics and Genetic Epidemiology, Brest, France) for access to surface plasmon resonance facilities.

References

- 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.
- Bocquier, A. A., Liu, L., Cann, I. K., Komori, K., Kohda, D., and Ishino, Y. (2001). Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr Biol* 11, 452-456.
- Böhlke, K., Pisani, F. M., Rossi, M., and Antranikian, G. (2002). Archaeal DNA replication: spotlight on a rapidly moving field. *Extremophiles* 6, 1-14.
- 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. O., and Ishino, Y. (1999). Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics* 152, 1249-1267.
- Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J. A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* 116, 39-50.
- Com, E., Evrard, B., Roepstorff, P., Aubry, F., and Pineau, C. (2003). New insights into the rat spermatogonial proteome: identification of 156 additional proteins. *Mol Cell Proteomics* 2, 248-261.
- De Falco, M., Fusco, A., De Felice, M., Rossi, M., and Pisani, F. M. (2004). The DNA primase of *Sulfolobus solfataricus* is activated by substrates containing a thymine-rich bubble and has a 3'-terminal nucleotidyl-transferase activity. *Nucleic Acids Res* 32, 5223-5230.
- Friedrich-Heineken, E., and Hubscher, U. (2004). The Fen1 extrahelical 3'-flap pocket is conserved from archaea to human and regulates DNA substrate specificity. *Nucleic Acids Res* 32, 2520-2528.
- Godfroy, A., Postec, A., and Raven, N. (2006). Growth of Hyperthermophilic Microorganisms for Physiological and Nutritional Studies. In *Methods in Microbiology, Extremophiles*, F. A. Rainey, and A. Oren, eds. (Oxford, England, Elsevier/Academic Press).
- Grogan, D. W. (2004). Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol* 6, 137-144.
- 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*. *European Journal of Biochemistry* 268, 5961-5969.
- Harrington, J. J., and Lieber, M. R. (1994). The characterization of a mammalian DNA structure-specific endonuclease. *Embo J* 13, 1235-1246.
- Henneke, G., Flament, D., Hubscher, 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*.
- Hosfield, D. J., Mol, C. D., Shen, B., and Tainer, J. A. (1998). Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell* 95, 135-146.
- Iwai, T., Kurosawa, N., Itoh, Y. H., Kimura, N., and Horiuchi, T. (2000). Sequence analysis of three family B DNA polymerases from the thermoacidophilic crenarchaeon *Sulfurisphaera ohwakuensis*. *DNA Res* 7, 243-251.
- Iyer, L. M., Koonin, E. V., Leipe, D. D., and Aravind, L. (2005). Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res* 33, 3875-3896.
- Kaiser, M. W., Lyamicheva, N., Ma, W., Miller, C., Neri, B., Fors, L., and Lyamichev, V. I. (1999). A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. *J Biol Chem* 274, 21387-21394.
- Kelman, Z., and White, M. F. (2005). Archaeal DNA replication and repair. *Curr Opin Microbiol*.
- Kirk, B. W., and Kuchta, R. D. (1999). Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases

- share significant sequence homology. *Biochemistry* 38, 7727-7736.
- 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 archaeon *Sulfolobus solfataricus*. *Mutat Res* 357, 245-253.
- Lao-Sirieix, S. H., and Bell, S. D. (2004). The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol* 344, 1251-1263.
- Lao-Sirieix, S. H., Pellegrini, L., and Bell, S. D. (2005). The promiscuous primase. *Trends Genet.*
- 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.*
- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* 362, 709-715.
- Liu, L., Komori, K., Ishino, S., Bocquier, A. A., Cann, I. K., Kohda, D., and Ishino, Y. (2001). The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J Biol Chem* 276, 45484-45490.
- Matsumoto, Y. (2001). Molecular mechanism of PCNA-dependent base excision repair. *Prog Nucleic Acid Res Mol Biol* 68, 129-138.
- Meslet-Cladiere, L., Norais, C., Kuhn, J., Briffotiaux, J., Sloostra, J. W., Ferrari, E., Hubscher, U., Flament, D., and Myllykallio, H. (2007). A Novel Proteomic Approach Identifies New Interaction Partners for Proliferating Cell Nuclear Antigen. *J Mol Biol.*
- Murante, R. S., Rust, L., and Bambara, R. A. (1995). Calf 5' to 3' exo/endonuclease must slide from a 5' end of the substrate to perform structure-specific cleavage. *J Biol Chem* 270, 30377-30383.
- Ogrunc, M., Becker, D. F., Ragsdale, S. W., and Sancar, A. (1998). Nucleotide excision repair in the third kingdom. *J Bacteriol* 180, 5796-5798.
- Ramadan, K., Shevelev, I., and Hubscher, U. (2004). The DNA-polymerase-X family: controllers of DNA quality? *Nat Rev Mol Cell Biol* 5, 1038-1043.
- Ranalli, T. A., Tom, S., and Bambara, R. A. (2002). AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair. *J Biol Chem* 277, 41715-41724.
- Sakurai, S., Kitano, K., Yamaguchi, H., Hamada, K., Okada, K., Fukuda, K., Uchida, M., Ohtsuka, E., Morioka, H., and Hakoshima, T. (2005). Structural basis for recruitment of human flap endonuclease 1 to PCNA. *Embo J* 24, 683-693.
- Sartori, A. A., and Jiricny, J. (2003). Enzymology of base excision repair in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J Biol Chem* 278, 24563-24576.
- Uemori, T., Ishino, Y., Doi, H., and Kato, I. (1995). The hyperthermophilic archaeon *Pyrodictium occultum* has two alpha-like DNA polymerases. *J Bacteriol* 177, 2164-2177.
- Warbrick, E., Heatherington, W., Lane, D. P., and Glover, D. M. (1998). PCNA binding proteins in *Drosophila melanogaster*: the analysis of a conserved PCNA binding domain. *Nucleic Acids Res* 26, 3925-3932.
- Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., d'Adda di Fagagna, F., et al. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686-1689.

2.2. Résultat complémentaire de l'étude : recherche de partenaires de p41

Afin d'avoir une vision complémentaire sur le rôle de l'ADN primase de *P. abyssi* et de confirmer l'interaction entre Fen1 et p41, nous avons recherché les partenaires de p41 dans le lysat cellulaire (figure 43). Les partenaires identifiés étaient les suivants : la reverse gyrase (*Pab2423*), RPA41 (*Pab2163*), RPA32 (*Pab2165*), les sous-unités A et B de l'ARN polymérase (*Pab0425* et *Pab0423*), la protéine *Pab1478* (voir le tableau 10). Fen1 n'a pas été identifié dans ce contexte.

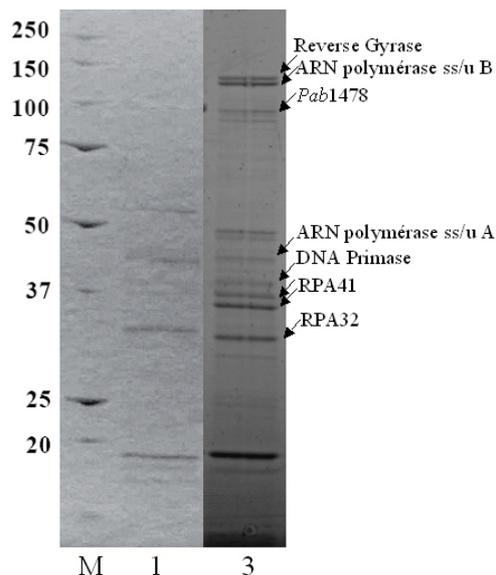


Figure 43 : Identification des partenaires potentiels de p41. La piste 1 correspond au contrôle négatif, la piste 2 correspond à l'expérience de pull-down avec p41 comme appât. Les flèches noires correspondent aux partenaires potentiels.

Appâts	Partenaires interactants					
	Protéines	Accession number	Masse Moléculaire	Score Total	Peptides matched	% Coverage
Flap	PCNA	gi 5458797	28 017	80	9/33	44
Endonuclease 1 gi 5458165	DNA primase catalytic subunit (p41)	gi 5457615	40 653	162	19/69	47
<i>Pab1877</i> (39 kDa)	Topoisomerase type VI, subunit B	gi 5458026	63 871	182	20/37	40
DNA primase p41 catalytic subunit gi 5457615	Reverse Gyrase	gi 14521467	139 900	317		43
<i>Pab2236</i> (41 kDa)	DNA-directed RNA polymerase subunit beta	gi 14520830	127 000	124		29
	<i>Pab1478</i>	gi 14521559	94 200	90,5		37,5
	Acetyltransferase (GNAT) family	gi 14521475	94 600	163		45,4
	Glucose-1-phosphate thymidyltransferase	gi 14521899	46 400	135		42
	PremRNA splicing protein Nop58plike	gi 14520278	47 200	81,6		37,6
	DNA-directed RNA polymerase subunit A	gi 14520832	44 600	105		49,9
	<i>Pab2163</i> RPA41	gi 14520503	43 000	101		48,8
	<i>Pab2165</i> RPA32	gi 14520501	31 800	153		63,2

Tableau 10 : partenaires identifiés par spectrométrie de masse, des protéines Fen1 et p41.

2.2.1. Les protéines RPA, partenaires de l'ADN primase

Les protéines RPA sont l'équivalent chez les archées et les eucaryotes des protéines SSB bactériennes. Ces protéines se lient et protègent l'ADN simple brin, elles empêchent la réassociation des deux brins d'ADN au niveau de la fourche de réplication. Le RPA est connu pour interagir avec de nombreux partenaires lors des processus de réplication et de réparation de l'ADN (Wold, 1997). Chez *P. abyssi*, le complexe RPA est un hétérotrimère composé de trois sous-unités : RPA41 (43 kDa), RPA32 (32 kDa) et RPA14 (14 kDa). Chez les eucaryotes, il a été démontré que RPA interagit fonctionnellement avec la polymérase- α primase. RPA stimule l'activité polymérase et augmente la processivité du complexe Pol α /Primase (Braun *et al.*, 1997). Il a également été proposé que RPA stabilise le complexe Pol α /Primase en se conduisant comme un *clamp*, et améliore la fidélité de la polymérase α (Maga *et al.*, 2001). Par ailleurs, une interaction directe entre la sous-unité p70 du RPA eucaryote et la sous-unité primase p48 du complexe Pol α /Primase a été démontrée (Nasheuer *et al.*, 1992). Une autre étude démontre l'interaction, par co-immunoprécipitation, entre le RPA41 et la sous-unité p41 de la primase chez *Pyrococcus furiosus* (Bocquier *et al.*, 2001; Komori and Ishino, 2001). Ces observations suggèrent une interaction similaire entre le RPA et le complexe Pol α /Primase eucaryote, et entre le RPA et la primase des archées, dans le contexte de l'initiation de la réplication et de la synthèse des fragments d'Okazaki. Mes résultats confirment cette hypothèse.

D'autre part, une étude a démontrée que les protéines SSB de *S. solfataricus* (équivalentes aux RPA des autres archées) peuvent déshybrider de façon spécifique de l'ADN contenant un mésappariement ou une lésion, *in vitro* (Cubeddu and White, 2005). Les auteurs suggèrent que les protéines SSB pourraient détecter l'ADN endommagé *in vivo* et recruter d'autres protéines spécifiques de la réparation. Les propriétés particulières des ADN primases d'archées, leurs capacités à se comporter comme des polymérases, suggèrent qu'elles pourraient participer à des mécanismes de réparation de l'ADN. Le RPA des archées pourrait, dans ce contexte, recruter les ADN primases pour combler une brèche de l'ADN.

2.2.2. La reverse gyrase est-elle est un nouveau partenaire de l'ADN primase ?

La reverse gyrase est une ADN topoisomérase de type I qui introduit des surenroulements positifs de l'ADN. C'est une protéine retrouvée uniquement chez les hyperthermophiles (archées et bactéries) (Forterre, 2002b). Son rôle dans l'adaptation aux hautes températures reste encore à préciser. Cependant, il est probable que le super-enroulement positif stabilise l'ADN aux températures élevées. L'étude de la séquence primaire de la reverse gyrase montre une organisation en deux domaines : un domaine C-terminal présentant une activité topoisomérase, et un domaine N-terminal présentant une activité hélicase. Cette organisation en deux domaines a été confirmée par l'étude de la structure de la reverse gyrase chez *Archaeoglobus fulgidus* (Rodriguez and Stock, 2002).

Dans le cas où l'interaction entre p41 et la reverse gyrase serait confirmée par une autre méthode, il s'agit d'un complexe non décrit à ce jour.

Une interaction fonctionnelle entre la reverse gyrase et les protéines SSB a été caractérisée chez *Sulfolobus* (Napoli *et al.*, 2005). Les auteurs montrent que les SSB stimulent l'activité de la reverse gyrase (stimulation de la liaison à l'ADN, du clivage, de la rotation des brins et de la ligature). Cette étude présente également des expériences de co-précipitation qui montrent une interaction entre la reverse gyrase et les SSB dépendantes de la présence d'ADN. Dans une autre étude, les auteurs montrent que la reverse gyrase est recrutée au niveau de l'ADN endommagé aux UV *in vitro* et *in vivo* (Napoli *et al.*, 2004). Ils suggèrent que la reverse gyrase peut participer de manière directe ou indirecte à la réponse cellulaire aux lésions de l'ADN induite par les UV. La stabilisation de régions simple brin sur l'ADN par les protéines SSB permettrait d'accroître l'efficacité de la liaison de la reverse gyrase sur l'ADN, et ainsi stimuler son activité.

Notre étude ne permet pas de déterminer si l'association entre la sous unité catalytique p41 de la primase avec la reverse gyrase est une interaction directe, ou si cette interaction est médiée par l'intermédiaire du RPA. La reverse gyrase intervient dans le dynamisme de la structure de l'ADN *in vivo* nécessaire à l'activité d'autres enzymes. Elle pourrait influencer l'activité de l'ADN primase, tant dans son rôle de synthèse d'amorces initiateuses que dans dans celui de colmatage de brèche lors de mécanisme de réparation de l'ADN.

Dans tous les cas, la recherche de partenaires de la reverse gyrase dans un extrait cellulaire total serait une investigation particulièrement intéressante afin d'améliorer les connaissances sur la fonction de cette protéine, et son implication éventuelle dans des mécanismes de réparation de l'ADN.

2.2.3. Les ARN polymérases, associées au RPA ou à l'ADN primase ?

Les ARN polymérases des archées sont des complexes enzymatiques qui au niveau de la séquence de leurs sous-unités et du point de vue fonctionnel ressemblent aux ARN polymérases des eucaryotes. L'homologue archéenne de la plus grande sous-unité de l'ARN polymérase eucaryote est partagé en deux sous-unités A' et A''. L'homologue chez les archées de la seconde plus grande sous unité ARN polymérase eucaryote est la sous unité B (c'est le cas de *P. abyssi*), ou deux sous unités, B' et B'' selon les espèces (Langer *et al.*, 1995).

Une interaction physique et fonctionnelle a été caractérisée entre les protéines SSB et les ARN polymérases chez *Sulfolobus solfataricus* (Richard *et al.*, 2004). Cette étude montre notamment une activation de la transcription, *in vitro*, en présence des protéines SSB de *S. solfataricus*. Les auteurs proposent que les SSB pourraient stimuler et recruter les ARN polymérases lors de la transcription. Dans notre cas, il est difficile de savoir si les ARN polymérases interagissent de façon directe ou indirecte avec la primase (via les RPA par

exemple). Ces résultats suggèrent néanmoins un couplage possible entre la transcription et la réplication ; ou un couplage entre la transcription et la réparation de l'ADN chez les Archées.

2.2.4. La protéine *Pab1478*

En comparant la séquence primaire de *Pab1478* avec les bases de données, il en ressort que cette protéine est une ATPase de la famille AAA+ et de la sous famille CDC48. Cette protéine, dont la fonction est inconnue chez les archées, a été retrouvée au sein de l'exosome chez *Sulfolobus solfataricus* (Evguenieva-Hackenberg *et al.*, 2003; Walter *et al.*, 2006). Chez les eucaryotes, l'exosome est un complexe de protéines intervenant dans le traitement et la dégradation des ARN. En utilisant un anticorps dirigé contre la protéine Rrp41 de *Sulfolobus solfataricus* (une protéine orthologue d'une sous unité de l'exosome eucaryote), les auteurs ont récupérés, par co-immunoprécipitation, trois autres protéines orthologues de l'exosome eucaryote (Rrp4, Rrp42 et Csl4), l'homologue de la primase bactérienne DnaG, une chaperonine Cpn et la protéine *Sso0421* (Cdc48). Les auteurs remarquent que la protéine Cdc48 de la levure présente une activité de chaperonine, cette activité pourrait entraîner une liaison non spécifique de la protéine lors des expériences d'interaction. Il est donc probable que cette interaction ne soit pas spécifique.

L'ADN primase de *P. abyssi* est donc retrouvé dans notre étude avec plusieurs protéines possiblement impliquées dans des mécanismes de détection ou de réparation de l'ADN endommagé, mais également avec des acteurs de l'initiation de la réplication. Ces données renforcent les conclusions énoncés dans le manuscrit présenté au chapitre III, concernant la dualité de l'ADN primase intervenant à la fois dans des mécanismes de réparation et au niveau de la synthèse des primers initiateurs.

* *
*

2.3. Fen1 est associé à la topoisomérase VI

L'interaction entre Fen1 et la topoisomérase VIB a été détectée par *pull-down* avec Fen1 comme appât. Cette interaction n'a jamais été décrite précédemment. Afin de la confirmer *in vitro*, la forme recombinante de la topoisomérase de *Sulfolobus shibatae* (provenant du laboratoire de P. Forterre) a été utilisée dans des expériences de résonance plasmonique de surface. Les sous-unités A et B de la topoisomérase de *P. abyssi* n'étaient en effet pas disponibles dans notre laboratoire. L'ADN topoisomérase VI est un hétérotétramère A2B2, et les sous-unités A et B de *S. Shibatae* partagent 35% d'identité de séquence avec les sous-unités de *P. abyssi*. L'analyse SPR indique que l'holoenzyme ADN topoisomérase VI et

Fen1 forme un complexe stable ($k_d = 3,6 \cdot 10^{-4} \text{ s}^{-1}$) (figure 44), alors que la sous-unité B seul ne semble pas interagir avec Fen1. Cela suggère que Fen1 est associé à l'holoenzyme via la sous-unité A ou via un changement conformationnel de la sous-unité B dans l'hétérotétramère. En effet, la sous-unité B de la TopoVI contient un site de fixation de l'ATP, nécessaire à l'activité catalytique (Buhler *et al.*, 2001). La fixation de l'ATP modifie en effet la structure de la topoisomérase VI et coordonne le clivage de l'ADN (Corbett *et al.*, 2007).

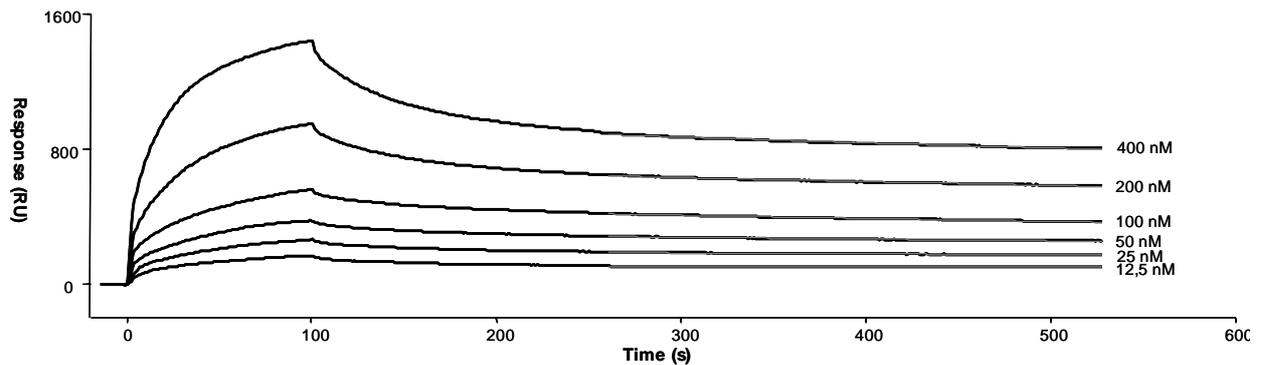


Figure 44 : sensorgramme représentant l'interaction physique entre Fen1 de *P. abyssi* et la topoisomérase VI de *S. shibatae*, grâce à la résonance plasmonique de surface. Différentes concentrations de la topoisomérase VI sont injectées, à temps 0, sur la puce recouverte de une protéine Fen1 fixée.

Les ADN topoisomérases agissent sur la topologie de l'ADN. Elles génèrent des coupures transitoires de l'ADN, modifiant la torsion de l'ADN en enlevant ou ajoutant des super-enroulements. Elles forment une liaison covalente transitoire de type phosphotyrosine avec l'ADN, cela afin d'éviter l'apparition d'extrémités libres de l'ADN au cours de cette réaction. Il existe deux familles d'ADN topoisomérases (I et II) décrites dans les trois domaines du vivant. Les topoisomérases VI, qui appartiennent à la famille II, sont présentes chez toutes les archées (à l'exception des Thermoplasmatales) (Bergerat *et al.*, 1997), elles sont également présentes chez les plantes. Chez les archées, ce sont les seules topoisomérases qui peuvent relâcher les super-enroulements positifs. Une étude chez *S. acidocaldarius* et *H. volcanii* a montré que le radicol inhibe l'activité de la topoisomérase VI *in vitro* et *in vivo*, empêche la croissance des archées étudiées en agissant sur la topoisomérase VI (Gadelle *et al.*, 2005). Cette protéine serait essentielle pour ces archées. Elle supprimerait les super-enroulements de l'ADN engendrés lors du passage de la fourche de réplication de l'ADN ou lors de la transcription, mais participerait également au désassemblage des chromosomes pour faciliter la division cellulaire.

Chez les eucaryotes, la topoisomérase I, qui se lie normalement de manière transitoire à l'ADN peut en présence d'une substance anticancéreuse (la camptothécine) rester accroché de manière covalente sur l'ADN et bloquer la réplication de l'ADN. Une étude a montré que les activités endonucléases de Mre11 et Slx4 étaient nécessaires pour enlever cette intermédiaire topoisomérase-ADN (Deng *et al.*, 2005). Il est possible d'imaginer que Fen1

intervient pour cliver les intermédiaires covalents ADN-topoisomérase, s'ils existent chez les archées, avec la topoisomérase VI. Cette hypothèse est supportée par des études récentes (Parrish *et al.*, 2003; Zheng *et al.*, 2005) démontrant que Fen1 eucaryote présente une activité *Gap endonuclease* (GEN), distincte de l'activité flap endonuclease et qui ne nécessite pas la présence d'une extrémité 5' simple brin. L'enzyme Fen1 des archées pourrait également présenter une telle activité et ainsi permettre de cliver les liaisons ADN-topoisomérase VI entraînant des arrêts de la fourche de réplication.

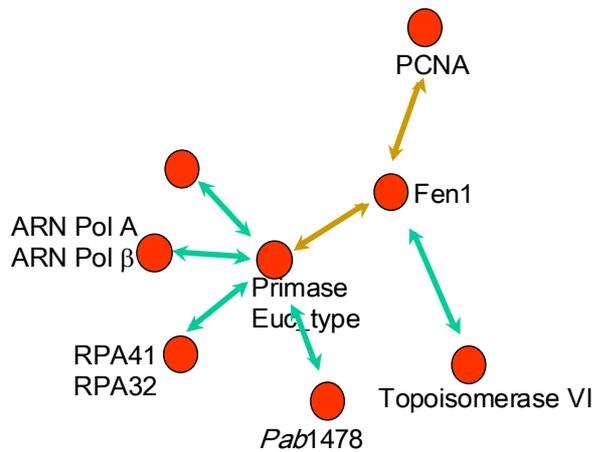


Figure 45 : Réseau d'interaction des partenaires potentiels des protéines Fen1 et p41, mis en évidence par *pull-down* chez *P. abyssi*.

* *
*

III. Analyse du protéome partiel

Une approche complémentaire à la technique de *pull-down* pour l'identification de protéines impliquées dans la maintenance génomique a été envisagée. Partant du postulat que ces protéines sont liées à l'ADN, nous avons entrepris l'analyse de l'intégralité des protéines associées à la chromatine, chez *Pyrococcus abyssi*. L'approche expérimentale consiste à fractionner l'extrait cellulaire, de manière à ne récupérer que l'ADN et les protéines associées, débarrassé des protéines solubles et des débris membranaires. Une électrophorèse bidimensionnelle a été effectuée sur cette fraction, ainsi que sur un lot témoin, ayant subi en plus du traitement précédent, une incubation en présence de DnaseI. La comparaison des profils électrophorétiques de ces deux échantillons nous permet de discriminer entre les faux positifs et les protéines réellement liées à la chromatine (figure 46 et tableau 11).

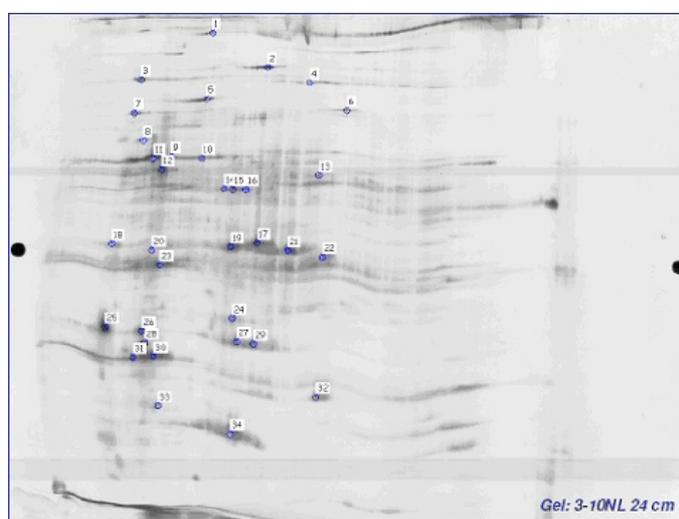


Figure 46 : Electrophorèse bidimensionnelle des protéines de *P. abyssi* liées à la chromatine. Les protéines sont séparées dans une première dimension par isoélectrofocalisation et dans une seconde dimension par SDS-PAGE. Les protéines ont ensuite été colorées par le nitrate d'argent. Les spots indiqués ont été analysés par spectrométrie de masse MALDI-TOF.

N°Spot	Protein ID	Species match	Accession number	Estimated Mw (Da)	Estimated pI	Score	Number peptides matched/ searched	% Coverage
11	flaD/E flagella-related protein D or E	<i>P. abyssi</i> GE5	giI5458909	48 341	4.66	140	17/62	35
14	Primase dnaG-like	<i>P. abyssi</i> GE5	giI14591462	50 281	5.90	72	9/44	31
16	Primase dnaG-like	<i>P. abyssi</i> GE5	giI14591462	50 281	5.90	130	13/40	41
23	PCNA	<i>P. abyssi</i> GE5	giI5458797	28 018	4.55	134	11/30	51
25	rpl12A LSU ribosomal protein L12A	<i>P. abyssi</i> GE5	giI5459201	11 293	4.28	99	7/36	73
27	psmA proteasome, subunit alpha	<i>P. abyssi</i> GE5	giI5458039	29 090	5.47	94	9/36	36
29	psmA proteasome, subunit alpha	<i>P. abyssi</i> GE5	giI5458039	29 090	5.47	144	13/43	55
31	PCNA	<i>P. abyssi</i> GE5	giI5458797	28 018	4.55	158	13/36	57

Tableau 11 : résultat analyse du protéome partiel

Cette étude est encore préliminaire. Par contre, il est intéressant de noter que :

- Le PCNA est identifié, ce qui souligne son rôle central dans le métabolisme de l'ADN.
- La protéine DnaG-like est présente en forte proportion

- La sous-unité alpha du protéasome est présente. Cette protéine est également retrouvée comme partenaire de RadA lors d'une expérience de *pull-down*, en utilisant RadA comme appât (résultat non présenté dans ce manuscrit). Bien que RadA n'ait pas été identifié lors de ce crible, ce résultat pourrait renforcer l'hypothèse que, de manière similaire aux eucaryotes, la quantité de RadA sous forme de filaments pourrait être régulée par le protéasome.
- D'autre part, les différents spots correspondant au PCNA et à DnaG laissent à penser que ces protéines subissent des modifications post-traductionnelles dont la nature n'est pas connue.

La mise au point de cette approche pourrait être optimisée, notamment pour ce qui concerne la préparation des extraits cellulaires, traiter dans les différentes conditions, afin de mettre en évidence plus de protéines spécifiquement liées à l'ADN.

* *

*

IV. Etude d'une ADN polymérase issue de virus thermophile

4.1. Introduction

Au cours de la thèse, j'ai collaboré avec les docteurs Viggó Marteinson et Ólafur Friðjónsson de la société Prokaria en Islande. Le but de cette étude était de cloner, exprimer et purifier une ADN polymérase issu d'un bactériophage thermophile, afin de déterminer les potentialités biotechnologiques de cette ADN polymérase.

Rhodothermus marinus est une Bactérie à coloration de Gram négatif, appartenant à l'ordre des Sphingobacteriales. Elle a été isolée à partir de sources chaudes sous-marines islandaises, son optimum de croissance est de 65°C (Bjornsdottir *et al.*, 2005). Comme tout être vivant, elle peut être infectée par des virus. Un bactériophage a été identifié au laboratoire Prokaria. Ce bactériophage appelé RM378 appartient à la famille des *Myoviridae*, qui correspond à la morphologie des phages tête-queue (groupe des phages T) (Hjorleifsdottir *et al.*, 2002). Son génome de 130 kb, qui correspond à de l'ADN double brin linéaire, a été séquencé afin d'y trouver notamment des gènes susceptibles de produire des protéines avec un intérêt biotechnologique. L'analyse de son génome a permis d'identifier plus de 200 ORFs et notamment deux gènes codant une ADN polymérase constitué de deux sous-unités : un gène de 1569 pb correspondant à une protéine de 522 acides aminés et de masse moléculaire théorique 61 kDa qui possède une activité polymérasique ; un gène de 942 pb codant une protéine de 313 acides aminés de masse moléculaire théorique 37 kDa possédant un domaine exonucléasique.

4.2. Clonage des deux sous-unités de l'ADN polymérase pour expression indépendante

Les gènes correspondant aux deux sous unités de l'ADN polymérase ont été clonés séparément dans des plasmides pET32 (carte du vecteur en annexe). Après induction de la production à l'IPTG dans les cellules d'expression *E. coli*, les cellules ont été lysées par sonication. Si la majorité de la sous unité polymérasique s'est trouvée dans la fraction insoluble, une part non négligeable de la protéine était trouvée dans la fraction soluble. Par contre, la sous unité exonucléasique était présente dans la fraction insoluble, ce qui laisse supposer des problèmes de repliement de cette protéine recombinante lors de l'expression dans *E. coli*.

La purification de la sous-unité polymérasique a donc été entreprise par chromatographie d'affinité sur colonne HisTrap. La figure 47 montre le profil protéique des fractions éluées à l'imidazole. La masse de la bande observée sur gel d'électrophorèse coloré au bleu de Coomassie migre entre 70 et 80 kDa, correspond à la masse théorique de 73 kDa, soit 61 kDa (ADN polymérase) + 12 kDa (thioredoxine fusionnée).

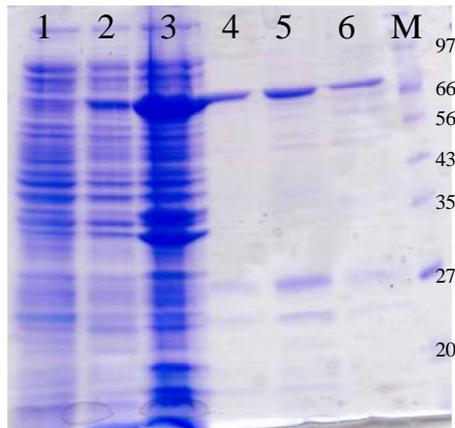


Figure 47 : Gel SDS-PAGE 12% permettant de visualiser l'expression et la purification de la sous unité polymérase à partir du lysat cellulaire de *E. coli*. Pistes : 1, cellules transformées par un plasmide non recombinant ; 2, fraction soluble du lysat cellulaire ; 3, fraction insoluble du lysat cellulaire ; 4, 5 et 6, protéines éluées après purification sur colonne HisTrap.

Des tests de mesure d'incorporation de dNTP radioactif sur matrice amorcée ont été entrepris. Ces tests ont été effectués à 50°C et à 60°C, sur les échantillons d'ADN polymérase avant et après purification. La figure 48 résume les résultats obtenus sur ces différentes fractions.

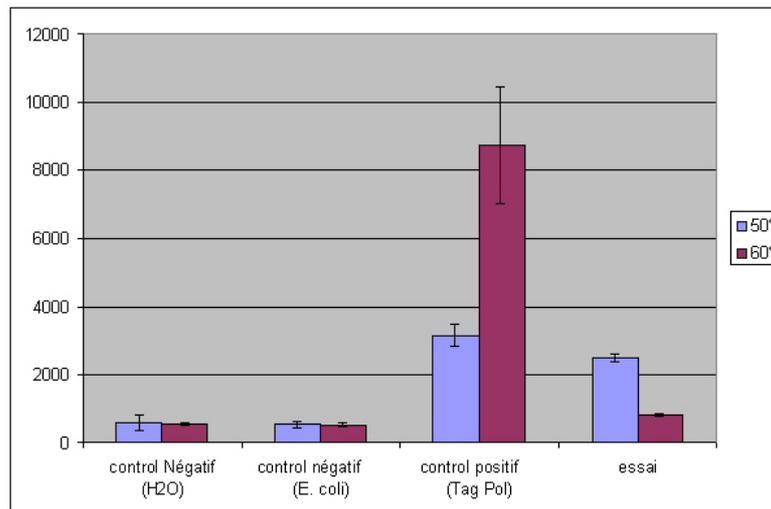


Figure 48 : activité de polymérisation de l'ADN de la sous-unité polymérasique (essai), avant purification, mesurée par scintillation liquide (radioactivité en cpm), à 50 et 60°C.

L'extrait cellulaire total de la souche de *E. coli* n'exprimant pas la protéine recombinante sert de témoin négatif, et permet de vérifier que les ADN polymérases de *E. coli* n'entraînent pas d'amplification positive à 50 et 60°C. La Tag Pol extraite de *Thermococcus aggregans* est une ADN polymérase thermostable utilisée ici comme témoin positif.

Ces résultats montrent que l'ADN polymérase virale présente une bonne activité à 50°C dans le surnageant bactérien (avant purification) mais une faible activité à 60°C. Par contre, cette protéine ne semble pas avoir d'activité significative après purification, dans les conditions testées, *in vitro*. Ce résultat peut paraître un peu décevant à première vue, mais une analyse plus approfondie mérite d'être engagée. En effet, l'ADN polymérase pourrait avoir

besoin de cofacteurs qui lui étaient fournis avant purification dans l'extrait bactérien. Le tampon de conservation de cette protéine après purification pourrait être mal adapté. La thioredoxine fusionnée pourrait aussi empêcher l'activité normale de l'ADN polymérase. Enfin, l'addition de la sous-unité exonucléasique pourrait améliorer l'activité et la stabilité à haute température (du complexe).

4.3. Clonage des deux sous-unités de l'ADN polymérase dans un même vecteur et sur-expression

Afin d'améliorer l'efficacité de la solubilisation des deux sous-unités lors de la production, de manière similaire à l'expression des protéines Mre11 et Rad50, les deux sous-unités de cette ADN polymérase ont été coexprimés dans une souche de *E. coli*.

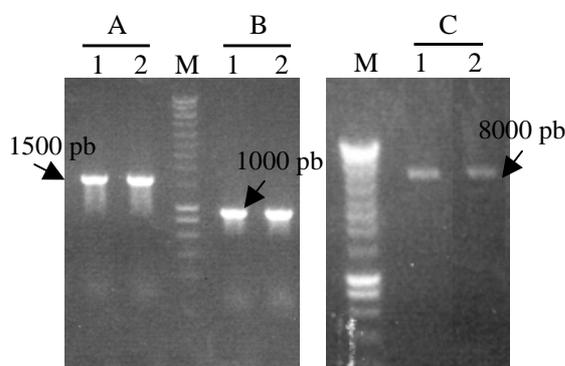


Figure 49 : vérification de la présence des deux gènes insérés dans le vecteur pET-30 par A, amplification du gène de la sous-unité polymérasique (1569 pb) à partir du plasmide recombinant ; B, amplification du gène de la sous-unité exonucléasique (942 pb) à partir du plasmide recombinant ; C, linéarisation du plasmide recombinant par l'enzyme de restriction SacI (la taille correspondant aux plasmides avec les deux inserts est de 8064 pb). Les pistes 1 et 2 correspondent aux plasmides isolés de deux clones différents de *E. coli*.

Un séquençage a été réalisé sur les ADN plasmidiques, afin de s'assurer qu'il n'y a pas eu de mutations au niveau des gènes d'intérêt. Le séquençage n'a pas révélé de mutations dans le gène codant le domaine exonucléasique, par contre trois mutations étaient présentes dans le gène correspondant au domaine polymérasique. Deux de ces mutations étaient silencieuses, les codons aboutissant au même acide aminé que la séquence originale. La mutation critique en position 744 du gène entraînait un codon AGC (codant une sérine) au lieu de AAC (codant une asparagine) ; elle a dû être corrigée par mutagenèse dirigée. Le kit Quick Change II Site Directed Mutagenesis (Stratagene) a permis la correction de cette mutation.

Le plasmide pET30-PolExo a été introduit dans des cellules d'expression *E. coli* BL21(DE3) et HMS174(DE3). L'expression des protéines a été testée à partir de ces cellules, à différentes températures, soit 25°C, 37°C et 42°C.

La figure 50 (piste 1 et 2) présente un test d'induction sur les souches BL21(DE3) à 25°C. Après lyse des cellules et migration des fractions solubles et insolubles sur gel SDS-PAGE, nous observons des bandes correspondant à la taille attendue des sous-unités exonucléasiques et polymérasiques, mais majoritairement présente sous forme insoluble.

Un western blot a été effectué sur les fractions solubles pour visualiser la forme recombinante de la sous unité polymérase. Cette sous unité présente en effet une étiquette histidine permettant une révélation grâce à l'utilisation d'anticorps anti-histidine. La figure 50 indique que cette sous unité est bien présente dans le surnageant de l'extrait cellulaire de *E. coli* après induction. Par contre, la sous-unité exonucléasique ne peut pas être visualisée par cette méthode car elle n'est pas marquée par cette étiquette histidine.

Une première étape de purification du complexe a enfin été effectuée sur la fraction soluble par *pull-down* en utilisant les billes magnétiques (Dynabeads® Talon™). Après coloration au bleu de Coomassie, il est possible d'observer la sous-unité polymérasique (61 kDa) co-purifiée avec la sous-unité exonucléasique (37 kDa) (figure 50, piste 3). Cependant, la purification n'est que partielle.

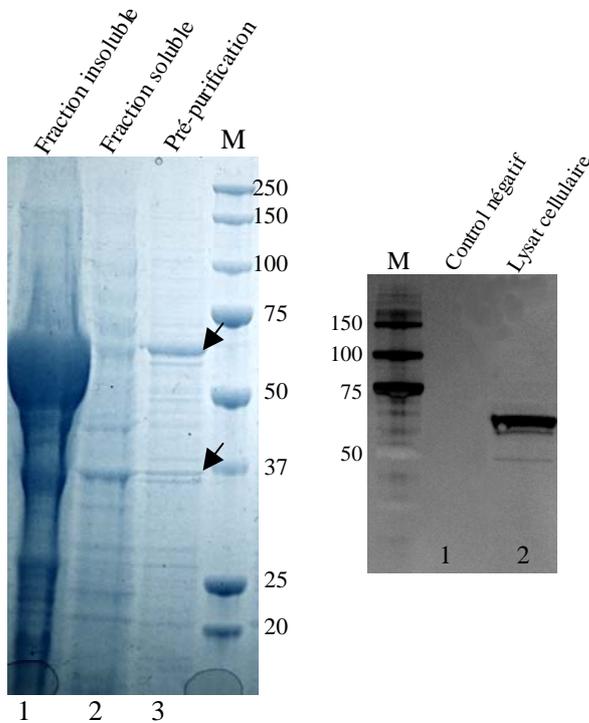


Figure 50 : analyse de la production des sous-unités polymérase et exonucléase par SDS-PAGE et western-blot.

A. Piste 1 : fraction insoluble du lysat cellulaire après 5 heures d'induction. Piste 2 : fraction soluble du lysat cellulaire. Piste 3 : fraction soluble après une première étape de purification sur billes magnétiques Co^{2+} (Dynabeads® Talon). Les sous unités polymérasique (61 kDa) et exonucléasique (37 kDa), indiquées par les flèches noires, apparaissent partiellement purifiées après coloration au bleu de Coomassie.

B. Piste 1 : le contrôle négatif correspond au lysat des cellules d'expression n'exprimant pas les protéines d'intérêt. Piste 2 : fraction soluble du lysat des cellules d'expression après induction. Seule la bande correspondant à la sous-unité polymérasique (61 kDa) est révélée par l'anticorps anti-histidine tag.

Ces travaux exploratoires sont plutôt encourageants. Les deux sous-unités d'une ADN polymérase ont été clonés dans un même vecteur, des tests de production à l'échelle pilote ont été effectués et les tests de purification préliminaires montrent qu'une co-purification des sous-unités semble possible. Faute de temps, les caractérisations biochimiques et fonctionnelles de cette ADN polymérase n'ont pu être réalisées au cours de ces travaux de thèse. Ces travaux préliminaires seront donc poursuivis par le laboratoire.

* *
*

Conclusions et Perspectives

A l'issue de ce travail, j'ai pu mettre en évidence un réseau d'interaction lié au métabolisme de l'ADN chez l'Archaea hyperthermophile *Pyrococcus abyssi*. Bien que préliminaire, ce réseau, représenté sur la figure 51, m'a permis d'identifier de nouveaux complexes impliqués dans la maintenance génomique.

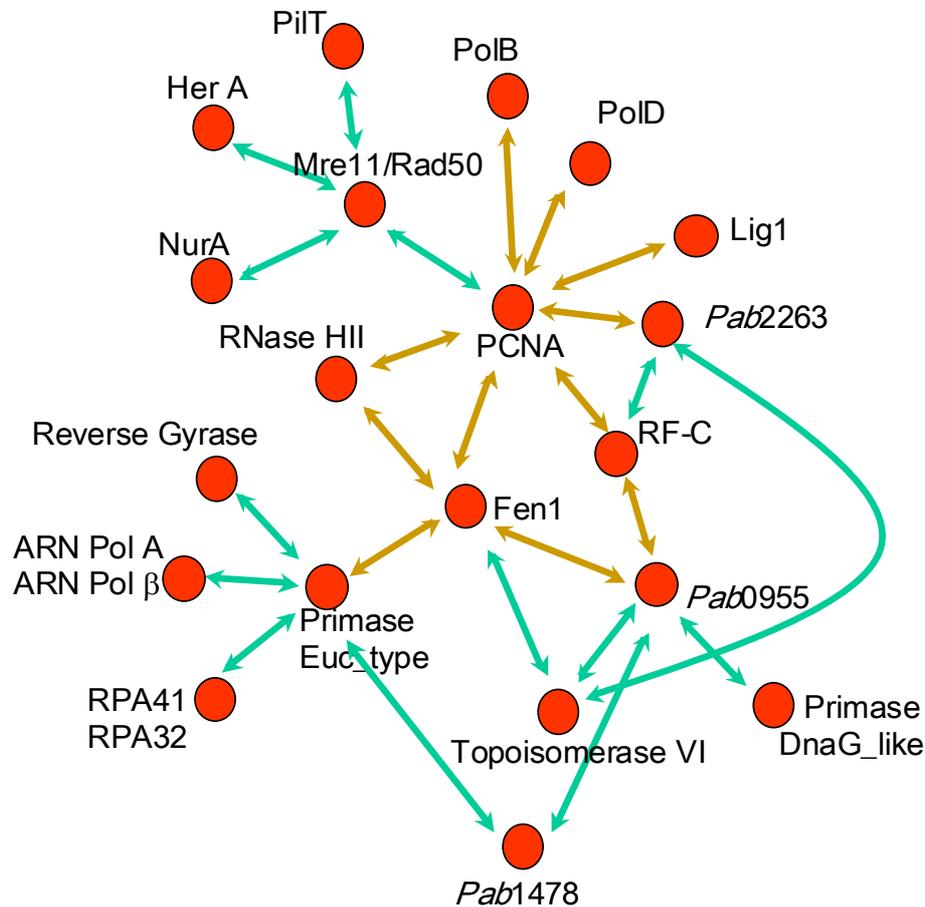


Figure 51 : réseau d'interaction montrant les protéines appâts testées et leurs partenaires identifiés au cours de ce travail. Le PCNA, Fen1, la sous-unité catalytique de la primase p41, le complexe Mre11/Rad50, *Pab0955*, *Pab2263*, ont été utilisés comme appâts dans les expériences de *pull-down*, lors de la capture de partenaires dans l'extrait cellulaire. Les flèches marron indiquent que l'interaction a été validée *in vitro*, entre protéines recombinantes, grâce à la technique de résonance plasmonique de surface ou par co-immunoprécipitation.

L'analyse globale de nos travaux nous permet de confirmer le rôle central des protéines PCNA (Moldovan *et al.*, 2007) et Fen1 (Liu *et al.*, 2004) dans la maintenance et la stabilité du génome des archées.

Les protéines appâts utilisées dans cette approche méthodologique interviennent dans la réplication de l'ADN, or la plupart des complexes découverts semblent avoir des implications dans les mécanismes de réparation. Le réseau de la maintenance génomique est combiné et coordonné par des interactions physiques avec des protéines du complexe de réplication. Nos résultats révèlent une intrication de ces processus. En effet, des erreurs de

fidélité peuvent apparaître lors de la duplication de l'ADN qui doivent être corrigées. Des lésions sur l'ADN peuvent bloquer la progression de la fourche de réplication et induire des voies de réparation et de recombinaison. Les modules fonctionnels, dans lesquels sont impliquées les protéines liées au métabolisme de l'ADN, semblent en perpétuelle interaction, afin de permettre l'organisation du vivant et l'adaptation aux conditions environnementales.

Les interactions mises en évidence apparaissent pour certaines comme potentiellement conservées chez les protéines d'archées et les protéines homologues eucaryotes comme Fen1 et le PCNA, le RF-C et le PCNA, et éventuellement Mre11/Rad50 et le PCNA. Par contre, l'interaction concernant les protéines Fen1 et l'ADN primase (p41) semble uniquement présente chez les archées. Ces résultats soulignent la dualité de ce modèle, qui permet à la fois de mieux comprendre des mécanismes de maintenance, mais également de découvrir de nouvelles voies de réparation de l'ADN. Dans ce contexte, il est à noter que cette étude ne permet pas de répondre aux questions, émanant de l'analyse génomique, sur les mécanismes de réparation de l'ADN, à savoir les protéines manquantes chez les *Archaea* comme celles dont le rôle est de reconnaître les lésions lors de la réparation par excision de nucléotides (NER) ou celles détectant les mésappariements dans la réparation des mésappariements (MMR) (Grogan, 2004). Cependant, les nouveaux complexes mis en évidence au cours de ce travail nous ont conduit à émettre de nouvelles hypothèses concernant la maintenance et la stabilité du génome.

Le PCNA pourrait associer le complexe Mre11/Rad50 à la réplication

Nous avons mis en évidence que le PCNA est associé au complexe ubiquitaire Mre11/Rad50 et aux protéines spécifiques des archées hyperthermophiles NurA et HerA. On peut penser que le PCNA recrute ces protéines intervenant dans des processus de réparation par recombinaison homologue. Ce phénomène serait profitable à la cellule lorsque la fourche de réplication crée des cassures double brin lors d'une rencontre avec une coupure sur un brin d'ADN, ou lors d'un blocage de l'avancée du replisome par une lésion sur l'ADN. Concernant cette étude, la recherche de laquelle des deux protéines, Mre11 ou Rad50, interagit directement avec le PCNA devra être entreprise. L'influence du PCNA sur l'activité biochimique du complexe Mre11/Rad50 sera menée en parallèle.

Les Protéines Fen1 et l'ADN primase des archées peuvent collaborer dans un mécanisme de réparation

Notre étude nous a permis de montrer que l'interaction physique entre les protéines Fen1 et la sous-unité catalytique de l'ADN primase p41, reflétait une interaction fonctionnelle. Cette observation conforte donc l'hypothèse selon laquelle l'ADN primase de

P. abyssi, habituellement connue comme participant uniquement à l'initiation de la synthèse d'ADN lors de la réplication, pourrait participer également à des mécanismes de réparation. Ainsi, si les protéines manquantes du NER restent à découvrir chez les Archées, nos résultats démontrent que la voie de réparation du BER est fonctionnelle. Cette voie doit être particulièrement importante chez les archées hyperthermophiles. En effet, le BER permet d'éliminer les bases modifiées par des processus chimiques, et le taux de ces réactions chimiques augmente avec l'augmentation de la température.

Il y a plusieurs perspectives à cette étude. Il serait possible d'effectuer des expériences de réparation de l'ADN *in vitro* en utilisant le lysat cellulaire de *P. abyssi* en immunodéplétant des protéines Fen1, l'ADN primase ou les autres ADN polymérases sur les différents substrats utilisés dans cette étude. Il serait, par ailleurs, intéressant de déterminer si l'ADN primase des archées exprime l'activité d-Rp (5'-deoxyribose-5-phosphate) lyase connue chez les ADN polymérase B2 et β . Cette activité peut en effet être impliquée dans la réparation par excision de base. Enfin, la coordination de l'ADN primase et de Fen1 avec les autres facteurs et enzymes du BER (ADN glycosylase, AP endonucléase et PCNA) sera étudiée.

Une nouvelle approche de protéomique permet la découverte d'une protéine liée au métabolisme de l'ADN

Par ailleurs, cette étude a également permis de valider une nouvelle approche protéomique originale pour déterminer, chez l'Homme et les archées, les partenaires du PCNA. Outre l'aspect technologique, cette approche nous a permis de mettre en évidence de nouvelles interactions entre les protéines impliquées dans la maturation des fragments d'Okazaki, mais également de lever des hypothèses sur de nouveaux acteurs du métabolisme de l'ADN. C'est le cas de *Pab2263*, une nouvelle nucléase, partenaire du PCNA, et présentant de l'affinité pour l'ADN simple brin. La caractérisation biochimique et la détermination de la structure tridimensionnelle de cette protéine fait actuellement l'objet d'une étude, en collaboration avec les docteurs H. Myllykallio et R. Ladenstein, afin de déterminer la fonction et la structure de *Pab2263*.

Vers une vision à plus grande échelle

A la suite de mes travaux, la production de protéines totales de *P. abyssi* en grande quantité, couplé à la recherche de partenaires par affinité est maintenant optimisée au laboratoire. Cette recherche peut, dès lors, être étendue à plus grande échelle. Dans ce contexte, nous sommes partenaires du programme « Marine Express » (réseau européen Marine Genomics) dont le but est de produire, à moyen débit, des protéines de modèles marins. Le laboratoire devrait disposer, à court terme, d'environ 40 protéines recombinantes

de *P. abyssi* présentant des fonctions potentielles dans les mécanismes de maintenance du génome. Ces protéines seront utilisées comme des nouveaux appâts afin d'élargir le réseau préliminaire identifié. Les protéines RPA, associées à l'ADN simple brin, sont susceptibles d'avoir un grand nombre d'interactions avec des composants des machineries de réplication et de réparation tout en participant à la coordination de ces mécanismes. La reverse gyrase, spécifique des thermophiles, pourrait participer, de manière directe ou indirecte, à la réponse cellulaire engendrée par la présence de lésions dans l'ADN. L'identification des complexes, dans lesquelles sont impliquées ces protéines, nous fournirait de nouvelles hypothèses de travail pour affiner nos connaissances et ouvrir de nouvelles perspectives. La production des polymérase B et D, fusionnées à une étiquette, histidine est également en bonne voie. L'identification des partenaires de ces composants majeurs du complexe de réplication sera couplée à leurs caractérisations fonctionnelles et à leurs comportements face à différents types de substrats. Ces deux polymérase semblent essentielles à la réplication de l'ADN (Henneke *et al.*, 2005), mais leurs implications dans des mécanismes éventuels de réparation restent encore à déterminer.

A plus long terme, l'analyse du réseau élargi nous permettra d'avoir une vision plus globale des mécanismes, des liens et des régulations qui sous-tendent les voies de réparation et de réplication de l'ADN au sein de la cellule.

Valorisations biotechnologiques potentielles

L'observation du mécanisme de réparation *in vitro* avec les protéines recombinantes ADN primase, Fen1 et ADN ligase de *P. abyssi* peut aboutir à des applications biotechnologiques originales. La récupération et l'amplification par PCR d'ADN ancien présente parfois un intérêt en archéologie, en paléontologie ou en médecine légale. Cependant, l'ADN ancien est fréquemment endommagé. Ces dommages incluent des modifications (oxydation et déamination) ou des pertes de bases (sites apuriques ou apyrimidiques) ou des cassures simples brins, des brèches de longueur variable (Mitchell *et al.*, 2005). Les techniques d'extraction d'ADN peuvent également entraîner des dommages. Ces lésions peuvent interférer et empêcher l'amplification par les enzymes de PCR. Une stratégie envisagée est de réparer l'ADN endommagé, préalablement à l'amplification par PCR. Une étude présente, par exemple, un comparatif d'amplification de l'ADN sur des échantillons provenant d'animaux de Pompei et Herculaneum, traité de façon préalable, ou non, par une polymérase I et une ADN ligase T4. Seuls les échantillons prétraités présentaient un résultat positif après amplification par PCR (Di Bernardo *et al.*, 2002). Des essais pour amplifier de l'ADN présentant différentes tailles de brèches, en utilisant préalablement les protéines ADN primase, Fen1 et ADN ligase, sont tout à fait envisageables.

La production de l'ADN polymérase du virus RM378 présente également des perspectives intéressantes. Une caractérisation de l'activité de ces enzymes sera effectuée. La composition du tampon ainsi que les conditions de réaction devront être optimisées. La thermostabilité et la fidélité de l'activité de polymérisation de l'ADN sera étudié.

L'ADN polymérase virale pourra également être utilisé comme appât pour rechercher des partenaires soit au sein d'un lysat cellulaire de *Rhodothermus marinus*, soit au sein d'un lysat viral. Sur le plan fondamental, cette étude nous permettrait de mieux comprendre la réplication du virus, sur le plan appliqué, les partenaires identifiés pourraient être utilisés pour améliorer la processivité de l'ADN polymérase lors d'utilisation en PCR. Par exemple, la protéine gp43 du bactériophage T4 de *E. coli*, est une ADN polymérase considéré comme fidèle et processive, et cette processivité est amélioré en présence de son facteur de processivité gp45. Le clonage et la production du facteur de processivité de l'ADN polymérase du virus RM378 pourra également être envisagé. Ainsi, l'utilisation des protéines thermostables du virus RM378, pour la synthèse d'ADN *in vitro*, présenterait un atout pour des applications biotechnologiques.

Perspectives personnelles

Je pourrai enfin valoriser les compétences acquises au cours de ce doctorat et m'initier à de nouveaux savoir-faire lors de ma future expérience post-doctorale à l'Université de Montréal. Je pourrai lors de cette expérience, mêler les notions d'interaction entre protéines et d'interaction hôte-pathogène. Je travaillerai notamment sur une protéine accessoire du virus de l'immunodéficience humaine et ces partenaires au sein de la cellule hôte. Il a en effet été montré, par une approche de TAP *pull-down*, que cette protéine virale est associée à une protéine de la cellule hôte qui interviendrait dans la maintenance du génome. Cette interaction entre la protéine virale et la protéine de l'hôte participe au blocage du cycle cellulaire.

* *
*

Matériels et méthodes

I. Techniques de Microbiologie

1.1. Culture de *Pyrococcus abyssi*

1.1.1. Culture de *P. abyssi* en batch

Les cultures de *Pyrococcus abyssi* (souche GE5) ont été réalisées en milieu liquide SME, dans des conditions d'anaérobiose stricte, à 95°C, pH 6,8 en fiole de type flacon pénicilline. Un volume de 50 mL de culture a été inoculé avec 2 mL de préculture contenant entre 10^8 et 10^9 cellules par mL. La densité cellulaire était évaluée par comptage direct par l'intermédiaire d'une cellule de Thomas.

Le milieu SME (composition exacte dans l'annexe) est un milieu riche contenant des sucres et acides aminés (extrait de levure, peptone), des sels minéraux, des vitamines et du soufre. Le soufre peut être remplacé par la L-cystéine quand les cultures sont réalisées en bioréacteur gas-lift (culture en continu). La résazurine est un indicateur de l'état de réduction du milieu et permet ainsi de visualiser l'anaérobiose du milieu. La résazurine est bleue à l'état oxydé, rose à 50% d'oxydation et incolore à l'état réduit (condition anaérobie). Le PIPES (acide piperazine-N, N'-bis-2-ethanesulfonique) est un tampon utilisable pour une gamme de pH de 6,1 à 7,5.

La stérilisation du milieu contenant du soufre s'effectue par tyndallisation. Cette étape consiste en deux séries de chauffage à 100°C dans l'autoclave pendant 30 min, séparées par une période de refroidissement complet. L'autoclavage n'est en effet pas possible car le soufre est sensible à la température. Afin de réaliser des cultures en anaérobiose, les milieux de culture sont répartis dans les flacons de culture à l'intérieur de l'enceinte anaérobie contenant un mélange gazeux composé de $N_2/H_2/CO_2$ dans les proportions 90/5/5. Milieux et matériels sont introduits dans l'enceinte par l'intermédiaire d'un sas. Le milieu de culture est réduit dans l'enceinte grâce à l'addition de 20 mL par litre de milieu de Na_2S à 2,5% (correspondant à une concentration finale de 5 mg/mL).

Les cultures en milieu liquide ont été réalisées dans des flacons pénicilline de 50 ou 100 mL. Ces flacons sont remplis respectivement avec 20 ou 50 mL de milieu de culture afin de garder un volume pour la phase gazeuse.

1.1.2. Culture de *P. abyssi* en bioréacteur

Afin de produire *P. abyssi* en grande quantité, une culture continue en bioréacteur *gas-lift* (figure 52) a été réalisée. Celui-ci a été mis au point pour la culture en continu de micro-organismes hyperthermophiles soufre-réducteurs (Godfroy *et al.*, 2006). Ce procédé permet l'apport de milieu nutritif frais en continu pendant que le milieu de culture dans le réacteur est soutiré simultanément en permanence. A l'équilibre, le volume du milieu de culture dans la

cuve et la densité microbienne sont ainsi maintenus constants. Le balayage gazeux à l'azote de la culture dans le bioréacteur gas-lift permet d'une part d'établir les conditions d'anaérobiose et d'autre part, de favoriser l'homogénéisation du milieu de culture et l'élimination des composés toxique (tel l' H_2S) élaborés par *P. abyssii*. Le procédé permet également de contrôler le pH, la température, les débits d'entrée et de sortie du milieu de culture.

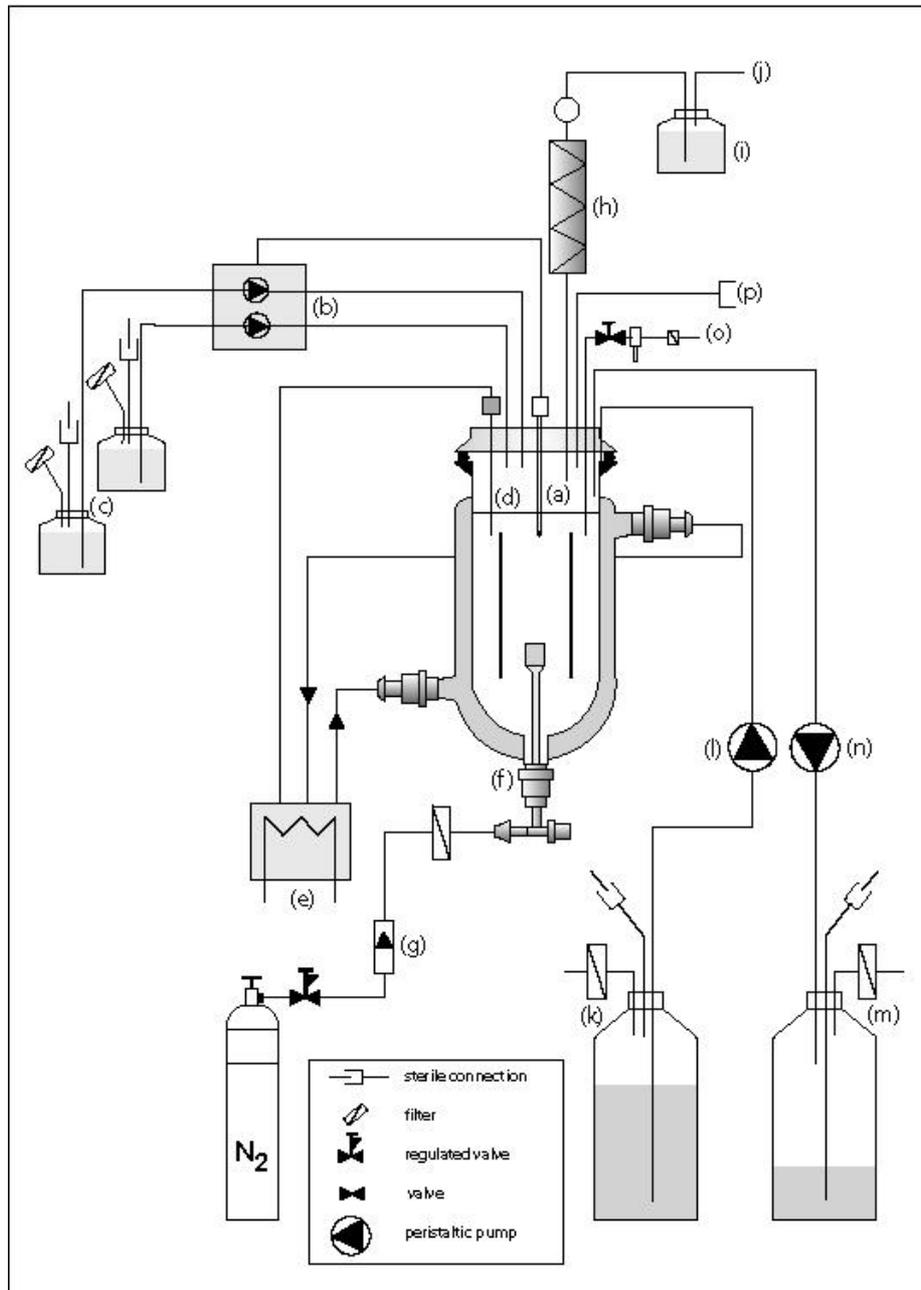


Figure 52 : schéma représentant le système expérimental pour la culture en continue de *P. abyssii* en bioréacteur *gas-lift*. (a) sonde à pH ; (b) contrôleur du pH ; (c) bouteilles d'acide et base ; (d) sonde de température ; (e) bain thermostaté contrôlant la température ; (f) robinet ; (g) débitmètre à gaz ; (h) condenseur ; (i) bullage dans l'eau ; (j) échappement d'eau ; (k) bouteille du milieu de culture ; (l) pompe ; (m) bouteille récupérant le milieu de culture ; (n) pompe ; (o) système de prélèvement d'échantillons ; (p) système pour l'inoculation (Godfroy *et al.*, 2006).

Les conditions de culture choisies étaient adaptées à la croissance de *P. abyssi* : une température de 90°C, un pH de 6,5, un balayage gazeux de la culture à l'azote permettait de maintenir l'anaérobiose tout en homogénéisant le milieu de culture. Un taux de dilution de 0,375 h⁻¹ (soit 750 mL.h⁻¹) a été appliqué, permettant d'apporter des nutriments en continu et de récupérer 18 L de milieu de culture par 24 heures. La densité cellulaire dans le fermenteur au cours de la production variait entre 1,5.10⁸ et 2.10⁸ cellules par mL. En sept jours de culture continue, après centrifugation, il a été possible d'obtenir 100 g de culots cellulaires. Après lyse cellulaire par sonication, nous avons récupéré 3,5 g de protéines solubles (approximativement dans 350 mL ; 10mg/mL).

1.2. Culture d'*Escherichia coli*

La bactérie *E. coli* a été utilisée pour le clonage et la production de protéines recombinantes. Les cellules compétentes NovaBlue GigaSinglesTM (Novagen) ont été utilisées pour le clonage. Les souches BL21-(DE3)-Codon Plus (Stratagene), HMS174-(DE3) et Rosetta(DE3) (Novagen) ont été utilisées pour la production des protéines régulé par le promoteur T7 et inductible à l'IPTG.

Les bactéries ont été cultivées à 37°C, en milieu LB (Luria Bertani) liquide sous agitation, ou en milieu gélosé en boîte de Petri, contenant l'antibiotique approprié pour sélectionner et/ou conserver les plasmides recombinants d'intérêts.

Préparation du milieu LB

Composants	Quantité pour 1 L
Tryptone	10 g
Extrait de levure	5 g
NaCl	5 g
Agar (milieu solide)	15 g

Antibiotique de sélection

Antibiotique	Concentration Solution Stock	Concentration de travail
Ampicilline	100 mg/mL dans H ₂ O	100 µg/mL
Kanamycine	30 mg/mL dans H ₂ O	30 µg/mL

* *
*

II. Techniques de Biologie Moléculaire

2.1. Extraction de l'ADN génomique de *P. abyssi*

L'extraction de l'ADN totale de *P. abyssi* comprend les étapes de lyse cellulaire, d'extraction de l'ADN, d'élimination des ARN et de précipitation de l'ADN génomique.

Une culture de 50 mL de *P. abyssi* a été arrêtée en phase exponentielle (entre 2.10^8 et 3.10^8 cellules/mL), puis centrifugée à 8000 g pendant 15 minutes à 4°C. Le culot a été repris dans 5 mL de tampon de lyse contenant du Tris pH 8,0 100 mM ; EDTA 50 mM ; NaCl 100 mM. Les cellules ont été lysées par ajout de 500µL de Sarkosyl 10% (1% final), 500µL de SDS 10% (1% final) et 100µL de protéinase K 20 mg/mL (0,4 mg/mL final). Ce mélange a été incubé pendant 3 heures à 45°C, puis centrifugé à 4800 g pendant 5 minutes. Un volume de PCI (25% Phénol, 24% chloroforme, 1% alcool isoamylique) a été ajouté pour un volume de surnageant de lyse, avant agitation par retournement jusqu'à émulsion. Après centrifugation à 8000 g pendant 20 min, la phase aqueuse a été récupérée et 10µg/mL de RNase lui a été ajouté afin d'éliminer les ARN. Une seconde extraction au PCI a ensuite été réalisée dans les mêmes conditions que la première. La phase aqueuse récupérée doit être limpide. L'ADN a ensuite été précipité en ajoutant 0,8 volumes d'isopropanol à la phase aqueuse puis placé pendant une nuit à -20°C, avant d'être centrifugé à 10 000 rpm pendant 30 min à 4°C. Après avoir été séchés au Speed vac, les culots ont finalement été repris dans 200 µL de TE (Tris 10 mM ; EDTA 1 mM). L'ADN génomique obtenu a été observé sur gel d'agarose 1%.

2.2. Amplification par PCR

Les réactions de polymérisation en chaîne (PCR) effectuées au cours de ce travail ont été réalisées sur un thermocycleur RoboCycler Gradient 96 (Stratagene).

Les gènes d'intérêts sont sélectionnés et amplifiés, à partir de l'ADN génomique, dans un mélange réactionnel d'un volume total de 50µL, en utilisant des amorces spécifiques et une ADN polymérase hautement fidèle afin d'obtenir des amplicons sans mutation. La polymérase *Pfu* (Promega), provenant de *Pyrococcus furiosus*, a été utilisée au cours de ce travail. Elle possède une activité correctrice d'erreur (activité exo 3' → 5') et une processivité de 500 pb/min. Les oligonucléotides de synthèse servant d'amorce provenait de la société Eurogentec. Le mélange réactionnel contenait 0,1 µM d'amorces, 100 µM de dNTP, 3 unités de *Pfu* polymérase. La réaction est initiée par une étape de dénaturation de 10 min à 94°C, puis les étapes du cycle sont les suivantes : une étape de dénaturation à 94°C pendant 30 secondes ; une étape d'hybridation au cours de laquelle un gradient de température entre 52°C et 63°C est programmé à l'aide du thermocycleur, afin de repérer les conditions optimales

d'hybridation des amorces ; une étape d'élongation à 72°C pendant 5 minutes. 25 cycles sont ainsi réalisés.

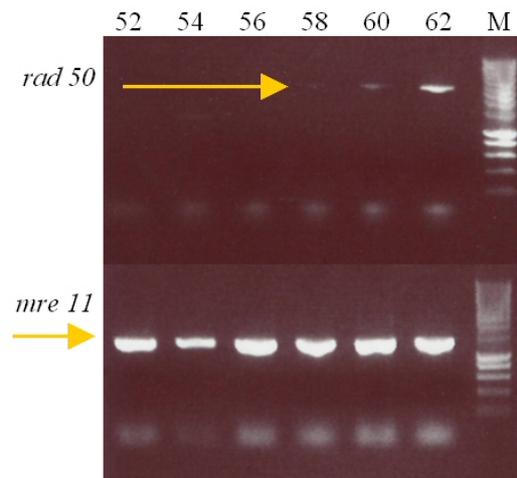


Figure 53 : visualisation de la présence des amplicons, issus de la PCR en gradient de température, sur gel d'agarose 1%. Les températures d'hybridation de 52 à 62°C sont indiquées.

2.3. Séparation et purification des fragments d'ADN

Dans des conditions de neutralité ou alcalines, les groupements phosphates de l'acide phosphorique de l'ADN sont ionisés. Les chaînes d'ADN constituent des polyanions qui, lorsqu'ils sont soumis à un champ électrique, se déplacent vers l'anode (reliée au pôle positif du générateur extérieur dans le cas d'un électrolyseur). L'électrophorèse en gel d'agarose permet de séparer des fragments d'ADN de taille comprise entre 0,2 et 50 kpb. Au cours de cette étude, les fragments d'ADN sont mis dans la solution de dépôt 6X (Tris-HCl 20 mM, EDTA 1 mM, glycérol 50% et 0,07% de bleu de bromophénol) puis déposés sur gel d'agarose 1% dans un tampon Tris-acétate TAE 1X (Tris-base 40 mM, EDTA pH8 1 mM, acide acétique glacial 0,1%). L'électrophorèse est réalisée sous voltage de 90 V. Le gel d'agarose contient du bromure d'éthidium (BET) qui est un agent s'intercalant dans l'ADN. Le BET présente la propriété d'émettre une fluorescence lorsqu'il est excité en lumière ultra-violette. Les bandes obtenues sont visualisées sous UV avec l'appareil GEL Doc 2000 Biorad, couplé à une caméra CCD ou avec le Fluor-STM MultiImager (Bio-Rad) permettant d'acquérir une image numérique.

L'estimation de la taille et de la quantité des fragments d'ADN ainsi séparés est effectuée grâce à un marqueur de taille (SmartLadder 200 - 10 000 pb, Eurogentec) (figure 54).

Tampon TAE 10X

Composants	Concentration	Quantité pour 1 L
Tris base	400 mM	48,4 g
EDTA	10 mM	20 mL EDTA 0,5 M
Acide acétique		57,1 mL

Tampon de dépôt

Composants	Concentration	Quantité pour 20 mL
Ficoll	20% Ficoll 400	4 g
Na ₂ EDTA	0.1M Na ₂ EDTA pH8.0	10 ml Na ₂ EDTA 0.2M pH8.0
SDS	1%	0,2 g
Bromophenol blue	0,25%	5 mL Bromophenol blue 1%

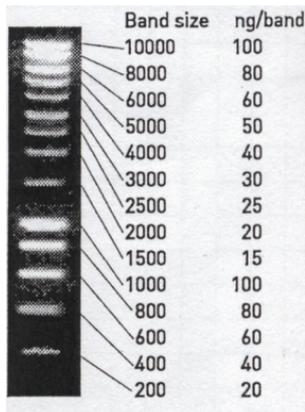


Figure 54 : Marqueur de poids moléculaire SmartLadder (Eurogenetec).

Les bandes d'ADN peuvent être récupérées en les découpant grâce à un scalpel, puis l'ADN est purifié sur colonne d'affinité avec l'aide du kit GeneClean Turbo kit (Q.biogene). Au cours de cette méthode, l'agarose est d'abord solubilisé, puis l'ADN se fixe à une colonne de silice à des concentrations élevées en sels chaotropiques. L'ADN est ensuite élué quand la concentration saline diminue avec un tampon peu salé ou avec de l'eau.

2.4. Techniques de Clonage

2.4.1. Vecteur

Le vecteur pET-30 Ek/LIC (Novagen) est un vecteur d'expression permettant un haut niveau d'expression des protéines (figure 55). Ce vecteur est commercialisé sous forme linéaire et permet le clonage d'un ou plusieurs gènes. Il confère une résistance à la kanamycine à la bactérie transformée. Il contient en amont du site de clonage le promoteur *T7lac* de l'ARN polymérase du Phage T7. Ce promoteur permet l'expression de protéines dans des bactéries hôtes de type DE3, possédant dans leur génome le gène codant l'ARN polymérase T7, ce gène étant placé sous contrôle du promoteur *lacUV5*, réprimé par le répresseur Lac I et induit par l'IPTG.

2.4.2. Clonage de deux gènes dans un même vecteur

Le système LIC Duet™ Adaptor Kits de la société Novagen, a été choisi pour cloner simultanément deux ORF permettant la coexpression des deux protéines dans *E. coli*. Cette stratégie permet le clonage direct des produits de PCR sans avoir recours à une enzyme de restriction ni réactions de ligation avant transformation, mais simplement par emboîtement entre extrémités cohésives dans les amorces PCR. Cette méthode favorise la solubilité et la stabilité de deux protéines formant un complexe, lors de la co-expression dans *E. coli*. La conception des amorces, pour amplifier les gènes à insérer, a été réalisée en fonction du kit de clonage. Le vecteur pET-30 Ek/LIC est utilisé pour le clonage des deux inserts (figure 25). Les inserts sont séparés par un adaptateur portant un promoteur *T7lac*, une séquence RBS (*ribosome binding sequence*) (figure 56).

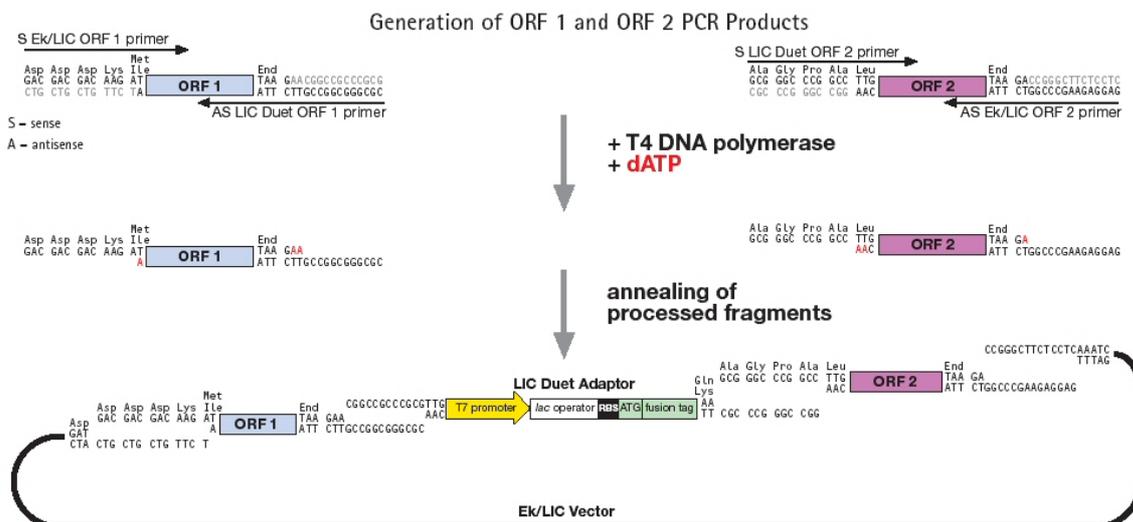


Figure 56 : Stratégie LIC Duet™ pour le clonage simultané de deux ORF.

2.4.3. Traitement des inserts par l'ADN polymérase T4

Le traitement par l'ADN polymérase T4, en présence d'un seul dNTP, permet de créer des extrémités cohésives dans les amorces PCR flanquant les ORF des gènes amplifiés. L'activité exonucléase 3'→5' de l'ADN polymérase T4 permet d'enlever les nucléotides aux extrémités des produits de PCR, jusqu'à rencontrer le premier résidu correspondant au seul dNTP présent dans le mix réactionnel (figure 29).

2.4.4. Réaction d'annealing du vecteur, de l'adaptateur et des inserts

L'annealing du vecteur avec les inserts et l'adaptateur a lieu par emboîtement des extrémités cohésives de chaque partenaire, comme indiqué dans le protocole fourni par

Novagen. Une étape de chauffage à 70°C pendant 30 seconde, puis l'ajout d'EDTA après 2 min à température ambiante favorisent l'hybridation entre les extrémités cohésives.

2.4.5. Purification de l'ADN plasmidique

Les minipréparations d'ADN plasmidiques ont été réalisées à l'aide du kit Qiagen Plasmid Purification. Cette technique est basée sur la technique à la lyse alcaline. Le SDS et la soude participent à la lyse alcaline des parois bactériennes. La soude dénature l'ADN plasmidique et chromosomique puis est neutralisée par l'acétate de sodium qui favorise également la renaturation de l'ADN plasmidique circulaire. L'ADN est lié sur une résine échangeuse d'anion dans des conditions appropriées de pH et à faible concentration en sel. L'ARN, les protéines et les molécules de bas poids moléculaire sont enlevés au cours de l'étape de lavage. L'ADN est élué avec un tampon hautement salé, puis concentré et désalé par une précipitation à l'isopropanol.

La purification de l'ADN plasmidique a été réalisée selon les recommandations du protocole fournies par Qiagen, à partir de 2 mL de culture de *E. coli*.

2.4.6. Séquençage de l'ADN

Un séquençage a été réalisé sur les ADN plasmidiques afin de s'assurer qu'il n'y a pas eu de mutations au niveau des gènes d'intérêts. Les amorces flanquantes à l'insert du vecteur pET-30 Ek/LIC, *T7 promoteur primer* #69348-3 et *T7 terminator primer* #69337-3 ont été utilisées. Ce séquençage a été effectué par la société Cogenics - Genome express (Grenoble). Le principe du séquençage est celui développé par Fred Sanger, basé sur la synthèse d'ADN en présence de didésoxynucléotides et réalisé actuellement grâce à des séquenceurs automatiques et à l'utilisation de fluorochromes.

2.4.7. Transformation des cellules compétentes

2.4.7.1. Transformation en cellules de stockage

Initialement, les ORF ont été clonés dans des cellules hôtes incapables d'exprimer les protéines recombinantes pour éviter des problèmes de toxicité de ces protéines qui pourrait interférer avec le fonctionnement cellulaire. Les cellules compétentes NovaBlue GigaSinglesTM fournies avec le kit LIC DuetTM Adaptator ont été utilisées pour le clonage initial avec le vecteur pET-30 Ek/LIC car elles présentent une grande efficacité de transformation. La transformation est réalisée par choc thermique. Le choc thermique fragilise les parois des bactéries compétentes, permettant aux plasmides recombinants d'intégrer les cellules. 1 µL de vecteur (réaction d'*annealing* précédente) a été mis en présence de 50 µL de cellules compétentes. Ce mélange a été incubé pendant 10 minutes à 4°C, sur la glace, puis le

vecteur est transfecté par choc thermique à 42°C pendant 45 secondes, suivi d'une incubation sur la glace pendant 5 minutes. 200 µL de milieu SOC ou de milieu LB ont ensuite été ajoutés, puis cette suspension a été incubé 1 heure à 37°C, sous agitation à 250 rpm. 100 µL de suspension bactérienne ont ensuite été étalés par boîtes de Pétri contenant du milieu LB agar avec de la kanamycine.

Les clones obtenues peuvent être contrôlés par extraction plasmidique associée à une double digestion et par séquençage. Les clones positifs, portant l'insert désiré, ont été stockés à -80°C, en milieu LB additionné de 20% de glycérol.

2.4.7.2. Transformation en cellules d'expression

Les cellules d'expression ont également été transformées par choc thermique avec une quantité comprise entre 100 ng et 1 µg de plasmides recombinants extraits des cellules de clonage et purification. Les souches d'*E. coli* transformées sont BL21(DE3)-Ril codon plus (Stratagene), HMS174 (DE3) et Rosetta(DE3) (Novagen).

* *
*

III. Technique de Biochimie

3.1. Technique de production des protéines recombinantes

3.1.1. Production des protéines en petit volume

Les premières productions de protéines ont été effectuées à l'échelle pilote, dans des volumes de 50 mL de culture, afin de tester plusieurs paramètres avant de passer à la production en grands volumes. Une pré-culture des cellules d'expression a été réalisée à 37°C, sur la nuit, dans 3 mL de milieu LB + antibiotique(s). Le lendemain, 50 mL de milieu LB ont étéensemencés avec 1 mL de pré-culture à 37°C sous agitation jusqu'à atteindre une DO_{600nm} de 0,6. L'ajout d'IPTG, à une concentration finale de 1 mM, permet l'induction de l'expression. Différentes températures d'expression avec les différentes souches d'expression ont été testées : 15°C, 25°C, 37°C et 42°C. La diminution de la température de croissance des bactéries lors de la production des protéines recombinantes permet souvent, en diminuant le taux de production de ces protéines, d'améliorer leurs repliements lors de leurs synthèses et éviter ainsi leurs séquestrations dans des corps d'inclusions insolubles. D'autre part, le choc thermique 37°C→42°C induirait l'expression de protéines de réponse au stress (*heat shock proteins*) qui pourraient aider au bon repliement des protéines d'intérêt et améliorer leur solubilité (Kube *et al.*, 2006). Dans tous les cas, des prélèvements de 2 mL ont été effectués régulièrement 2 h, 3h, 4h et une nuit après induction. Ces suspensions ont ensuite été centrifugées pendant 10 minutes à 12 000 rpm sur une petite centrifugeuse de paillasse (Jouan A14). Les culots étaient repris dans 500 µL de tampon PBS puis ont subi une lyse par sonication (sonicateur vibracell, 375 W, 40% amplitude) par injection de pulse de 0,5 seconde d'ultrasons pendant 1 à 2 minutes. Lors de cette étape, les tubes ont été maintenus dans la glace pour éviter la dégradation des protéines. Une centrifugation de 8 minutes à 12 000 rpm a ensuite permis de sédimenter les débris cellulaires, les protéines solubles restant dans le surnageant. Afin de visualiser l'expression des protéines d'intérêt par les cellules hôtes, les protéines des surnageants et culots sont séparées par électrophorèse sur gel de polyacrylamide 12%, suivie d'une coloration au bleu de Coomassie.

Tampon PBS

Composants	Concentration	Quantité pour 1 L
NaCl	140 mM	8 g
KCl	2,7 mM	0,2 g
Na ₂ HPO ₄ , 7 H ₂ O	4 mM	1,15 g
K ₂ HPO ₄	1,5 mM	0,2 g

3.1.2. Production des protéines en grand volume

Après les expressions à l'échelle pilote, les cellules et les conditions d'expression présentant le meilleur rendement ont été choisies pour la production en grand volume. Les conditions expérimentales sont similaires, les pré-cultures se font en cascades jusqu'à la phase de production proprement dite, dans 6 litres de milieu LB. Après une nuit d'expression, les cultures bactériennes ont été centrifugées à 5 000 g pendant 45 minutes (centrifugeuse KR 422) puis les culots cellulaires ont été repris dans du tampon de broyage (Tris-HCl pH 8,0 20 mM ; imidazole 20 mM, DTT 1 mM compléter d'antiprotéase : *Complete Protease Inhibitor Cocktail* Roche Applied Science). La lyse des cellules a été réalisée dans ce cas avec une presse de French (Aminco®-SLM Instruments). Trois passages successifs sur la presse de French ont été réalisés à une pression de 1000 psi.

3.2. Purification des protéines

3.2.1. Précipitation des protéines non thermostables

Les surnageants issus de la lyse des culots cellulaires ont été traités à 80°C pendant 20 minutes au bain-marie dans le but de faire précipiter les protéines non thermostables de *E. coli*. Les surnageants contenant les protéines d'intérêt thermostables sont récupérés après centrifugation à 14 000 g pendant 10 min. Cette étape de purification n'est que partielle mais permet de clarifier grandement l'extrait protéique en se débarrassant d'une grande partie des protéines de *E. coli*.

3.2.2. Purification par chromatographie

La purification par chromatographie FPLC (*Fast Protein Liquid Chromatography*) a été réalisée sur une production de protéines recombinantes en grand volume (6 litres). Les cellules ont été lysées à la presse de French, l'extrait protéique dans du tampon de broyage (Tris-HCl pH 8,0 20 mM ; imidazole 20 mM, DTT 1 mM + antiprotéase) a été clarifié par précipitation des protéines non thermostables à 80°C.

Le système ÄktaFPLCTM (Amersham Biosciences) (figure 57) est un système de chromatographie liquide qui a été utilisé pour la purification des protéines lors de cette étude. La phase stationnaire en chromatographie FPLC est une colonne d'affinité pour queue histidine renfermant du nickel (*HisTrap column* fournie par Amersham). Celle-ci a été équilibrée par un tampon d'équilibration (tampon A : Tris-HCl pH 8,0 20 mM, imidazole 20 mM, DTT 1 mM) avant passage de l'échantillon à purifier. Les protéines produites ont ensuite été chargées sur cette colonne, puis éluées par un gradient d'imidazole à un débit de 1 mL/min avec du tampon d'éluion (tampon B : Tris-HCl pH 8,0 20 mM ; imidazole 1 M ; DTT 1 mM ; NaCl 150 mM). Les fractions de 1 mL sont récupérées dans un rack et sont

ensuite analysées par SDS-PAGE. La visualisation des profils de migration de ces fractions a permis de localiser les protéines d'intérêts au niveau du profil d'éluion.

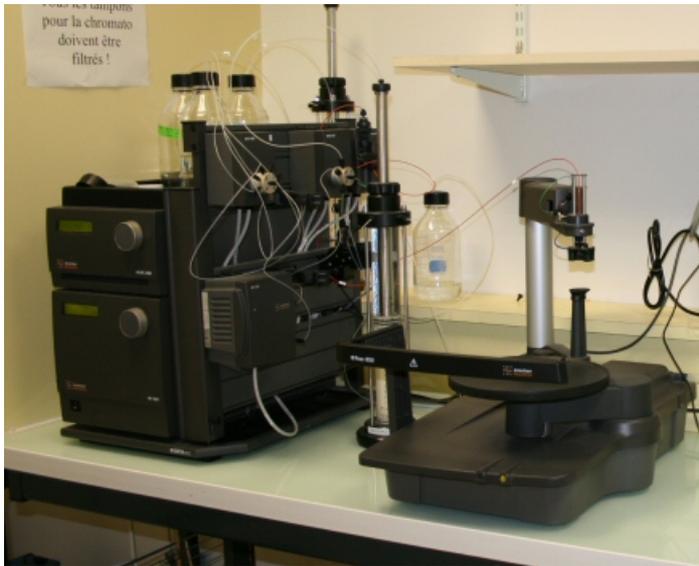


Figure 57 : Système ÄktaFPLC™ (Amersham Biosciences) utilisé pour la purification des protéines par chromatographie.

HisTrap Mre11 Rad50 : UV
 HisTrap Mre11 Rad50 : [imidazole]
 HisTrap Mre11 Rad50 : Injection

HisTrap Mre11 Rad50 : Conductance (en fonction de NaCl)
 HisTrap Mre11 Rad50 : Fractions

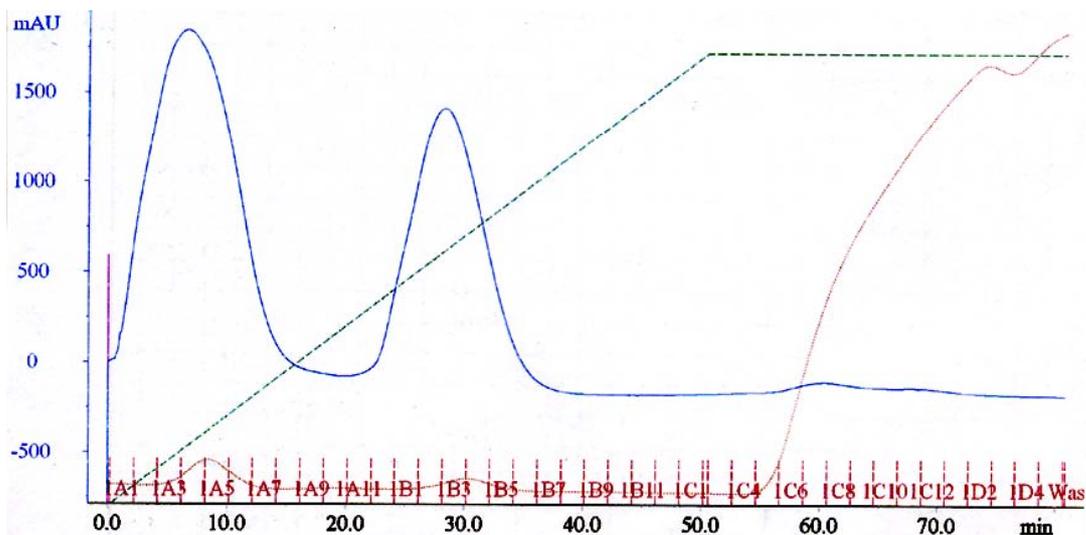


Figure 58 : chromatogramme de la purification de *PabMre11* et *PabRad50* par FPLC.

3.2.3. Dialyse

Les fractions contenant la protéine d'intérêt ont été regroupées puis dialysées de façon séquentielle contre différents tampons afin de se débarrasser du nickel éventuellement présent qui pourrait précipiter ainsi que de l'imidazole.

La dialyse permet de modifier le solvant dans lequel sont dissoutes les protéines. Les petites molécules vont traverser une membrane semi-perméable par effet d'osmose et d'agitation moléculaire, tandis que les macromolécules resteront dans le boudin de dialyse.

Les boudins de dialyse ont subi préalablement un traitement afin d'éliminer les protéases, de la façon suivante : les boudins ont été placés dans une solution contenant 5% de Na_2CO_3 à ébullition pendant 15 minutes puis ont été rincés à l'eau distillée. Ils ont ensuite été placés dans une solution contenant 1% d'EDTA pH7 mises à ébullition pendant 15 minutes. Après un dernier rinçage à l'eau distillée, les boudins de dialyse peuvent être conservés à 4°C dans une solution contenant 0,03% d'EDTA.

Une première dialyse, pendant 2 fois deux heures dans 2 litres de tampon B a permis de se débarrasser du nickel. Une nouvelle dialyse dans deux fois deux litres d'un tampon contenant Tris-HCl pH 8,0 20 mM ; imidazole 500 mM ; DTT 1 mM ; NaCl 150 mM puis deux autres dans du tampon dépourvu d'imidazole (Tris-HCl pH 8,0 20 mM ; DTT 1 mM ; NaCl 150 mM) ont permis de diminuer progressivement la concentration en imidazole du tampon. Une dernière dialyse dans ce même tampon complété de 50% de glycérol nous a permis de concentrer les protéines purifiées mais également de les conserver à -80°C à long terme. Une migration sur gel SDS-PAGE nous a permis de vérifier l'intégrité des protéines en fin de purification et un dosage colorimétrique nous a permis d'évaluer la quantité de protéines purifiées.

3.3. Dosage des protéines

Tous les dosages de protéines ont été réalisés par des méthodes colorimétriques, soit avec le kit DC Protein Assay (Bio-Rad), soit avec le kit Quick Start™ Bradford Protein Assay (Bio-Rad) selon les recommandations du fournisseur, notamment en fonction de la composition du tampon contenant les protéines à doser.

Avec le kit DC Protein Assay, basée sur la méthode de Lowry, le dosage a été réalisé avec 25 µL d'échantillon, la DO étant mesurée à 750 nm après réaction. Une gamme d'étalonnage a été réalisée avec des quantités connues de BSA (*Bovin Serum Albumin*). Si le tampon contenant les protéines contenait du glycérol, le dosage était effectué avec le kit Quick Start™ Bradford Protein Assay avec 20 µL d'échantillon, la DO étant mesurée à 595 nm après réaction (figure 59).

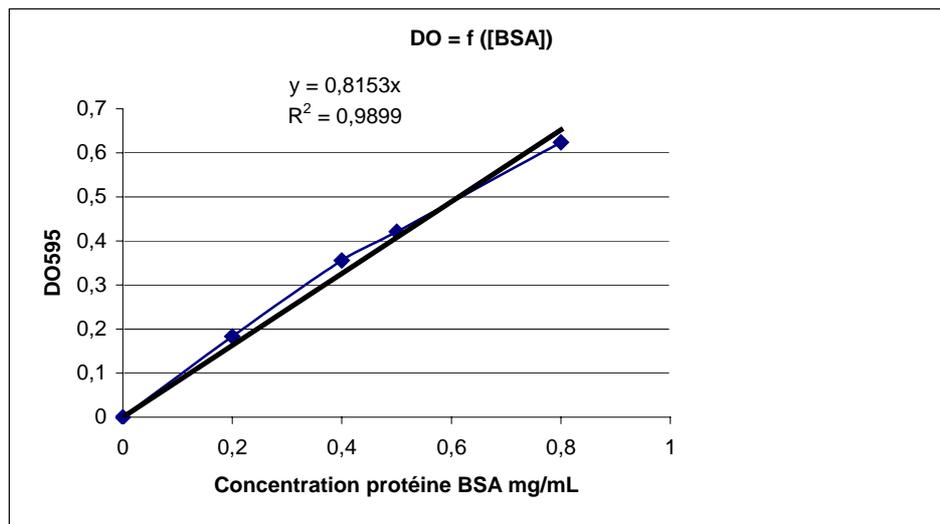


Figure 59 : courbe d'étalonnage pour le dosage des protéines par le kit Quick Start™ Bradford Protein Assay (Bio-Rad).

3.4. Séparation et visualisation des protéines par électrophorèse sur gel de polyacrylamide

L'électrophorèse est basée sur le déplacement de molécules électriquement chargées dans un champ électrique. Les protéines peuvent être ainsi séparées en fonction de leur masse, par migration sur un gel de polyacrylamide en présence de SDS (*Sodium Dodecyl Sulfate*) dans un champ électrique (SDS-PAGE : *Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis*) (Laemmli, 1970). Le SDS est un détergent qui se lie aux protéines en les transformant en polyanions de charge globale négative et de même mobilité électrophorétique. Les protéines sont ainsi séparées uniquement en fonction de leur masse moléculaire par effet de gel-filtration.

3.4.1. Préparation des échantillons

Le tampon de chargement de Laemmli a été ajouté à 10 µL d'extrait protéique, puis les protéines ont été dénaturées par un chauffage à 98°C pendant 10 minutes. Un marqueur de poids moléculaire (*Precision Plus Protein All Blue Standards 10-250 kDa* Bio-rad) a été déposé sur le gel afin de déterminer la taille des protéines d'intérêts (figure 60).

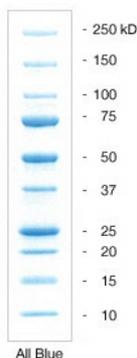


Figure 60 : marqueur de taille *Precision Plus Protein™ Standards* (Bio-Rad).

3.4.2. Migration

Les électrophorèses ont été réalisées dans des gels de polyacrylamide 12%, prêts à l'emploi, Bis-Tris *Criterion™ XT Precast Gel* commercialisés par la société Bio-Rad. Ces gels ont été placés dans une cuve Bio-Rad contenant du tampon MOPS-SDS. La migration a eu lieu à voltage constant, à 200 Volts, pendant environ 1 heure, jusqu'à ce que le front de migration ait atteint le bas du gel.

Tampon MOPS-SDS 20X

Composants	Concentration	Quantité pour 1 L
MOPS	1 M	209,2 g
Tris-base	0,5 M	121,2 g
SDS	70 mM	20 g
EDTA	16 mM	6 g

3.4.3. Coloration et séchage des gels

Les gels ont été démoulés puis placés dans une solution de bleu de Coomassie à 1g/L dans un mélange contenant 10% d'acide acétique, 50% d'éthanol et 40% d'eau, pendant 1 heure, sous agitation constante. Les gels ont ensuite été décolorés par des bains successifs dans une solution contenant 10% d'acide acétique, 30% d'éthanol et 60% d'eau, jusqu'à visualisation des bandes correspondantes aux protéines. Les gels peuvent être ensuite conservés quelques jours dans l'eau ou séchés directement, par chauffage sous vide, avec le sécheur de gel *Gel Dryeur 583* (Bio-Rad).

3.5. Technique de Western Blot

La technique du Western Blot permet d'identifier de façon spécifique une protéine d'intérêt par immuno-marquage avec l'aide d'un anticorps dirigé contre cette protéine, après électrophorèse sur gel de polyacrylamide, et transfert des protéines sur un support immobilisant (membrane de nitrocellulose ou membrane PolyVinylidene DiFluoride PVDF).

3.5.1. Préparation des anticorps

La production des anticorps utilisés au cours de cette étude a été effectuée par la société Eurogentec. Les lapins ont été immunisés par injection de 200 µg de protéines recombinantes, puis le sérum prélevé au bout de 87 jours contient les anticorps polyclonaux dirigés contre la protéine d'intérêt.

Les anticorps ont été purifiés sur colonne sépharose protéine A (HiTrap™ Protein A HP Amersham Biosciences). 10 mL de sérum ont été dilués dans 10 mL de tampon de fixation (NaH₂PO₄ 20 mM, pH7) et centrifugé à 10 000 g pendant 10 minutes afin de culotter les éléments non protéiques insolubles. Après une étape de lavage avec 50 mL, l'échantillon a

été déposé sur la colonne de tampon de fixation, les anticorps sont élués avec 25 mL de tampon d'éluion (Acide citrique 0,1 M pH 3,0) et récolté dans 7 mL de solution de neutralisation (Tris-HCl 1 M pH 9,0). Les solutions d'anticorps ont ensuite été aliquotées et conservées à -20°C . La vérification de la purification a été observée après migration en SDS-PAGE.

3.5.2. Electro-transfert des protéines sur membrane

Après migration des protéines sur gel de polyacrylamide SDS-PAGE, celles-ci ont été transférées sous l'influence d'un champ électrique sur une membrane de PVDF. Le gel, situé du côté de l'anode et la membrane située du côté de la cathode, sont pris en sandwich entre des papiers Whatman imbibés de tampon de transfert et les électrodes. Le transfert a lieu dans une cuve (*Criterion Blotter*, Bio-Rad) contenant du tampon de transfert (Tris base 25 mM ; glycine 192 mM ; méthanol 20%), sous voltage constant 100 V pendant 45 minutes.

Tampon de transfert

Composants	Concentration	Quantité pour 1 L
Tris-base	25 mM	3,03 g
Glycine	192 mM	14,4 g
Méthanol	20 %	200 mL

3.5.3. Immunodétection des protéines

Avant de sonder les protéines avec l'anticorps spécifique, tous les sites de fixations potentiels de la membrane doivent être bloqués. La membrane a ainsi été saturée avec des protéines du lait en l'incubant pendant au moins 30 minutes dans un tampon TBST (Tris-HCl pH 7,5 10 mM ; NaCl 150 mM ; 0,05% Tween 20) contenant 5% de lait en poudre. La membrane a ensuite été placée dans une solution de TBST contenant l'anticorps primaire pendant 1 heure à température ambiante ou sur la nuit à 4°C . Après 3 étapes de lavages dans des bains successifs de TBST de 10 minutes, et afin de détecter le complexe antigène-anticorps formé, la membrane a été incubé dans du tampon TBST contenant l'anticorps secondaire dirigé contre l'anticorps primaire, pendant 1 heure à température ambiante. L'anticorps secondaire est couplé à une enzyme HRP (*Horse Radish Peroxydase*) qui permet sa détection.

Les dilutions pour chaque anticorps sont référencées dans le tableau 7.

La révélation des protéines a été rendue possible grâce au kit ECL Plus (Amersham Biosciences). Le principe de ce kit est basé sur la chimiluminescence amplifiée (*Enhanced Chemiluminescence*). Le substrat Lumigen PS-3 (luminol) apporté par la solution de détection est oxydé en ester d'acridium sous l'action de peroxydes et de l'enzyme peroxydase. Les esters formés réagissent avec le peroxyde d'hydrogène H_2O_2 et conduisent à des produits

excités, qui en redescendant à l'état fondamental, émettent une fluorescence détectée par un phospho-fluo-imageur. La membrane a ainsi été incubé en présence de 2,5 % de luminol et 97,5% d'H₂O₂ pendant 2 minutes, puis la révélation a été acquise sous forme d'image numérisée grâce au phospho-fluo-imager Typhoon 9400 (Amersham Biosciences).

Anticorps	Dilution
Anti-PCNA	1/10 000
Anti Fen1	1/5 000
Anti p41	1/2 500
Anti His-Tag	1/2 000
Anti <i>rabbit HRP</i> (Amersham)	1/20 000
Anti <i>mouse HRP</i> (GE HealthcareScience)	1/5 000

Tableau 7 : conditions d'utilisation des anticorps

3.6. Préparation de l'extrait cellulaire total de *Pyrococcus abyssi*

Durant le processus de culture continu en bioréacteur gas-lift, la suspension cellulaire en phase exponentielle de croissance ($2 \cdot 10^8$ cellules par mL) est collectée et centrifugée à 5000 g pendant 45 minutes. Le culot est resuspendu dans du tampon Tris-HCl 50 mM, pH 8,0 contenant du DTT 0,5 mM, NaCl 150 mM. Les cellules sont lysées par sonication, par l'intermédiaire d'ultrason, grâce au sonicateur vibracell (15 pulses de 30 secondes, 375 W, 40% amplitude). La lyse cellulaire peut être contrôlée en microscopie optique. Les débris cellulaires sont enlevés par centrifugation à 10 000 g pendant 20 minutes. Le surnageant est utilisée comme extrait cellulaire total. La concentration en protéine totale de cet extrait peut être évaluée grâce à un dosage colorimétrique de type Bradford (Bio-Rad DC protein Assay) avec la BSA comme standard.

3.7. Capture de partenaires par *pull-down*

Le principe de la technique présentée ici est basé sur la capture, par affinité, de partenaires interactants à une protéine appât fixée sur des billes magnétiques. Cette technologie est dérivée des techniques de chromatographie IMAC (*Immobilized Metal Affinity Chromatography*) (figure 61). 20 µg de protéines recombinantes présentant une étiquette histidine ont été fixées sur 40 µL de billes magnétiques sensibilisées au cobalt (Dynabeads® Talon™, Invitrogen). Le couplage était réalisé dans 1 mL de tampon *binding and washing buffer* (Na-phosphate 50 mM pH 8,0 ; NaCl 300 mM ; 0,01% tween-20) durant 20 minutes à 25°C. Pour procéder au lavage, les billes magnétiques ont été séparées du tampon en utilisant un aimant, et le tampon a été enlevé à la pipette. Les billes sont ensuite resuspendues dans 1 mL de tampon *binding and washing buffer*. Cette procédure a été effectuée trois fois de suite,

puis 4 mL d'extrait cellulaire total de *P. abyssi* (contenant environ 10 mg/mL de protéines) ont été ajoutés aux billes, ce mélange ayant été incubé pendant deux heures à 4°C sous rotation. Les complexes appât-protéines capturées se sont formés lors de cette étape. Après 5 lavages dans du tampon *binding and washing buffer*, les billes ont été resuspendues dans 20 µL de ce même tampon, auquel ont été ajoutés 10 µL de tampon de dénaturation (tampon de chargement de Laemmli). Les protéines ont été éluées des billes par chauffage à 100°C pendant 10 minutes puis ont été séparées sur gel de polyacrylamide 12% (Bis-Tris *CriterionTM XT Precast Gel*, Bio-Rad).

En parallèle, un témoin négatif, traité dans les mêmes conditions, a été réalisé. Il correspond à l'incubation de l'extrait cellulaire avec les billes magnétiques sans appâts. Il permet de discriminer entre les interactions non spécifiques avec les billes et les partenaires spécifiques des appâts testés.

Binding and washing buffer

Composants	Concentration	Quantité pour 1 L
Na-Phosphate	50 mM	500 mL NaP 0,1 M pH 8,0
NaCl	300 mM	17,5 g
Tween 20	0,01%	100 µL

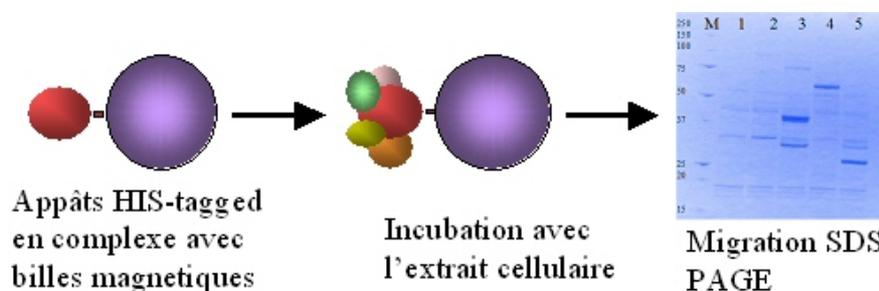


Figure 61 : principe de la technique de *pull down* à l'aide des billes magnétiques.

3.8. Capture de partenaires par Co-immunoprécipitation

La co-immunoprécipitation (Co-IP) est une technique dérivée de l'immunoprécipitation. Le principe de l'immunoprécipitation est simple. Un anticorps dirigé contre une protéine spécifique forme un complexe immun avec la protéine cible dans un échantillon ou un extrait cellulaire. Ce complexe immun est capturé par une protéine A ou une protéine G immobilisé sur un support. Après des étapes de lavage, les complexes immuns sont élués du support et analysés après électrophorèse SDS-PAGE et western-blot. La co-immunoprécipitation est conduite de la même manière, mais dans ce cas une protéine partenaire de la protéine cible est co-précipité avec le complexe immun (figure 62).

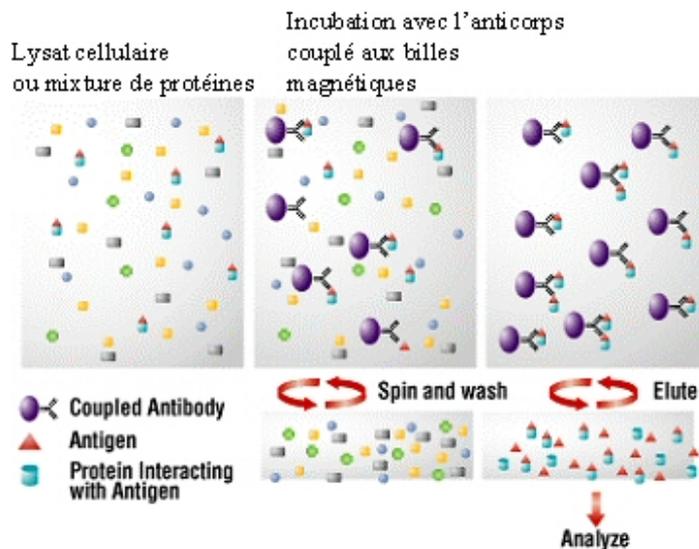


Figure 62 : principe schématisé de la co-immunoprécipitation de deux protéines partenaires.

Des billes magnétiques recouvertes de la protéine A (Dynabeads protein A, Novagen) ont été équilibrés avec du tampon PBS (NaCl 140 mM ; KCl 2,7 mM, Na₂HPO₄ 4 mM ; K₂HPO₄ 1,5 mM) et incubés avec des anticorps purifiés avec un rapport de 2,5 µg d'anticorps pour 1 µL de billes, à température ambiante, pendant 20 minutes, sous rotation. Après trois étapes de lavage avec le tampon PBS, les billes ont été suspendues dans 10 µL de tampon PBS et 10 µL de tampon de Laemmli. L'élution des complexes immuns a eu lieu lors de l'étape de chauffage à 98°C pendant 10 minutes, puis les protéines ont été séparées sur un gel de polyacrylamide-SDS et transférés sur une membrane PVDF par western blot.

3.9. Identification des protéines en spectrométrie de masse

Les bandes correspondantes aux protéines à identifier ont été découpées au scalpel à partir du gel de polyacrylamide après coloration au bleu de Coomassie et placées dans une solution contenant 0,1% d'acide acétique. Ces échantillons ont ensuite été envoyés à la plateforme de protéomique Innova Proteomics Ouest-genopole®. Les protéines ont été digérées par un traitement à la trypsine, puis le mélange de peptides ainsi formé a été placé sur une cible MALDI pour analyse en spectrométrie de masse par MALDI-TOF (*Matrix Assisted Laser Desorption Ionisation - Time of Flight*) (figure 63). Le principe de cette technique est la suivante : les peptides sont mélangés à une matrice qui est capable d'absorber à une longueur d'onde du laser utilisé. Le laser a pour effet d'évaporer la matrice qui entraîne avec elle l'échantillon en phase gazeuse. Les molécules ionisées de la matrice transfèrent un proton aux peptides qui deviennent eux même ionisés. Les ions ainsi formés sont accélérés grâce à une haute tension et envoyés dans un tube de vol où ils volent jusqu'au détecteur. Puis, ils sont triés en fonction de leur rapport masse (M) / charge (z). La technique du MALDI-TOF prend en compte la différence de temps (temps de vol) que mettent des ions pour parcourir le chemin de la chambre d'ionisation au détecteur, sachant que les plus gros ions mettent plus de

temps que les plus petits. Le détecteur collecte les ions et l'enregistreur converti ces informations en spectre de masse. Le MALDI-TOF mesure donc les masses des peptides et le logiciel ProteinScape (Bruker Daltonics) permet de réaliser une empreinte peptidique expérimentale pour chaque protéine à identifier. Cette empreinte peptidique (carte peptidique massique) est comparée aux empreintes peptidiques théoriques des protéines (digestion virtuelle par la trypsine) connues et référencées dans les bases de données internationales via des moteurs de recherche (Mascot, ProFound) (figure 64).

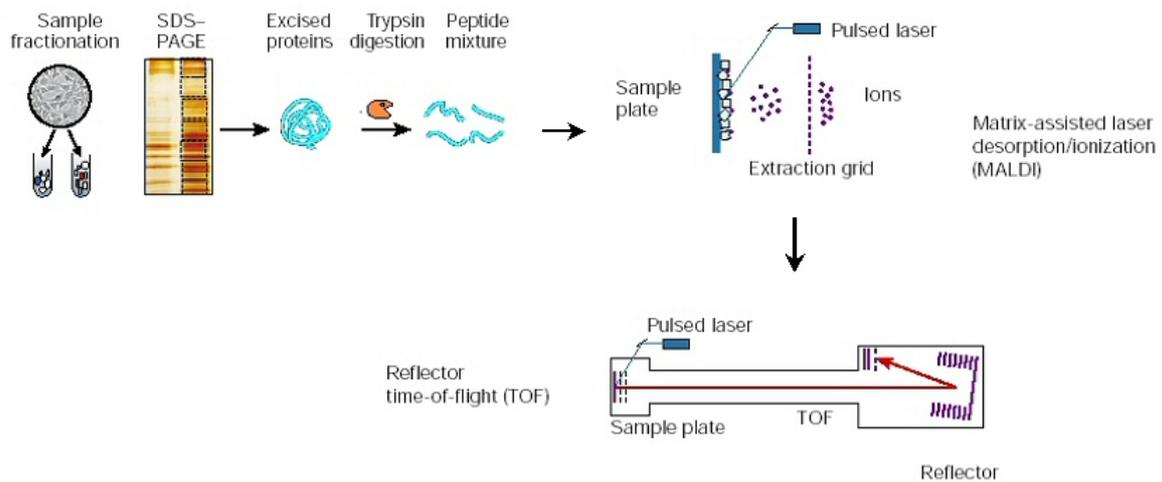


Figure 63 : principe schématisé de l'identification des protéines par spectrométrie de masse MALDI-TOF. Les protéines sont digérées en peptides qui sont ionisés et séparés dans un tube, en fonction de leur temps de vol. D'après (Aebersold and Mann, 2003).

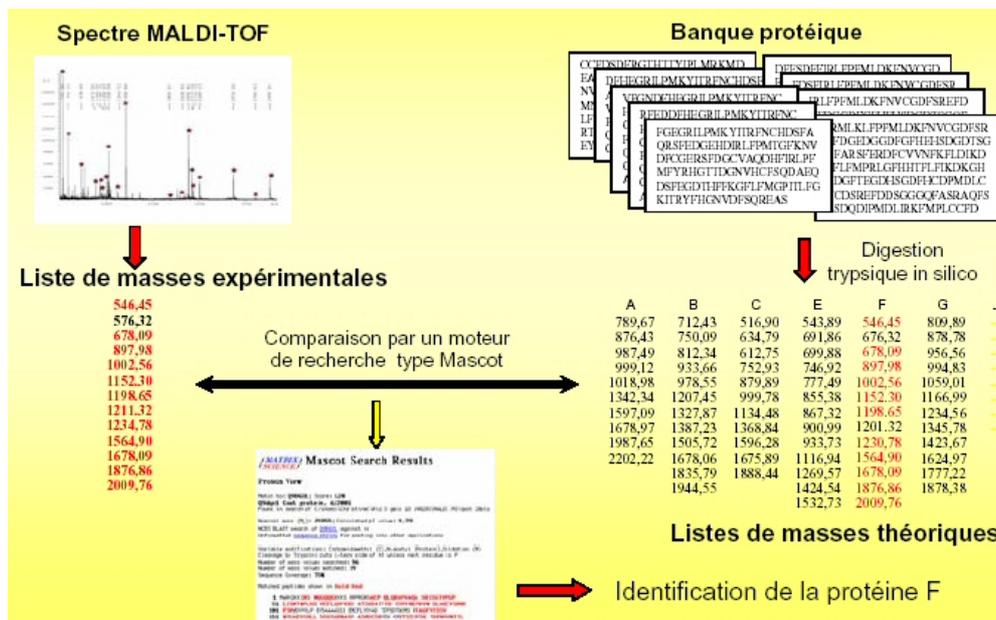


Figure 64 : stratégie d'identification des protéines par empreinte peptidique massique. La liste de masses expérimentales des peptides est comparée à la liste de masses théoriques des protéines des banques de données afin d'identifier la protéine digérée.

3.10. Analyse du protéome partiel de *P. abyss*

Afin de procéder à l'identification de l'ensemble des protéines associées à l'ADN, une extraction de ces protéines a été entreprise. L'extrait cellulaire a donc été fractionné afin de ne récupérer que l'ADN et les protéines associées. Un témoin négatif traité à la DNase puis analysé dans les mêmes conditions permettra de discriminer les protéines réellement associées à l'ADN lors d'une analyse comparative en protéomique grâce à un gel 2D.

Les culots cellulaires ont été lavés dans du tampon (Hepes 25 mM, pH7,0 ; MgCl₂ 15 mM ; sorbitol 400 mM ; DTT 0,5 mM). La suspension a ensuite été centrifugée à basse vitesse (500 rpm) pendant 2 minutes. Le surnageant contenant les cellules a ensuite été centrifugé 15 minutes à 3000 g. Elles ont ensuite été reprises dans 2 mL de tampon de lyse (Hepes 25 mM, pH7,0 ; MgCl₂ 15 mM, NaCl 100 mM, sorbitol 400 mM, DTT 0,5 mM) auxquels ont été ajoutés 2 mL du même tampon additionné de triton x100 1%. Le tout a été incubé 15 min sur glace. Un lot témoin a été traité à la DNase : 100 µL de CaCl₂ 1 M et 100 µL de DNase ont été ajoutés à ce lot qui a ensuite été incubé à 37°C pendant 1 heure. Les deux lots (avec ou sans traitement à la DNase) ont ensuite été centrifugés pendant 25 min à 14000 g à 4°C. Au cours de cette étape, il y a séparation des protéines solubles (dans le surnageant) et des protéines associées à l'ADN. Le culot contenant la fraction insoluble (ADN et protéines associées à la chromatine) a été repris dans 1 mL de tampon sans triton, puis centrifugé de nouveau pendant 25 min à 14000 g à 4°C. Le surnageant est soigneusement enlevé, puis les protéines associées à la chromatine ont été reprises dans 500 µL de tampon Tris HCl 50 mM, pH 8,0 ; DDT 0,5 mM. Les deux échantillons ont été envoyés à la plateforme de protéomique Ouest Genopole pour analyse comparative en électrophorèse bidimensionnelle.

3.11. Etude des interactions protéine-protéine par résonance plasmonique de surface (SPR)

3.11.1. Principe de la résonance plasmonique de surface

La technique de résonance plasmonique de surface permet la mise en évidence et le suivi en temps réel de l'interaction entre deux molécules (Smith and Corn, 2003) (Maier, 2007). Le principe de cette technologie est le suivant : lorsqu'un faisceau de lumière polarisée monochromatique illumine une interface entre deux milieux d'indice de réfraction différent, une partie de la lumière incidente est réfléchi sur l'interface et l'autre partie de la lumière est réfractée à travers la surface. Selon l'angle d'incidence du faisceau toute la lumière peut être réfléchi. Lorsqu'il n'y a pas de réfraction, une des composantes électromagnétiques de la lumière, l'onde évanescente, se propage perpendiculairement à l'interface sur une distance équivalente à sa longueur d'onde. La zone balayée par l'onde est appelée champ évanescent. Si

une couche fine de métal, riche en électrons libres est déposée à l'interface, ceux-ci entrent en résonance avec les photons du faisceau incident, ce phénomène est appelé résonance plasmonique de surface (le nuage électronique du métal est appelé plasmon). Une conséquence énergétique de cette résonance est visible dans le faisceau réfléchi qui, analysé avec une barrette de diodes présente une chute d'intensité à un angle défini. Cet angle d'intensité minimum est l'angle de résonance. Il varie en fonction de l'indice de réfraction du milieu présent dans le champ évanescent (figure 65).

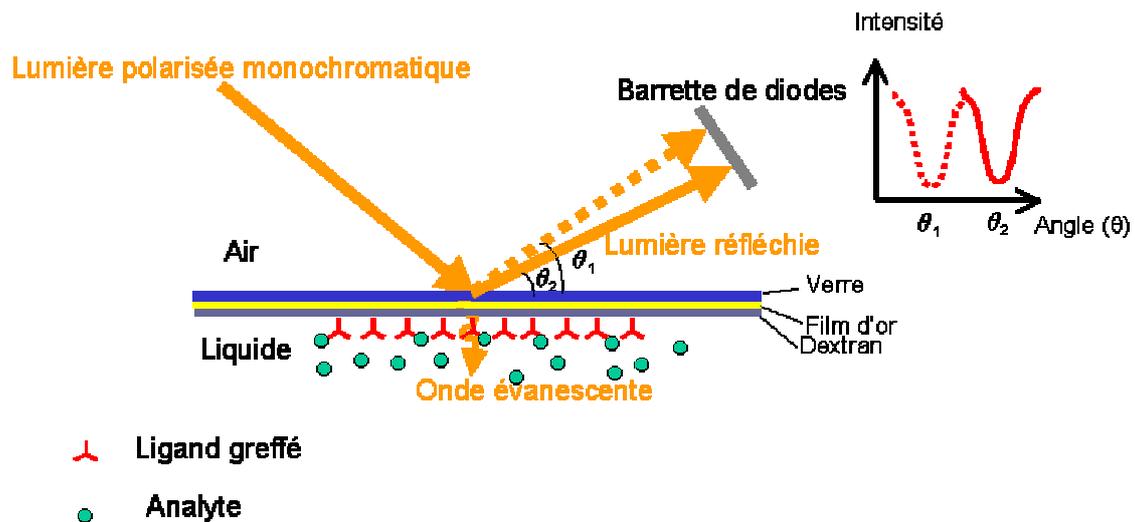


Figure 65 : représentation schématique du principe de la résonance plasmonique de surface : la lumière polarisée est réfractée selon un angle d'incidence variable en fonction de l'indice de réfraction du milieu.

3.11.2. Principe de fonctionnement de l'appareil utilisé

Le Biacore (BIA : Biospecific Interaction Analysis) a pour fonction de visualiser en temps réel des interactions entre biomolécules non marquées dans un débit continu de tampon. Un des réactifs, le ligand, est retenu de manière spécifique sur une interface appelée *sensor chip* (biocapteur). Les autres partenaires, les analytes, sont injectés à débit constant par un circuit microfluidique au contact de cette interface. L'analyte dilué dans un tampon circule à flux constant à la surface du biocapteur. Les changements de masse induits par l'association ou la dissociation des complexes modifient la réfringence du milieu et décalent la position de l'angle de résonance. L'enregistrement de la variation de l'angle de résonance permet de suivre en temps réel la fixation des molécules injectées sur le biocapteur. Le signal de résonance est exprimé en unités de résonance (RU). Une variation de 1000 RU correspond à une déviation de l'angle de 0,1 degré, ainsi qu'à une fixation de 1 ng de protéine par mm^2 de surface. L'enregistrement de ce signal s'appelle un sensorgramme (figure 66).

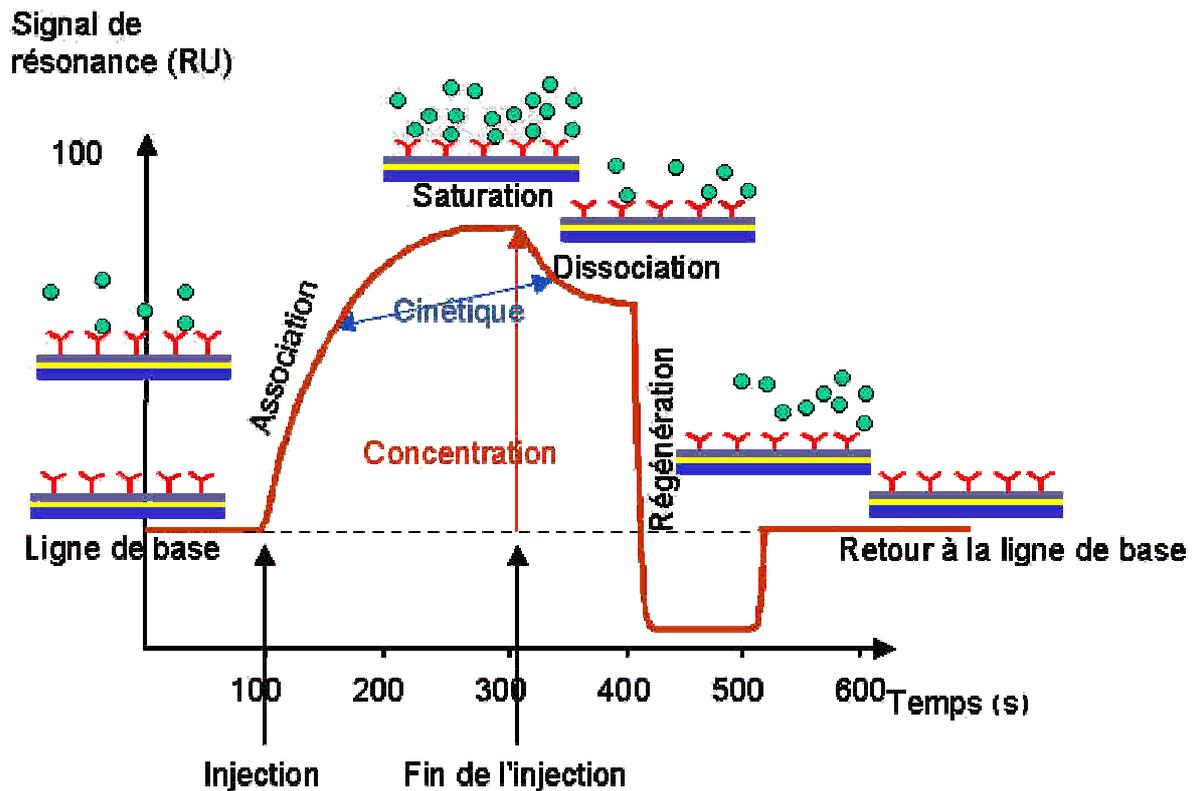


Figure 66 : sensorgramme obtenu après injection d'analytes sur un biocapteur recouvert de ligands. Il est possible d'observer après injection des analytes, une phase d'association des complexes, montrant l'interaction, puis une phase de saturation. La phase de dissociation est observée à la fin de l'injection.

3.11.3. Expériences réalisées au cours de cette étude

Les analyses étaient effectuées à 25°C avec un appareil BIAcore X. Les protéines (PCNA ou Fen1) étaient couplées de manière covalente, via une liaison amine, à la surface d'une puce CM5 (BIAcore) dans du tampon acétate de sodium à 10 mM (pH 4,5). Le taux d'immobilisation du PCNA était de 150 unité de résonance (RU). Le taux d'immobilisation pour Fen1 était de 2000 RU. Diverses concentrations des partenaires potentiels étaient injectés dans un flux de 30 µL/min dans du tampon HBS-P (Hepes 10 mM (pH 7,4) ; NaCl 150 mM ; 0,005% (v/v) P20). Les puces étaient régénérées entre les essais d'interactions successives avec du tampon NaOH 50 mM (8 µl at 10 µl/min). Les données expérimentales étaient analysées avec le logiciel BIAevaluation software (V3.2).

* *
*

3.12. Préparation des substrats

Les oligonucléotides synthétiques utilisés au cours de cette étude ont été synthétisés par la société Eurogentec, chaque nucléotide étant couplé séquentiellement de 3' vers 5' par une réaction chimique impliquant le β -cyanoethyl. Certains de ces oligonucléotides ont été couplés à un fluorochrome (FAMTM en 5' ou cy5 en 3') par cette même société. D'autres ont été marqués à la fluorescéine en 5' au laboratoire. Les oligonucléotides ont été purifiés par HPLC-RP (High Performance Liquid Chromatography - reverse phase). Après réception sous forme lyophilisée, ils sont stockés à une concentration de 100 μ M, à -20°C, dans de l'eau MilliQ ou du tampon TE.

Oligonucléotides utilisés au cours de ce travail

Oligonucléotides	taille	Séquences (5' vers 3')
L87	87-mer	CAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTACCC GGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCA
L57	57-mer	TCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATG CAAGCTTGGCA
M13 CutFAM	32-mer	TGCCAAGCTTGCATGCCTGCAGGTCGACTCTA
M13 prim	30-mer	ATTCGTAATCATGGTCATAGCTGTTTCCTG
Nick-55	55-mer	GAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATAG CTGTTTCCTG
THFnick-cy5	55-mer	THF5'GAG-GAT-CCC-CGG-GTA-CCG-AGC-TCG-AAT-TCG-TAA-TCA- TGG-TCA-TAG-CTG-TTT-CCT-G 3'CY5
5nt-cy5	50-mer	5'TCC-CCG-GGT-ACC-GAG-CTC-GAA-TTC-GTA-ATC-ATG-GTC- ATA-GCT-GTT-TCC-TG 3'CY5
THF5nt-cy5	50-mer	THF5'TCC-CCG-GGT-ACC-GAG-CTC-GAA-TTC-GTA-ATC-ATG-GTC- ATA-GCT-GTT-TCC-TG 3'CY5

3.12.1. Marquage des oligonucléotides en 5'

Le marquage d'un oligonucléotide en 5' à la fluorescéine permet la visualisation de cet oligonucléotide sans interférer avec une polymérisation de l'ADN en 3' lors des tests fonctionnels. Au cours de ce processus, un thiophosphate est transféré de l'ATP γ S au groupement 5'hydroxyl de l'oligonucléotide par une T4 polynucleotide kinase (PNK). Après addition de ce groupement thiol, la fluorescéine est couplée chimiquement à l'extrémité 5' de cet oligonucléotide. Au cours de ce travail, le mélange réactionnel contenait 1,5 μ L d'oligonucléotide à 100 pmol/ μ L ; 2 μ L d'ATP γ S (375 μ M) ; 1,4 μ L de T4PNK ; 8,6 μ L d'H₂O. Ce mélange a été incubé 30 min à 37°C, puis 30 min à 65°C après ajout de 1 μ L de fluorescéine. Afin d'obtenir une concentration finale en oligonucléotide de 5 pmol/ μ L, 15 μ L d'H₂O ont été ajoutés. Cet ADN marqué a ensuite été purifié sur des microcolonnes de résine sephadex (MicroSpin G-25 columns, GE Healthcare) afin de se débarrasser de l'excès de fluorochrome libre.

Les oligonucléotides ainsi marqués peuvent servir d'amorce pour des tests d'élongation, mais également de marqueur de taille après migration en gel de polyacrylamide.

3.12.2. Préparation de la matrice ADN amorcé

La matrice ADN utilisé au cours de cette étude était le vecteur M13mp18(+) simple brin circulaire (Amersham), dérivé du bactériophage filamenteux M13 (Yanisch-Perron *et al.*, 1985). Ce vecteur, qui comporte 7 250 bases, a été hybridé avec un oligonucléotide constituant une zone double brin pouvant servir d'amorce, et éventuellement d'autres oligonucléotides en aval de façon à créer une brèche à remplir par une enzyme présentant une activité de polymérisation de l'ADN (figure 67).

23,8 μL de vecteur M13mp18(+) à 200 ng/ μL (2 pmol), 6 pmol d'oligonucléotide, 5 μL d'*annealing buffer* 10X (Tris pH 8 100 mM ; NaCl 500 mM) ont été mélangés dans un volume réactionnel total de 50 μL (QSP H₂O 50 μL). Ce mélange réactionnel a été incubé à 75°C pendant 15 minutes puis laissé reposer à température ambiante pendant au moins 1 heure. La matrice M13 amorcé aura ainsi une concentration finale de 0,04 pmol/ μL .

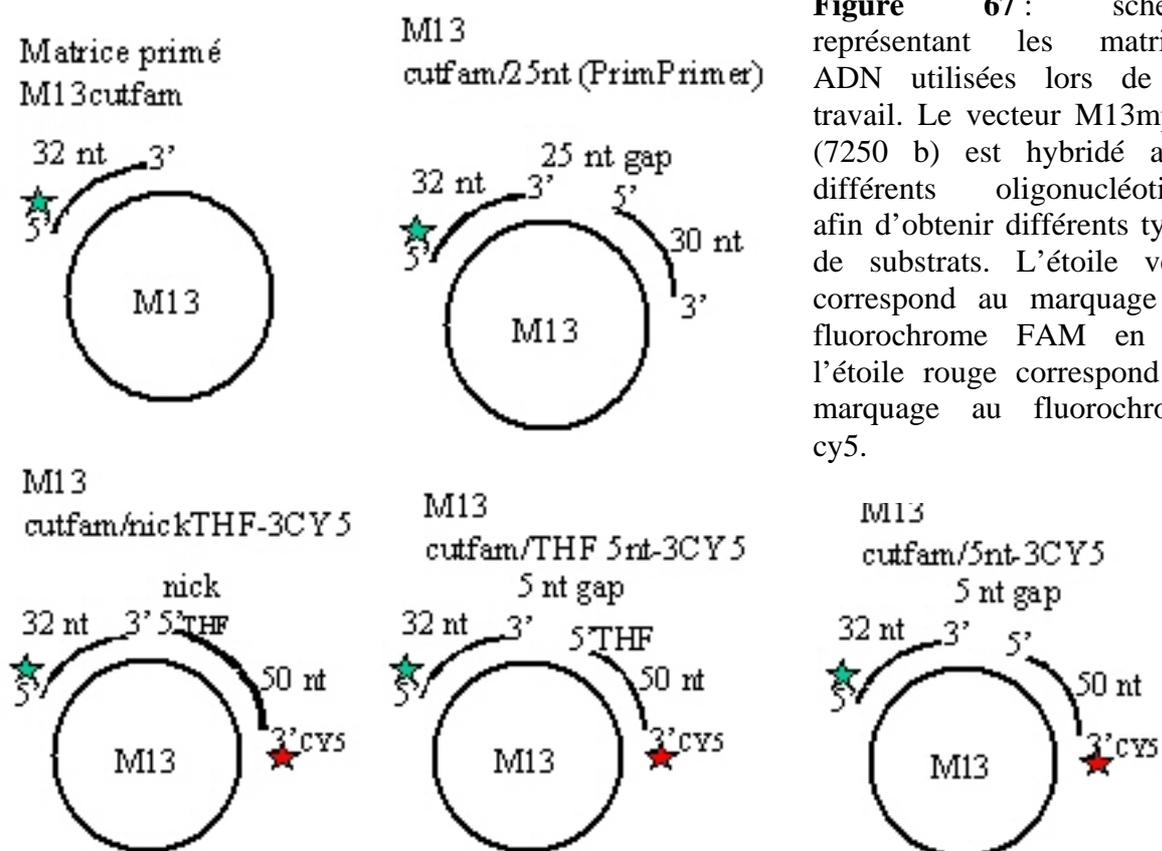


Figure 67 : schéma représentant les matrices ADN utilisées lors de ce travail. Le vecteur M13mp18 (7250 b) est hybridé avec différents oligonucléotides afin d'obtenir différents types de substrats. L'étoile verte correspond au marquage au fluorochrome FAM en 5', l'étoile rouge correspond au marquage au fluorochrome cy5.

3.13. Protocole d'élongation de la matrice amorcée

Le mélange réactionnel contenait dans un volume total de 10 μ L, 1 μ L de matrice M13 amorcé (soit 0,04 pmol), 1 μ L de tampon de réaction 10X (Tris 500 mM ;beta mercapto 10 mM ;MgCl₂ 50 mM) et typiquement 5 pmol de protéines. L'élongation était typiquement effectuée à 60°C pendant 15 à 30 minutes. La réaction était ensuite arrêtée sur glace en ajoutant 10 μ L de tampon stop (98% formamide ; EDTA 10 mM).

3.14. Révélation des produits de réaction obtenus en électrophorèse sur gel de polyacrylamide dénaturant

Les produits de réaction ont été dénaturés à 98°C pendant 10 minutes, puis déposés sur gel de polyacrylamide 14% contenant 7 M d'urée et 3,075 M de formamide (condition dénaturante). Des oligonucléotides de 87 nucléotides et 57 nucléotides, préalablement marqué à la fluorescéine sont utilisés comme marqueur de masse moléculaire.

Tampon TBE 10 X

Composants	Concentration	Quantité pour 1 L
Tris base	890 mM	108 g
Acide borique	890 mM	55 g
EDTA	20 mM	40 mL EDTA 0,5 M pH 8,0

Préparation d'un gel acrylamide 14%-urée-formamide

Composants	Quantité pour 40 mL
Formamide	5 mL
Acrylamide bis-acrylamide 40%	14 mL
TBE 10X	4 mL
Urée	16,82 g
H ₂ O	Qsp 40 mL
Ammonium persulfate (APS) 10%	400 μ L
Temed	60 μ L

Le phospho-fluo-imager Typhoon 9400 (Amersham Biosciences) permet la détection des produits et l'acquisition des images sous formes numérisées. Les fluorochromes sont excités par un laser, à une longueur d'onde donnée, en absorbant de l'énergie lumineuse sous forme d'un photon. Puis, cette énergie est restitué rapidement, par émission d'un photon, mais à une longueur d'onde plus élevée. Les différents fluorochromes sont caractérisés par leur spectre d'excitation et leur spectre d'émission (figure 68). Selon les fluorochromes utilisés, le

phospho-fluo-imager permet de sélectionner une longueur d'onde d'excitation et des filtres optiques pour la détection de la fluorescence émise. Au cours du scan, l'émission de la lumière fluorescente est convertie en signal électrique, proportionnel au nombre de photons détectés, puis en signal digital. Les résultats sont visualisables sous formes d'images numériques et analysés grâce au logiciel Imagequant (Molecular Dynamics®).

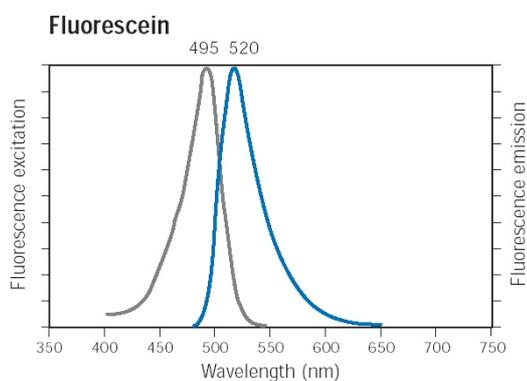


Figure 68 : spectre d'excitation (ligne grise) et spectre d'émission (ligne bleue) de la fluorescéine.

Fluorochromes	Excitation max (nm)	Emission max (nm)	Excitation (Typhoon) (nm)	Emission (Typhoon) (filtre)
Fluorescéine	495	520	532	520 BP 40
FAM TM	495	535	532	520 BP 40
Cy5	649	670	633	670 BP 30

Tableau x : conditions d'utilisation des fluorochromes utilisés au cours de cette étude.

2.15. Test d'activité de l'ADN polymérase

L'activité de polymérisation de l'ADN a été déterminée par mesure de l'incorporation de nucléotide radioactif par scintillation liquide. La technique de mesure de radioactivité par scintillation liquide consiste à mélanger la solution radioactive à étudier avec un liquide scintillant, et à transformer les rayonnements ionisants, consécutifs aux désintégrations radioactives, en lumière, détectable et quantifiable.

Les tests de polymérisation sont menés à 50°C et 60°C pendant 30 min, dans 50 µL de volume réactionnel. Le tampon de réaction contenait 1 mM dNTP ; 20 mM Tris HCl pH 8,0 ; 2 mM MgCl₂ ; 10 mM KCl ; 1 mM dithiothreitol ; 20 µM [³H]dTTP. La réaction était arrêtée sur glace, et 50 µL des produits de réaction étaient déposés sur une papier filtre Whatman. La membrane a été lavée deux fois avec du tampon contenant 100 mM NaPO₄ (pH 7) puis séché à 65°C. La quantité de radioactivité incorporé était mesuré par le compteur à scintillation.

* *

*

Références Bibliographiques

- Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**: 198-207.
- Antranikian, G., and Egorova, K. (2007) Extremophiles, a unique resource of biocatalysts for industrial biotechnology. In *Physiology and biochemistry of extremophile*. Gerday, C. and Glansdorff, N. (eds). Washington D.C.: ASM Press.
- Aravind, L., Makarova, K.S., and Koonin, E.V. (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res* **28**: 3417-3432.
- Arcus, V.L., Backbro, K., Roos, A., Daniel, E.L., and Baker, E.N. (2004) Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem* **279**: 16471-16478.
- Arezi, B., and Kuchta, R.D. (2000) Eukaryotic DNA primase. *Trends Biochem Sci* **25**: 572-576.
- Assenmacher, N., and Hopfner, K.P. (2004) MRE11/RAD50/NBS1: complex activities. *Chromosoma* **113**: 157-166.
- 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.
- Avery, O.T., Macleod, C.M., and McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. Oswald Theodore Avery (1877-1955). *Journal of Experimental Medicine* **79**: 137-158.
- 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.
- Bauer, M.W., Driskill, L.E., Callen, W., Snead, M.A., Mathur, E.J., and Kelly, R.M. (1999) An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyzes beta-1,4 bonds in mixed-linkage (1-->3),(1->4)-beta-D-glucans and cellulose. *J Bacteriol* **181**: 284-290.
- Beese, L.S., Derbyshire, V., and Steitz, T.A. (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**: 352-355.
- Belay, N., Johnson, R., Rajagopal, B.S., de Macario, E.C., and Daniels, L. (1988) Methanogenic bacteria from human dental plaque. *Appl Environ Microbiol* **54**: 600-603.
- Belay, N., Mukhopadhyay, B., Conway de Macario, E., Galask, R., and Daniels, L. (1990) Methanogenic bacteria in human vaginal samples. *J Clin Microbiol* **28**: 1666-1668.
- Bell, S.P., and Dutta, A. (2002) DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-374.
- Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A., and Forterre, P. (1997) An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* **386**: 414-417.
- Bessman, M.J., Kornberg, A., Lehman, I.R., and Simms, E.S. (1956) Enzymic synthesis of deoxyribonucleic acid. *Biochim Biophys Acta* **21**: 197-198.
- Bjornsdottir, S.H., Blondal, T., Hreggvidsson, G.O., Eggertsson, G., Petursdottir, S., Hjorleifsdottir, S., Thorbjarnardottir, S.H., and Kristjansson, J.K. (2005) *Rhodothermus marinus*: physiology and molecular biology. *Extremophiles*.
- Bloch, 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.
- Bocquier, A.A., Liu, L., Cann, I.K., Komori, K., Kohda, D., and Ishino, Y. (2001) Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr Biol* **11**: 452-456.

- Böhlke, K., Pisani, F.M., Rossi, M., and Antranikian, G. (2002) Archaeal DNA replication: spotlight on a rapidly moving field. *Extremophiles* **6**: 1-14.
- Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E., Nesbo, C.L., Case, R.J., and Doolittle, W.F. (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet* **37**: 283-328.
- Braun, K.A., Lao, Y., He, Z., Ingles, C.J., and Wold, M.S. (1997) Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry* **36**: 8443-8454.
- Bravo, R., and Celis, J.E. (1980) A search for differential polypeptide synthesis throughout the cell cycle of HeLa cells. *J Cell Biol* **84**: 795-802.
- 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.
- Brown, S.H., and Kelly, R.M. (1993) Characterization of Amylolytic Enzymes, Having Both alpha-1,4 and alpha-1,6 Hydrolytic Activity, from the Thermophilic Archaea *Pyrococcus furiosus* and *Thermococcus litoralis*. *Appl Environ Microbiol* **59**: 2614-2621.
- Budd, M.E., Tong, A.H., Polaczek, P., Peng, X., Boone, C., and Campbell, J.L. (2005) A Network of Multi-Tasking Proteins at the DNA Replication Fork Preserves Genome Stability. *PLoS Genet* **1**: e61.
- Buhler, C., Lebbink, J.H., Bocs, C., Ladenstein, R., and Forterre, P. (2001) DNA topoisomerase VI generates ATP-dependent double-strand breaks with two-nucleotide overhangs. *J Biol Chem* **276**: 37215-37222.
- Burgers, P.M., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C., Lawrence, C.W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C.A., Sugino, A., Todo, T., Wang, Z., Weill, J.C., and Woodgate, R. (2001) Eukaryotic DNA polymerases: proposal for a revised nomenclature. *J Biol Chem* **276**: 43487-43490.
- Burggraf, S., Huber, H., and Stetter, K.O. (1997) Reclassification of the crenarchaeal orders and families in accordance with 16S rRNA sequence data. *Int J Syst Bacteriol* **47**: 657-660.
- Cairns, J., and Davern, C.I. (1967) The mechanics of DNA replication in bacteria. *J Cell Physiol* **70**: Suppl:65-76.
- Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. (2005) Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes Dev* **19**: 1905-1919.
- 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., Yuasa, M., Daiyasu, H., Toh, H., and Ishino, Y. (2001a) Biochemical analysis of replication factor C from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* **183**: 2614-2623.
- Cann, I.K.O., and Ishino, Y. (1999) Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics* **152**: 1249-1267.
- Cann, I.K.O., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H., and Ishino, Y. (2001b) Biochemical analysis of Replication Factor C from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Journal of Bacteriology* **183**: 2614-2623.
- Cariello, N.F., Swenberg, J.A., and Skopek, T.R. (1991) Fidelity of *Thermococcus litoralis* DNA polymerase (Vent) in PCR determined by denaturing gradient gel electrophoresis. *Nucleic Acids Res* **19**: 4193-4198.
- Cavalier-Smith, T. (2006) Rooting the tree of life by transition analyses. *Biol Direct* **1**: 19.
- Cavicchioli, R., Curmi, P.M., Saunders, N., and Thomas, T. (2003) Pathogenic

- archaea: do they exist? *Bioessays* **25**: 1119-1128.
- Chapados, B.R., Hosfield, D.J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* **116**: 39-50.
- 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.
- Chien, A., Edgar, D.B., and Trela, J.M. (1976) Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* **127**: 1550-1557.
- Cho, S., Park, S.G., Lee, D.H., and Park, B.C. (2004) Protein-protein interaction networks: from interactions to networks. *J Biochem Mol Biol* **37**: 45-52.
- Chung, Y.C., Kobayashi, T., Kanai, H., Akiba, T., and Kudo, T. (1995) Purification and Properties of Extracellular Amylase from the Hyperthermophilic Archaeon *Thermococcus profundus* DT5432. *Appl Environ Microbiol* **61**: 1502-1506.
- Clissold, P.M., and Ponting, C.P. (2000) PIN domains in nonsense-mediated mRNA decay and RNAi. *Curr Biol* **10**: R888-890.
- Cohen, G.N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., Poch, O., Prieur, D., Querellou, J., Ripp, R., Thierry, J.C., Van, D.O.J., Weissenbach, J., Zivanovic, Y., and Forterre, P. (2003) An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Molecular Microbiology* **47**: 1495-1512.
- Com, E., Evrard, B., Roepstorff, P., Aubry, F., and Pineau, C. (2003) New insights into the rat spermatogonial proteome: identification of 156 additional proteins. *Mol Cell Proteomics* **2**: 248-261.
- Constantinesco, F., Forterre, P., and Elie, C. (2002) NurA, a novel 5'-3' nuclease gene linked to rad50 and mre11 homologs of thermophilic Archaea. *EMBO Rep* **3**: 537-542.
- Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L., and Elie, C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Res* **32**: 1439-1447.
- Corbett, K.D., Benedetti, P., and Berger, J.M. (2007) Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI. *Nat Struct Mol Biol* **14**: 611-619.
- Corlis, J.B., and Ballard, R.D. (1977) Oasis of life in the cold abyss. *National Geographic*: 441-453.
- Cubeddu, L., and White, M.F. (2005) DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol* **353**: 507-516.
- Cullmann, G., Fien, K., Kobayashi, R., and Stillman, B. (1995) Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. *Mol Cell Biol* **15**: 4661-4671.
- Daimon, K., Kawarabayasi, Y., Kikuchi, H., Sako, Y., and Ishino, Y. (2002) Three proliferating cell nuclear antigen-like proteins found in the hyperthermophilic archaeon *Aeropyrum pernix*: interactions with the two DNA polymerases. *J Bacteriol* **184**: 687-694.
- D'Amours, D., and Jackson, S.P. (2002) The Mre11 complex: at the crossroads of dna repair and checkpoint signalling. *Nat Rev Mol Cell Biol* **3**: 317-327.
- Davey, M.J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002) Motors and switches: AAA+ machines within the replisome. *Nat Rev Mol Cell Biol* **3**: 826-835.
- De Falco, M., Fusco, A., De Felice, M., Rossi, M., and Pisani, F.M. (2004) The DNA primase of *Sulfolobus solfataricus* is activated by substrates containing a thymine-rich bubble and has a 3'-terminal nucleotidyl-transferase activity. *Nucleic Acids Res* **32**: 5223-5230.
- De Falco, M., Ferrari, E., De Felice, M., Rossi, M., Hubscher, U., and Pisani, F.M. (2007) The human GINS complex binds to and specifically stimulates

- human DNA polymerase alpha-primase. *EMBO Rep* **8**: 99-103.
- 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., Esposito, L., Pucci, B., Carpentieri, F., De Falco, M., Rossi, M., and Pisani, F.M. (2003) Biochemical characterization of a CDC6-like protein from the crenarchaeon *Sulfolobus solfataricus*. *J Biol Chem* **278**: 46424-46431.
- de Laat, W.L., Appeldoorn, E., Jaspers, N.G., and Hoeijmakers, J.H. (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *J Biol Chem* **273**: 7835-7842.
- De Wachter, R., Willekens, P., and Zillig, W. (1989) Nucleotide sequence of the 5S ribosomal RNA of the archaebacterium *Pyrococcus woesei*. *Nucleic Acids Res* **17**: 5848.
- Demple, B., Herman, T., and Chen, D.S. (1991) Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: definition of a family of DNA repair enzymes. *Proc Natl Acad Sci U S A* **88**: 11450-11454.
- Deng, C., Brown, J.A., You, D., and Brown, J.M. (2005) Multiple endonucleases function to repair covalent topoisomerase I complexes in *Saccharomyces cerevisiae*. *Genetics* **170**: 591-600.
- Dervyn, E., Noirot-Gros, M.F., Mervelet, P., McGovern, S., Ehrlich, S.D., Polard, P., and Noirot, P. (2004) The bacterial condensin/cohesin-like protein complex acts in DNA repair and regulation of gene expression. *Mol Microbiol* **51**: 1629-1640.
- Desogus, G., Onesti, S., Brick, P., Rossi, M., and Pisani, F.M. (1999) Identification and characterization of a DNA primase from the hyperthermophilic archaeon *Methanococcus jannaschii*. *Nucleic Acids Res* **27**: 4444-4450.
- Di Bernardo, G., Del Gaudio, S., Cammarota, M., Galderisi, U., Cascino, A., and Cipollaro, M. (2002) Enzymatic repair of selected cross-linked homoduplex molecules enhances nuclear gene rescue from Pompeii and Herculaneum remains. *Nucleic Acids Res* **30**: e16.
- Di Giulio, M. (2003) The universal ancestor was a thermophile or a hyperthermophile: tests and further evidence. *J Theor Biol* **221**: 425-436.
- Dib, R., Chobert, J.M., Dalgalarondo, M., Barbier, G., and Haertle, T. (1998) Purification, molecular properties and specificity of a thermoactive and thermostable proteinase from *Pyrococcus abyssi*, strain st 549, hyperthermophilic archaea from deep-sea hydrothermal ecosystem. *FEBS Lett* **431**: 279-284.
- 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.
- Dionne, I., Nookala, R.K., Jackson, S.P., Doherty, A.J., and Bell, S.D. (2003) A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Molecular Cell* **11**: 275-282.
- Ditzel, L., Lowe, J., Stock, D., Stetter, K.O., Huber, H., Huber, R., and Steinbacher, S. (1998) Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell* **93**: 125-138.
- Dominguez, O., Ruiz, J.F., Lain de Lera, T., Garcia-Diaz, M., Gonzalez, M.A., Kirchhoff, T., Martinez, 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.
- Dong, G., Vieille, C., and Zeikus, J.G. (1997) Cloning, sequencing, and expression of the gene encoding amylopullulanase from *Pyrococcus furiosus* and biochemical characterization of the recombinant enzyme. *Appl Environ Microbiol* **63**: 3577-3584.
- Drewes, G., and Bouwmeester, T. (2003) Global approaches to protein-protein interactions. *Curr Opin Cell Biol* **15**: 199-205.

REFERENCES BIBLIOGRAPHIQUES

- Eckburg, P.B., Lepp, P.W., and Relman, D.A. (2003) Archaea and their potential role in human disease. *Infect Immun* **71**: 591-596.
- Edgell, D.R., and Doolittle, W.F. (1997) Archaea and the origin(s) of DNA replication proteins. *Cell* **89**: 995-998.
- Egorova, K., and Antranikian, G. (2005) Industrial relevance of thermophilic Archaea. *Curr Opin Microbiol*.
- Erauso, G., Reysenbach, A.L., Godfroy, A., Meunier, J.R., Crump, B., Partensky, F., Baross, J.A., Marteinson, V., Barbier, G., Pace, N.R., and Prieur, D. (1993) *Pyrococcus abyssi* sp. nov. , a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**: 338-349.
- Erauso, G., Godfroy, A., Raguenes, G., and Prieur, D. (1995) Plate cultivation technique for strictly anaerobic, thermophilic, sulfur-metabolizing archaea isolated from a deep-sea hydrothermal vent. In *Protocol for Archabacterial Resaerch.* Fleischmann, E., Place, A., Robb, F. and Scheir, H. (eds). New York: Cold Spring harbor Laboratory Press, pp. 25-29.
- Evguenieva-Hackenberg, E., Walter, P., Hochleitner, E., Lottspeich, F., and Klug, G. (2003) An exosome-like complex in *Sulfolobus solfataricus*. *EMBO Rep* **4**: 889-893.
- Fiala, G., and Stetter, K.O. (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Arch. Microbiol.* **145**: 56-61.
- Fitz-Gibbon, S.T., Ladner, H., Kim, U.J., Stetter, K.O., Simon, M.I., and Miller, J.H. (2002) Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *Proc Natl Acad Sci U S A* **99**: 984-989.
- Forterre, P., Elie, C., Sioud, M., and Hamal, A. (1989) Studies on DNA polymerases and topoisomerases in archaeobacteria. *Can J Microbiol* **35**: 228-233.
- Forterre, P., Bergerat, A., and Lopez-Garcia, P. (1996) The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea. *FEMS Microbiol Rev* **18**: 237-248.
- Forterre, P., de la Tour, C.B., Philippe, H., and Duguet, M. (2000) Reverse gyrase from hyperthermophiles - probable transfer of a thermoadaptation trait from Archaea to Bacteria. *Trends in Genetics* **16**: 152-154.
- Forterre, P., Zivanovic, Y., Matte-Tailliez, O., Matsunaga, F., and Myllykallio, H. (2001) Ce qui est vrai pour l'Archaea *Pyrococcus* est (presque) vrai pour l'éléphant ! *médecine/sciences* **17**: 939-945.
- Forterre, P. (2002a) The origin of DNA genomes and DNA replication proteins. *Current Opinion in Microbiology* **5**: 525-532.
- Forterre, P. (2002b) A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. *Trends Genet* **18**: 236-237.
- Forterre, P., Brochier, C., and Philippe, H. (2002) Evolution of the archaea. *Theoretical Population Biology* **61**: 409-422.
- 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.
- Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R., and Steitz, T.A. (1988) Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci U S A* **85**: 8924-8928.
- Frey, A., and Suppmann, B. (1996) Demonstration of the Expand PCR System's greater fidelity and higher yields with a lacI-based-fidelity assay. *Biochemica* **2**: 34-35.
- Friedberg, E.C. (2006) The eureka enzyme: the discovery of DNA polymerase. *Nat Rev Mol Cell Biol* **7**: 143-147.
- Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., and Ellenberger, T. (2006) Mechanism of nucleotide excision repair in eukaryotes. In *DNA Repair and Mutagenesis*. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A. and Ellenberger, T. (eds). Washington.
- Friedrich-Heineken, E., and Hubscher, U. (2004) The Fen1 extrahelical 3'-flap pocket is conserved from archaea to human and regulates DNA substrate

- specificity. *Nucleic Acids Res* **32**: 2520-2528.
- Fukuchi, S., and Nishikawa, K. (2001) Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria. *J Mol Biol* **309**: 835-843.
- Fukui, T., Yamauchi, K., Muroya, T., Akiyama, M., Maki, H., Sugino, A., and Waga, S. (2004) Distinct roles of DNA polymerases delta and epsilon at the replication fork in *Xenopus* egg extracts. *Genes Cells* **9**: 179-191.
- Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. (2005) Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res* **15**: 352-363.
- Gadelle, D., Bocs, C., Graille, M., and Forterre, P. (2005) Inhibition of archaeal growth and DNA topoisomerase VI activities by the Hsp90 inhibitor radicicol. *Nucleic Acids Res* **33**: 2310-2317.
- Galtier, N., and Lobry, J.R. (1997) Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* **44**: 632-636.
- Gambus, A., Jones, R.C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R.D., and Labib, K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8**: 358-366.
- Garcia-Diaz, M., Dominguez, O., Lopez-Fernandez, L.A., de Lera, L.T., Saniger, M.L., Ruiz, J.F., Parraga, M., Garcia-Ortiz, M.J., Kirchhoff, T., del Mazo, J., Bernad, A., and Blanco, L. (2000) DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis. *J Mol Biol* **301**: 851-867.
- Garg, P., and Burgers, P.M. (2005) DNA polymerases that propagate the eukaryotic DNA replication fork. *Crit Rev Biochem Mol Biol* **40**: 115-128.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141-147.
- 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.
- Godfroy, A., Postec, A., and Raven, N. (2006) Growth of Hyperthermophilic Microorganisms for Physiological and Nutritional Studies. In *Methods in Microbiology, Extremophiles*. Rainey, F.A. and Oren, A. (eds). Oxford, England: Elsevier/Academic Press.
- Gonzalez, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., Tamaoka, J., and Kato, C. (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2**: 123-130.
- Grogan, D.W. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol* **8**: 180-185.
- Grogan, D.W., Carver, G.T., and Drake, J.W. (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. *Proc Natl Acad Sci U S A* **98**: 7928-7933.
- Grogan, D.W. (2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol* **6**: 137-144.
- 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*. *European Journal of Biochemistry* **268**: 5961-5969.
- Harrington, J.J., and Lieber, M.R. (1994) The characterization of a mammalian DNA structure-specific endonuclease. *Embo J* **13**: 1235-1246.
- Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999) From molecular to modular cell biology. *Nature* **402**: C47-52.
- 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. *Journal of Molecular Biology* **323**: 795-810.
- Henneke, G., Friedrich-Heineken, E., and Hubscher, U. (2003) Flap endonuclease 1: a novel tumour suppresser protein. *Trends Biochem Sci* **28**: 384-390.
- Henneke, G., Flament, D., Hubscher, 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.*
- Herndl, G.J., Reinthaler, T., Teira, E., van Aken, H., Veth, C., Pernthaler, A., and Pernthaler, J. (2005) Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303-2309.
- Hickey, D.A., and Singer, G.A. (2004) Genomic and proteomic adaptations to growth at high temperature. *Genome Biol* **5**: 117.
- Hjorleifsdottir, S., Hreggvidsson, G.O., Fridjonsson, O., Aevarsson, A., and Kristjansson, J.K. (2002) Bacteriophage RM378 of a thermophilic host organism. US Patent 6,492,161.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., Vo, M., Taggart, J., Goudreault, M., Muskat, B., Alfarano, C., Dewar, D., Lin, Z., Michalickova, K., Willems, A.R., Sassi, H., Nielsen, P.A., Rasmussen, K.J., Andersen, J.R., Johansen, L.E., Hansen, L.H., Jespersen, H., Podtelejnikov, A., Nielsen, E., Crawford, J., Poulsen, V., Sorensen, B.D., Matthiesen, J., Hendrickson, R.C., Gleeson, F., Pawson, T., Moran, M.F., Durocher, D., Mann, M., Hogue, C.W., Figeys, D., and Tyers, M. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**: 180-183.
- Hoeijmakers, J.H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature* **411**: 366-374.
- Hopfner, K.P., Karcher, A., Shin, D., Fairley, C., Tainer, J.A., and Carney, J.P. (2000a) Mre11 and Rad50 from *Pyrococcus furiosus*: cloning and biochemical characterization reveal an evolutionarily conserved multiprotein machine. *J Bacteriol* **182**: 6036-6041.
- Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. (2000b) Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* **101**: 789-800.
- Hopfner, K.P., Karcher, A., Craig, L., Woo, T.T., Carney, J.P., and Tainer, J.A. (2001) Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell* **105**: 473-485.
- Hosfield, D.J., Mol, C.D., Shen, B., and Tainer, J.A. (1998) Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell* **95**: 135-146.
- 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.
- Huber, H., Hohn, M.J., Stetter, K.O., and Rachel, R. (2003) The phylum Nanoarchaeota: present knowledge

REFERENCES BIBLIOGRAPHIQUES

- and future perspectives of a unique form of life. *Res Microbiol* **154**: 165-171.
- Hubscher, U., and Seo, Y.S. (2001) Replication of the lagging strand: a concert of at least 23 polypeptides. *Mol Cells* **12**: 149-157.
- Hubscher, U., Maga, G., and Spadari, S. (2002) Eukaryotic DNA polymerases. *Annu Rev Biochem* **71**: 133-163.
- Hurst, L.D., and Merchant, A.R. (2001) High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes. *Proc Biol Sci* **268**: 493-497.
- Ishino, Y., Tsurimoto, T., Ishino, S., and Cann, I.K. (2001) Functional interactions of an archaeal sliding clamp with mammalian clamp loader and DNA polymerase delta. *Genes Cells* **6**: 699-706.
- Iwai, T., Kurosawa, N., Itoh, Y.H., Kimura, N., and Horiuchi, T. (2000) Sequence analysis of three family B DNA polymerases from the thermoacidophilic crenarchaeon *Sulfurisphaera ohwakuensis*. *DNA Res* **7**: 243-251.
- Iyer, L.M., Koonin, E.V., Leipe, D.D., and Aravind, L. (2005) Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res* **33**: 3875-3896.
- Jain, R., Rivera, M.C., and Lake, J.A. (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci U S A* **96**: 3801-3806.
- Jeong, H., Mason, S.P., Barabasi, A.L., and Oltvai, Z.N. (2001) Lethality and centrality in protein networks. *Nature* **411**: 41-42.
- Johnson, A., and O'Donnell, M. (2005) Cellular DNA replicases: components and dynamics at the replication fork. *Annu Rev Biochem* **74**: 283-315.
- Johnson, K.A., and Johnson, D.E. (1995) Methane emissions from cattle. *J Anim Sci* **73**: 2483-2492.
- Jonsson, Z.O., Hindges, R., and Hubscher, U. (1998) Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *Embo J* **17**: 2412-2425.
- Jorgensen, S., Vorgias, C.E., and Antranikian, G. (1997) Cloning, sequencing, characterization, and expression of an extracellular alpha-amylase from the hyperthermophilic archaeon *Pyrococcus furiosus* in *Escherichia coli* and *Bacillus subtilis*. *J Biol Chem* **272**: 16335-16342.
- Joyce, C.M. (2004) T4 replication: what does "processivity" really mean? *Proc Natl Acad Sci U S A* **101**: 8255-8256.
- Kahn, P. (1995) From genome to proteome: looking at a cell's proteins. *Science* **270**: 369-370.
- Kaiser, M.W., Lyamicheva, N., Ma, W., Miller, C., Neri, B., Fors, L., and Lyamichev, V.I. (1999) A comparison of eubacterial and archaeal structure-specific 5'-exonucleases. *J Biol Chem* **274**: 21387-21394.
- Kaledin, A.S., Sliusarenko, A.G., and Gorodetskii, S.I. (1980) [Isolation and properties of DNA polymerase from extreme thermophilic bacteria *Thermus aquaticus* YT-1]. *Biokhimiia* **45**: 644-651.
- Kao, H.I., and Bambara, R.A. (2003) The protein components and mechanism of eukaryotic Okazaki fragment maturation. *Crit Rev Biochem Mol Biol* **38**: 433-452.
- Karner, M.B., DeLong, E.F., and Karl, D.M. (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-510.
- Kasiviswanathan, R., Shin, J.H., and Kelman, Z. (2005) Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic Acids Res* **33**: 4940-4950.
- Kawai, G., Ue, H., Yasuda, M., Sakamoto, K., Hashizume, T., McCloskey, J.A., Miyazawa, T., and Yokoyama, S. (1991) Relation between functions and conformational characteristics of modified nucleosides found in tRNAs. *Nucleic Acids Symp Ser*: 49-50.
- Kelman, Z., and O'Donnell, M. (1994) DNA replication: enzymology and mechanisms. *Curr Opin Genet Dev* **4**: 185-195.

- Kelman, Z., and O'Donnell, M. (1995) Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps. *Nucleic Acids Res* **23**: 3613-3620.
- Kelman, Z. (1997) PCNA: structure, functions and interactions. *Oncogene* **14**: 629-640.
- Kelman, Z. (2000) DNA replication in the third domain (of life). *Curr Protein Pept Sci* **1**: 139-154.
- Kelman, Z., and Hurwitz, J. (2000) A unique organization of the protein subunits of the DNA polymerase clamp loader in the archaeon *Methanobacterium thermoautotrophicum* deltaH. *J Biol Chem* **275**: 7327-7336.
- Kelman, Z., and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*.
- Kirk, B.W., and Kuchta, R.D. (1999) Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. *Biochemistry* **38**: 7727-7736.
- Klungland, A., and Lindahl, T. (1997) Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *Embo J* **16**: 3341-3348.
- Komori, K., Sakae, S., Shinagawa, H., Morikawa, K., and Ishino, Y. (1999) A Holliday junction resolvase from *Pyrococcus furiosus*: functional similarity to *Escherichia coli* RuvC provides evidence for conserved mechanism of homologous recombination in Bacteria, Eukarya, and Archaea. *Proc Natl Acad Sci U S A* **96**: 8873-8878.
- Komori, K., and Ishino, Y. (2001) Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem* **276**: 25654-25660.
- Kong, X.P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992) Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69**: 425-437.
- Koonin, E.V. (2003) Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat Rev Microbiol* **1**: 127-136.
- Kornberg, A., Bessman, M.J., Lehman, I.R., and Simms, E.S. (1956) Enzymic synthesis of deoxyribonucleic acid. *Biochim Biophys Acta* **21**: 197-198.
- Krishna, T.S., Kong, X.P., Gary, S., Burgers, P.M., and Kuriyan, J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**: 1233-1243.
- Krishnan, L., Sad, S., Patel, G.B., and Sprott, G.D. (2001) The potent adjuvant activity of archaeosomes correlates to the recruitment and activation of macrophages and dendritic cells in vivo. *J Immunol* **166**: 1885-1893.
- Kube, J., Brokamp, C., Machielsen, R., van der Oost, J., and Markl, H. (2006) Influence of temperature on the production of an archaeal thermoactive alcohol dehydrogenase from *Pyrococcus furiosus* with recombinant *Escherichia coli*. *Extremophiles* **10**: 221-227.
- Kubota, Y., Takase, Y., Komori, Y., Hashimoto, Y., Arata, T., Kamimura, Y., Araki, H., and Takisawa, H. (2003) A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes Dev* **17**: 1141-1152.
- 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.
- Kunkel, T.A. (2004) DNA replication fidelity. *J Biol Chem* **279**: 16895-16898.
- Kurosawa, Y., Ogawa, T., Hirose, S., Okazaki, T., and Okazaki, R. (1975) Mechanism of DNA chain growth. XV. RNA-linked nascent DNA pieces in *Escherichia coli* strains assayed with spleen exonuclease. *J Mol Biol* **96**: 653-664.
- Labib, K., and Gambus, A. (2007) A key role for the GINS complex at DNA replication forks. *Trends Cell Biol*.
- Ladrat, C., Alayse-Danet, A.M., Dietrich, J., and Barbier, G. (1996) A new thermostable glucose-activated β -glucosidase from the

- hyperthermophilic marine archaeobacterium *Pyrococcus abyssi*: purification and characterization. *Journal of Marine Biotechnology* **4**: 192-199.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.
- Lange, M., Westermann, P., and Ahring, B.K. (2005) Archaea in protozoa and metazoa. *Appl Microbiol Biotechnol* **66**: 465-474.
- Langer, D., Hain, J., Thuriaux, P., and Zillig, W. (1995) Transcription in archaea: similarity to that in eucarya. *Proc Natl Acad Sci U S A* **92**: 5768-5772.
- Lao-Sirieix, S.H., and Bell, S.D. (2004) The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol* **344**: 1251-1263.
- Lao-Sirieix, S.H., Nookala, R.K., Roversi, P., Bell, S.D., and Pellegrini, L. (2005a) Structure of the heterodimeric core primase. *Nat Struct Mol Biol* **12**: 1137-1144.
- Lao-Sirieix, S.H., Pellegrini, L., and Bell, S.D. (2005b) The promiscuous primase. *Trends Genet.*
- Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R., and Gelfand, D.H. (1989) Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase gene from *Thermus aquaticus*. *J Biol Chem* **264**: 6427-6437.
- 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.*
- Leipe, D.D., Aravind, L., and Koonin, E.V. (1999) Did DNA replication evolve twice independently? *Nucleic Acids Res* **27**: 3389-3401.
- Lepp, P.W., Brinig, M.M., Ouverney, C.C., Palm, K., Armitage, G.C., and Relman, D.A. (2004) Methanogenic Archaea and human periodontal disease. *Proc Natl Acad Sci U S A* **101**: 6176-6181.
- Lieber, M.R. (1997) The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* **19**: 233-240.
- Limoli, C.L., Giedzinski, E., Bonner, W.M., and Cleaver, J.E. (2002) UV-induced replication arrest in the xeroderma pigmentosum variant leads to DNA double-strand breaks, gamma-H2AX formation, and Mre11 relocalization. *Proc Natl Acad Sci U S A* **99**: 233-238.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**: 709-715.
- Liu, L., Komori, K., Ishino, S., Bocquier, A.A., Cann, I.K., Kohda, D., and Ishino, Y. (2001) The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J Biol Chem* **276**: 45484-45490.
- Liu, Y., Kao, H.I., and Bambara, R.A. (2004) Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem* **73**: 589-615.
- Liu, Y., and West, S.C. (2004) Happy Hollidays: 40th anniversary of the Holliday junction. *Nat Rev Mol Cell Biol* **5**: 937-944.
- Loh, E., and Loeb, L.A. (2005) Mutability of DNA polymerase I: implications for the creation of mutant DNA polymerases. *DNA Repair (Amst)* **4**: 1390-1398.
- Loor, G., Zhang, S.J., Zhang, P., Toomey, N.L., and Lee, M.Y.W.T. (1997) Identification of DNA replication and cell cycle proteins that interact with PCNA. *Nucleic Acids Research* **25**: 5041-5046.
- 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.
- Maga, G., Frouin, I., Spadari, S., and Hubscher, U. (2001) Replication protein A as a "fidelity clamp" for DNA polymerase alpha. *J Biol Chem* **276**: 18235-18242.
- Maga, G., and Hubscher, U. (2003) Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J Cell Sci* **116**: 3051-3060.
- Maier, S. (2007) *Plasmonics: Fundamentals and Applications*.
- Makarova, K.S., Aravind, L., Grishin, N.V., Rogozin, I.B., and Koonin, E.V. (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res* **30**: 482-496.
- Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2003) Potential genomic determinants of hyperthermophily. *Trends Genet* **19**: 172-176.
- Manzan, A., Pfeiffer, G., Hefferin, M.L., Lang, C.E., Carney, J.P., and Hopfner, K.P. (2004) MlaA, a hexameric ATPase linked to the Mre11 complex in archaeal genomes. *EMBO Rep* **5**: 54-59.
- Marhuenda-Egea, F.C., and Bonete, M.J. (2002) Extreme halophilic enzymes in organic solvents. *Curr Opin Biotechnol* **13**: 385-389.
- Marians, K.J. (1992) Prokaryotic DNA replication. *Annu Rev Biochem* **61**: 673-719.
- Marinsek, N., Barry, E.R., Makarova, K.S., Dionne, I., Koonin, E.V., and Bell, S.D. (2006) GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*.
- Marti, T.M., Kunz, C., and Fleck, O. (2002) DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol* **191**: 28-41.
- Maser, R.S., Mirzoeva, O.K., Wells, J., Olivares, H., Williams, B.R., Zinkel, R.A., Farnham, P.J., and Petrini, J.H. (2001) Mre11 complex and DNA replication: linkage to E2F and sites of DNA synthesis. *Mol Cell Biol* **21**: 6006-6016.
- Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., and Elion, J. (1997) Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **277**: 1833-1834.
- Matsui, E., Kawasaki, S., Ishida, H., Ishikawa, K., Kosugi, Y., Kikuchi, H., Kawarabayashi, Y., and Matsui, I. (1999) Thermostable flap endonuclease from the archaeon, *Pyrococcus horikoshii*, cleaves the replication fork-like structure endo/exonucleolytically. *J Biol Chem* **274**: 18297-18309.
- Matsui, E., Nishio, M., Yokoyama, H., Harata, K., Darnis, S., and Matsui, I. (2003) Distinct domain functions regulating de novo DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*. *Biochemistry* **42**: 14968-14976.
- Matsumiya, S., Ishino, Y., and Morikawa, K. (2001) Crystal structure of an archaeal DNA sliding clamp: proliferating cell nuclear antigen from *Pyrococcus furiosus*. *Protein Sci* **10**: 17-23.
- Matsumiya, S., Ishino, S., Ishino, Y., and Morikawa, K. (2002) Physical interaction between proliferating cell nuclear antigen and replication factor C from *Pyrococcus furiosus*. *Genes Cells* **7**: 911-922.
- Matsumoto, Y. (2001) Molecular mechanism of PCNA-dependent base excision repair. *Prog Nucleic Acid Res Mol Biol* **68**: 129-138.
- Matsunaga, F., Forterre, P., Ishino, Y., and Myllykallio, H. (2001) *In vivo* interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 11152-11257.
- Matte-Tailliez, O., Zivanovic, Y., and Forterre, P. (2000) Mining archaeal proteomes for eukaryotic proteins with novel functions: the PACE case. *Trends Genet* **16**: 533-536.
- McGlynn, P. (2004) Links between DNA replication and recombination in prokaryotes. *Curr Opin Genet Dev* **14**: 107-112.
- Meselson, M., and Stahl, F.W. (1958) The replication of DNA. *Cold Spring Harb Symp Quant Biol* **23**: 9-12.

- Meslet-Cladiere, L., Norais, C., Kuhn, J., Briffotiaux, J., Sloostra, J.W., Ferrari, E., Hubscher, U., Flament, D., and Myllykallio, H. (2007) A Novel Proteomic Approach Identifies New Interaction Partners for Proliferating Cell Nuclear Antigen. *J Mol Biol*.
- Messer, W., Blaesing, F., Jakimowicz, D., Krause, M., Majka, J., Nardmann, J., Schaper, S., Seitz, H., Speck, C., Weigel, C., Wegrzyn, G., Welzeck, M., and Zakrzewska-Czerwinska, J. (2001) Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading. *Biochimie* **83**: 5-12.
- Miller, T.L., Wolin, M.J., de Macario, E.C., and Macario, A.J. (1982) Isolation of *Methanobrevibacter smithii* from human feces. *Appl Environ Microbiol* **43**: 227-232.
- Mirzoeva, O.K., and Petrini, J.H. (2001) DNA damage-dependent nuclear dynamics of the Mre11 complex. *Mol Cell Biol* **21**: 281-288.
- Mirzoeva, O.K., and Petrini, J.H. (2003) DNA replication-dependent nuclear dynamics of the Mre11 complex. *Mol Cancer Res* **1**: 207-218.
- Mitchell, D., Willerslev, E., and Hansen, A. (2005) Damage and repair of ancient DNA. *Mutat Res* **571**: 265-276.
- Miyachi, K., Fritzler, M.J., and Tan, E.M. (1978) Autoantibody to a nuclear antigen in proliferating cells. *J Immunol* **121**: 2228-2234.
- Miyata, T., Oyama, T., Mayanagi, K., Ishino, S., Ishino, Y., and Morikawa, K. (2004) The clamp-loading complex for processive DNA replication. *Nat Struct Mol Biol* **11**: 632-636.
- Moldovan, G.L., Pfander, B., and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell* **129**: 665-679.
- Motz, M., Kober, I., Girardot, C., Loeser, E., Bauer, U., Albers, M., Moeckel, G., Minch, E., Voss, H., Kilger, C., and Koegl, M. (2002) Elucidation of an archaeal replication protein network to generate enhanced PCR enzymes. *Journal of Biological Chemistry* **277**: 16179-16188.
- Mueller, P., Egorova, K., Vorgias, C.E., Boutou, E., Trauthwein, H., Verseck, S., and Antranikian, G. (2006) Cloning, overexpression, and characterization of a thermoactive nitrilase from the hyperthermophilic archaeon *Pyrococcus abyssi*. *Protein Expr Purif* **47**: 672-681.
- 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.
- Murante, R.S., Rust, L., and Bambara, R.A. (1995) Calf 5' to 3' exo/endonuclease must slide from a 5' end of the substrate to perform structure-specific cleavage. *J Biol Chem* **270**: 30377-30383.
- Mushegian, A.R., and Koonin, E.V. (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci U S A* **93**: 10268-10273.
- Myllykallio, H., Lopez, P., Lopez Garcia, 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.
- Nagasawa, K., Kitamura, K., Yasui, A., Nimura, Y., Ikeda, K., Hirai, M., Matsukage, A., and Nakanishi, M. (2000) Identification and characterization of human DNA polymerase beta 2, a DNA polymerase beta -related enzyme. *J Biol Chem* **275**: 31233-31238.
- Nakashima, H., Fukuchi, S., and Nishikawa, K. (2003) Compositional changes in RNA, DNA and proteins for bacterial adaptation to higher and lower temperatures. *J Biochem (Tokyo)* **133**: 507-513.
- Napoli, A., Valenti, A., Salerno, V., Nadal, M., Garnier, F., Rossi, M., and Ciaramella, M. (2004) Reverse gyrase recruitment to DNA after UV light irradiation in *Sulfolobus solfataricus*. *J Biol Chem* **279**: 33192-33198.
- Napoli, A., Valenti, A., Salerno, V., Nadal, M., Garnier, F., Rossi, M., and Ciaramella, M. (2005) Functional interaction of reverse gyrase with single-strand binding protein of the archaeon

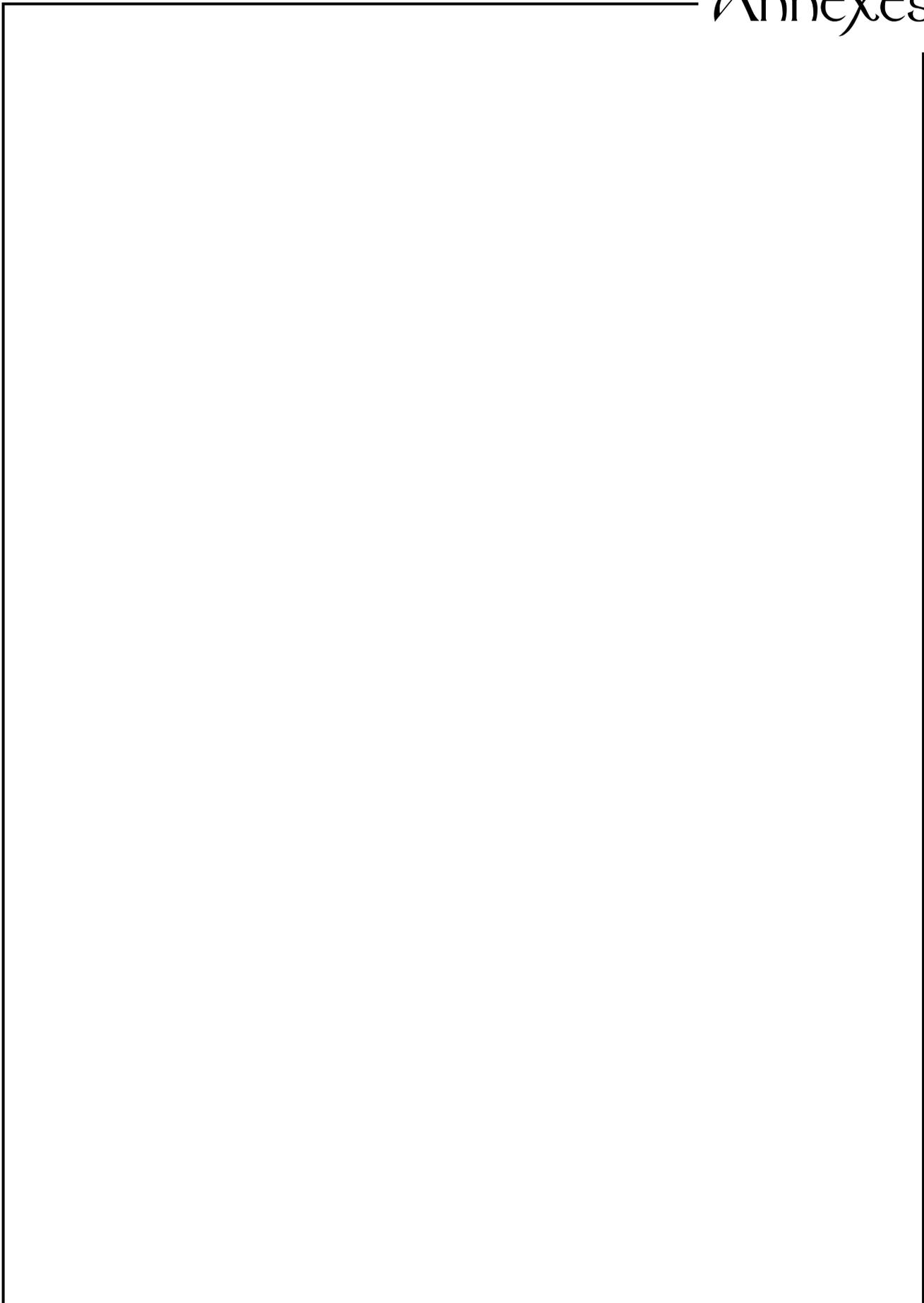
- Sulfolobus. *Nucleic Acids Res* **33**: 564-576.
- Nasheuer, H.P., von Winkler, D., Schneider, C., Dornreiter, I., Gilbert, I., and Fanning, E. (1992) Purification and functional characterization of bovine RP-A in an in vitro SV40 DNA replication system. *Chromosoma* **102**: S52-59.
- Nazarkina Zh, K., Khodyreva, S.N., Marsin, S., Radicella, J.P., and Lavrik, O.I. (2007) Study of interaction of XRCC1 with DNA and proteins of base excision repair by photoaffinity labeling technique. *Biochemistry (Mosc)* **72**: 878-886.
- Neuwald, A.F., Aravind, L., Spouge, J.L., and Koonin, E.V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**: 27-43.
- Niehaus, F., Bertoldo, C., Kahler, M., and Antranikian, G. (1999) Extremophiles as a source of novel enzymes for industrial application. *Appl Microbiol Biotechnol* **51**: 711-729.
- Niture, S.K., Doneanu, C.E., Velu, C.S., Bailey, N.I., and Srivenugopal, K.S. (2005) Proteomic analysis of human O6-methylguanine-DNA methyltransferase by affinity chromatography and tandem mass spectrometry. *Biochem Biophys Res Commun* **337**: 1176-1184.
- Noirot, P., and Noirot-Gros, M.F. (2004) Protein interaction networks in bacteria. *Curr Opin Microbiol* **7**: 505-512.
- Noirot-Gros, M.F., Dervyn, E., Wu, L.J., Mervelet, P., Errington, J., Ehrlich, S.D., and Noirot, P. (2002) An expanded view of bacterial DNA replication. *Proc Natl Acad Sci U S A* **99**: 8342-8347.
- 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**: e77.
- O'Connor, E.M., and Shand, R.F. (2002) Halocins and sulfolobocins: the emerging story of archaeal protein and peptide antibiotics. *J Ind Microbiol Biotechnol* **28**: 23-31.
- O'Donnell, M., Jeruzalmi, D., and Kuriyan, J. (2001) Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Curr Biol* **11**: R935-946.
- Ogrunc, M., Becker, D.F., Ragsdale, S.W., and Sancar, A. (1998) Nucleotide excision repair in the third kingdom. *J Bacteriol* **180**: 5796-5798.
- 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., Prakash, L., Prakash, S., Todo, T., Walker, G.C., Wang, Z., and Woodgate, R. (2001) The Y-family of DNA polymerases. *Mol Cell* **8**: 7-8.
- Ohta, S., Shiomi, Y., Sugimoto, K., Obuse, C., and Tsurimoto, T. (2002) A Proteomics Approach to Identify Proliferating Cell Nuclear Antigen (PCNA)-binding Proteins in Human Cell Lysates. Identification of the human CHL12/RFCs2-5 complex as a novel PCNA-binding protein. *Journal of Biological Chemistry* **277**: 40362-40367.
- Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G., and Steitz, T.A. (1985) Structure of large fragment of Escherichia coli DNA polymerase I complexed with dTMP. *Nature* **313**: 762-766.
- Oren, A. (1994) The ecology of the extremely halophilic archaea. *FEMS Microbiol Rev* **13**: 415-440.
- Parrish, J.Z., Yang, C., Shen, B., and Xue, D. (2003) CRN-1, a Caenorhabditis elegans FEN-1 homologue, cooperates with CPS-6/EndoG to promote apoptotic DNA degradation. *Embo J* **22**: 3451-3460.
- Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., and Yeates, T.O. (1999) Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A* **96**: 4285-4288.
- Perler, F.B., Kumar, S., and Kong, H. (1996) Thermostable DNA polymerases. *Adv Protein Chem* **48**: 377-435.
- Perry, J.J., Staley, J.T., and Lory, S. (2004) *Microbiologie*. Paris: Dunod.

- Phipps, B.M., Hoffmann, A., Stetter, K.O., and Baumeister, W. (1991) A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaeobacteria. *Embo J* **10**: 1711-1722.
- Phizicky, E.M., and Fields, S. (1995) Protein-protein interactions: methods for detection and analysis. *Microbiol Rev* **59**: 94-123.
- Pisani, F.M., De Felice, M., Carpentieri, F., and Rossi, M. (2000) Biochemical characterization of a clamp-loader complex homologous to eukaryotic replication factor C from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J Mol Biol* **301**: 61-73.
- Pomerantz, R.T., and O'Donnell, M. (2007) Replisome mechanics: insights into a twin DNA polymerase machine. *Trends Microbiol* **15**: 156-164.
- Prangishvili, D., Holz, I., Stieger, E., Nickell, S., Kristjansson, J.K., and Zillig, W. (2000) Sulfobibicins, specific proteinaceous toxins produced by strains of the extremely thermophilic archaeal genus *Sulfolobus*. *J Bacteriol* **182**: 2985-2988.
- Prasad, R., Dianov, G.L., Bohr, V.A., and Wilson, S.H. (2000) FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. *J Biol Chem* **275**: 4460-4466.
- Prosperi, E. (2006) The fellowship of the rings: distinct pools of proliferating cell nuclear antigen trimer at work. *Faseb J* **20**: 833-837.
- Ramadan, K., Shevelev, I., and Hubscher, U. (2004) The DNA-polymerase-X family: controllers of DNA quality? *Nat Rev Mol Cell Biol* **5**: 1038-1043.
- Ranalli, T.A., Tom, S., and Bambara, R.A. (2002) AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair. *J Biol Chem* **277**: 41715-41724.
- Rao, H.G., Rosenfeld, A., and Wetmur, J.G. (1998) *Methanococcus jannaschii* flap endonuclease: expression, purification, and substrate requirements. *J Bacteriol* **180**: 5406-5412.
- Reddy, T.R., and Suryanarayana, T. (1989) Archaeobacterial histone-like proteins. Purification and characterization of helix stabilizing DNA binding proteins from the acidothermophile *Sulfolobus acidocaldarius*. *J Biol Chem* **264**: 17298-17308.
- Richard, D.J., Bell, S.D., and White, M.F. (2004) Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* **32**: 1065-1074.
- Rieu-Lesme, F., Delbes, C., and Sollelis, L. (2005) Recovery of partial 16S rDNA sequences suggests the presence of Crenarchaeota in the human digestive ecosystem. *Curr Microbiol* **51**: 317-321.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**: 1030-1032.
- Robertson, C.E., Harris, J.K., Spear, J.R., and Pace, N.R. (2005) Phylogenetic diversity and ecology of environmental Archaea. *Curr Opin Microbiol*.
- Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R., and Bell, S.D. (2004) Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**: 25-38.
- Robinson, N.P., and Bell, S.D. (2005) Origins of DNA replication in the three domains of life. *Febs J* **272**: 3757-3766.
- Rodriguez, A.C., and Stock, D. (2002) Crystal structure of reverse gyrase: insights into the positive supercoiling of DNA. *Embo J* **21**: 418-426.
- Ropp, P.A., and Copeland, W.C. (1996) Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. *Genomics* **36**: 449-458.
- Rothwell, P.J., and Waksman, G. (2005) Structure and mechanism of DNA polymerases. *Adv Protein Chem* **71**: 401-440.
- 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.*
- 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.
- Sakurai, S., Kitano, K., Yamaguchi, H., Hamada, K., Okada, K., Fukuda, K., Uchida, M., Ohtsuka, E., Morioka, H., and Hakoshima, T. (2005) Structural basis for recruitment of human flap endonuclease 1 to PCNA. *Embo J* **24**: 683-693.
- Sandman, K., and Reeve, J.N. (2005) Archaeal chromatin proteins: different structures but common function? *Curr Opin Microbiol.*
- Sartori, A.A., and Jiricny, J. (2003) Enzymology of base excision repair in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J Biol Chem* **278**: 24563-24576.
- Schaeffer, P.M., Headlam, M.J., and Dixon, N.E. (2005) Protein-protein interactions in the eubacterial replisome. *IUBMB Life* **57**: 5-12.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C.P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**: 5857-5864.
- Seybert, A., Scott, D.J., Scaife, S., Singleton, M.R., and Wigley, D.B. (2002) Biochemical characterisation of the clamp/clamp loader proteins from the euryarchaeon *Archaeoglobus fulgidus*. *Nucleic Acids Res* **30**: 4329-4338.
- Sharief, F.S., Vojta, P.J., Ropp, P.A., and Copeland, W.C. (1999) Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase. *Genomics* **59**: 90-96.
- Sharples, G.J., and Leach, D.R. (1995) Structural and functional similarities between the SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) recombination and repair proteins of yeast. *Mol Microbiol* **17**: 1215-1217.
- Shen, B., Qiu, J., Hosfield, D., and Tainer, J.A. (1998) Flap endonuclease homologs in archaeobacteria exist as independent proteins. *Trends Biochem Sci* **23**: 171-173.
- Shen, B., Singh, P., Liu, R., Qiu, J., Zheng, L., Finger, L.D., and Alas, S. (2005) Multiple but dissectible functions of FEN-1 nucleases in nucleic acid processing, genome stability and diseases. *Bioessays* **27**: 717-729.
- Shockley, K.R., Ward, D.E., Chhabra, S.R., Conners, S.B., Montero, C.I., and Kelly, R.M. (2003) Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. *Appl Environ Microbiol* **69**: 2365-2371.
- Singer, G.A., and Hickey, D.A. (2003) Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. *Gene* **317**: 39-47.
- Smith, E.A., and Corn, R.M. (2003) Surface Plasmon Resonance Imaging as a Tool to Monitor Biomolecular Interactions in an Array Based Format. *Appl. Spectroscopy* **57**: 320A-332A.
- Sprinzak, E., and Margalit, H. (2001) Correlated sequence-signatures as markers of protein-protein interaction. *J Mol Biol* **311**: 681-692.
- Stetter, K.O. (1999) Extremophiles and their adaptation to hot environments. *FEBS Lett.* **452**: 22-25.
- Stillman, B. (2005) Origin recognition and the chromosome cycle. *FEBS Lett* **579**: 877-884.
- Sung, J.S., and Demple, B. (2006) Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *Febs J* **273**: 1620-1629.
- Szilagyi, A., and Zavodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8**: 493-504.
- Takagi, M., Nishioka, M., Kakihara, H., Kitabayashi, M., Inoue, H., Kawakami, B., Oka, M., and Imanaka, T. (1997) Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and its application to PCR. *Appl Environ Microbiol* **63**: 4504-4510.
- Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki,

- H. (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* **17**: 1153-1165.
- Tekaia, F., Yeramian, E., and Dujon, B. (2002) Amino acid composition of genomes, lifestyles of organisms, and evolutionary trends: a global picture with correspondence analysis. *Gene* **297**: 51-60.
- Tom, S., Henricksen, L.A., and Bambara, R.A. (2000) Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J Biol Chem* **275**: 10498-10505.
- Toueille, M., and Hubscher, U. (2004) Regulation of the DNA replication fork: a way to fight genomic instability. *Chromosoma* **113**: 113-125.
- Trakselis, M.A., and Bell, S.D. (2004) Molecular biology: the loader of the rings. *Nature* **429**: 708-709.
- Trenz, K., Smith, E., Smith, S., and Costanzo, V. (2006) ATM and ATR promote Mre11 dependent restart of collapsed replication forks and prevent accumulation of DNA breaks. *Embo J* **25**: 1764-1774.
- Trivedi, S., Gehlot, H.S., and Rao, S.R. (2006) Protein thermostability in Archaea and Eubacteria. *Genet Mol Res* **5**: 816-827.
- Uemori, T., Ishino, Y., Doi, H., and Kato, I. (1995) The hyperthermophilic archaeon *Pyrodictium occultum* has two alpha-like DNA polymerases. *J Bacteriol* **177**: 2164-2177.
- Usui, K., Katayama, S., Kanamori-Katayama, M., Ogawa, C., Kai, C., Okada, M., Kawai, J., Arakawa, T., Carninci, P., Itoh, M., Takio, K., Miyano, M., Kidoaki, S., Matsuda, T., Hayashizaki, Y., and Suzuki, H. (2005) Protein-protein interactions of the hyperthermophilic archaeon *Pyrococcus horikoshii* OT3. *Genome Biol* **6**: R98.
- van den Burg, B. (2003) Extremophiles as a source for novel enzymes. *Curr Opin Microbiol* **6**: 213-218.
- Van Houten, B., Croteau, D.L., DellaVecchia, M.J., Wang, H., and Kisker, C. (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutat Res* **577**: 92-117.
- Vetriani, C. (2001) Origin of Archaea. In *Encyclopedia of Biodiversity*. Levin, S.A. (ed).
- Vivona, J.B., and Kelman, Z. (2003) The diverse spectrum of sliding clamp interacting proteins. *FEBS Lett* **546**: 167-172.
- von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. (2002) Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**: 399-403.
- Wachtershauser, G. (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog Biophys Mol Biol* **58**: 85-201.
- Walter, P., Klein, F., Lorentzen, E., Ilchmann, A., Klug, G., and Evguenieva-Hackenberg, E. (2006) Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol* **62**: 1076-1089.
- Wang, Z., Castano, I.B., De Las Penas, A., Adams, C., and Christman, M.F. (2000) Pol kappa: A DNA polymerase required for sister chromatid cohesion. *Science* **289**: 774-779.
- Warbrick, E., Heatherington, W., Lane, D.P., and Glover, D.M. (1998) PCNA binding proteins in *Drosophila melanogaster*: the analysis of a conserved PCNA binding domain. *Nucleic Acids Res* **26**: 3925-3932.
- Warbrick, E. (2000) The puzzle of PCNA's many partners. *Bioessays* **22**: 997-1006.
- Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M., Beeson, K.Y., Bibbs, L., Bolanos, R., Keller, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G.G., Simon, M., Soll, D., Stetter, K.O., Short, J.M., and Noordewier, M. (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A* **100**: 12984-12988.
- Watson, J.D., and Crick, F.H. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**: 737-738.

- Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., Devine, S.K., Day, J.P., Wilkinson, A., d'Adda di Fagagna, F., Devine, K.M., Bowater, R.P., Jeggo, P.A., Jackson, S.P., and Doherty, A.J. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**: 1686-1689.
- White, M.F. (2003) Archaeal DNA repair: paradigms and puzzles. *Biochemical Society Transactions* **31**: 690-693.
- White, M.F. (2007) DNA Repair. In *Archaea, Evolution, Physiology, and Molecular Biology*. Garrett, R.A. and Klenk, H.P. (eds).
- Whitford, M.F., Teather, R.M., and Forster, R.J. (2001) Phylogenetic analysis of methanogens from the bovine rumen. *BMC Microbiol* **1**: 5.
- Wilson, S.H. (1998) Mammalian base excision repair and DNA polymerase beta. *Mutat Res* **407**: 203-215.
- 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 Archae, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**: 4576-4579.
- Woese, C.R. (2000) Interpreting the universal phylogenetic tree. *Proc Natl Acad Sci U S A* **97**: 8392-8396.
- Wold, M.S. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**: 61-92.
- Xu, H., Zhang, P., Liu, L., and Lee, M.Y. (2001) A novel PCNA-binding motif identified by the panning of a random peptide display library. *Biochemistry* **40**: 4512-4520.
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.
- Zavodszky, P., Kardos, J., Svingor, and Petsko, G.A. (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci U S A* **95**: 7406-7411.
- Zechner, E.L., Wu, C.A., and Marians, K.J. (1992) Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. II. Frequency of primer synthesis and efficiency of primer utilization control Okazaki fragment size. *J Biol Chem* **267**: 4045-4053.
- Zhang, R., and Zhang, C.T. (2003) Multiple replication origins of the archaeon *Halobacterium* species NRC-1. *Biochem Biophys Res Commun* **302**: 728-734.
- Zheng, L., Zhou, M., Chai, Q., Parrish, J., Xue, D., Patrick, S.M., Turchi, J.J., Yannone, S.M., Chen, D., and Shen, B. (2005) Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. *EMBO Rep* **6**: 83-89.



Milieu SME pour la culture de *P. abyssii*

	pour 1 L	pour 20 L
➤ NaCl	28 g	560 g
➤ Extrait de levure	1 g	20 g
➤ Peptone	2 g	40 g
➤ L-cystéine	0.5 g	10 g
➤ Résazurine	0.5 ml	10 ml
➤ Magnesium Salt Sol.	10 ml	200 ml
➤ Vitamines SME	0.5 ml	10 ml
➤ Solution A	1 ml	20 ml
➤ Solution B	1 ml	20 ml
➤ Solution C	1 ml	20 ml
➤ Solution D	1 ml	20 ml
➤ eau distillée	qsp 1 litre	qsp 20 litres

Il est nécessaire de réaliser des solutions "mère" :

Magnesium Salt Stock Solution

- MgSO₄, 7 H₂O 180 g
- MgCl₂, 6 H₂O 140 g
- eau distillée qsp 1 l

Solution A

- MnSO₄, 4 H₂O 9 g
- ZnSO₄, 7 H₂O 2.5 g
- NiCl₂, 6 H₂O 2.5 g
- AlK(SO₄)₂, 12 H₂O 0.3 g
- CoCl₂, 6 H₂O 0.3 g
- CuSO₄, 5 H₂O 0.15 g
- (Trisodium citrate) 4 g
- (sans effet, peut être omis)
- eau distillée qsp 1 l

Solution B

- CaCl₂, 2 H₂O 56 g
- NaBr 25 g
- KCl 16 g
- KI 10 g
- SrCl₂, 6 H₂O 4 g
- eau distillée qsp 1 l

Solution C

- K₂HPO₄ 50 g
- H₃BO₃ 7.5 g
- Na₂WO₄, 2 H₂O 3.3 g
- Na₂MoO₄, 2 H₂O 0.15 g
- Na₂SeO₃ 0.005 g
- eau distillée qsp 1 l

Solution D

- FeCl₂, 4 H₂O 10 g
- Hydrochloric acid 1M qsp 1 l

Vitamine solution pour SME

- 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
- éthanol dans eau distillée 50 % V/V qsp 1 l.

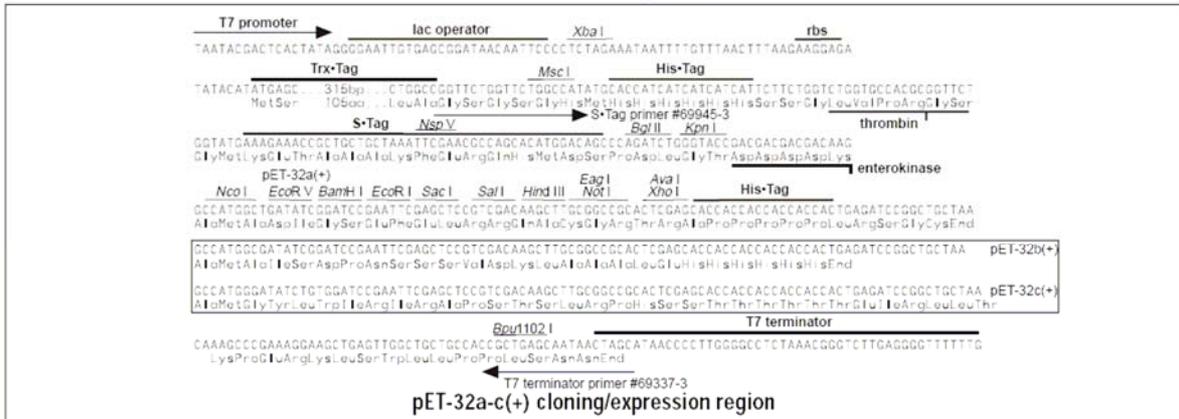
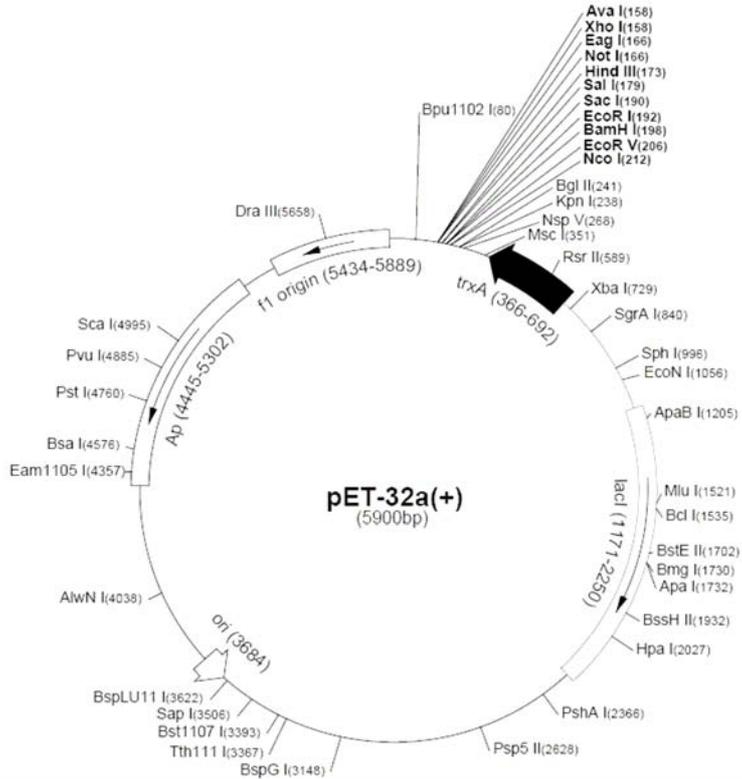
Solution de résazurine

- résazurine 1 g
- eau distillée qsp 1 l

pET-32a(+) sequence landmarks

T7 promoter	764-780
T7 transcription start	763
Trx•Tag coding sequence	366-692
His•Tag coding sequence	327-344
S•Tag coding sequence	249-293
Multiple cloning sites (<i>Nco</i> I - <i>Xho</i> I)	158-217
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lac</i> I coding sequence	1171-2250
pBR322 origin	3684
<i>bla</i> coding sequence	4445-5302
f1 origin	5434-5889

The maps for pET-32b(+) and pET-32c(+) are the same as pET-32a(+) (shown) with the following exceptions: pET-32b(+) is a 5899bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-32c(+) is a 5901bp plasmid; add 1bp to each site beyond *Bam*H I at 198 except for *Eco*R V, which cuts at 209.



Carte du vecteur restriction du vecteur pET-32 utilisé pour le clonage et l'expression de protéines fusionnées à la thioredoxine.

Résumé

Les Archaea sont des micro-organismes rencontrés dans tous les écosystèmes, mais qui apparaissent comme majoritaires dans les environnements dits extrêmes. Les archaea hyperthermophiles, comme *Pyrococcus abyssi* sont en permanence exposés à des températures qui peuvent augmenter le taux de dommages de l'ADN, pourtant, le taux de mutations spontanés chez ces micro-organismes est similaire à celui des espèces modèles mésophiles. Il est ainsi probable que les hyperthermophiles possèdent des systèmes particulièrement efficaces pour dupliquer, maintenir et stabiliser leur génome. L'objectif de ce projet était d'explorer le réseau d'interaction impliqué dans les processus de réplication et de réparation de l'ADN. L'approche méthodologique mise en œuvre a consisté à coupler la capture de partenaires d'interaction par pull-down avec leur identification par spectrométrie de masse. J'ai pu ainsi mettre en évidence, au sein l'extrait cellulaire de *P. abyssi*, un réseau préliminaire reliant des protéines de la maintenance génomique. Nous avons non seulement mis en évidence de nouvelles protéines impliquées probablement dans des mécanismes de réparation, mais également des nouvelles interactions non suspectées entre des composants déjà caractérisés.

Les principaux résultats sont les suivants : (1) La nucléase *Pab2263*, partenaire du PCNA, est un nouvel acteur du métabolisme de l'ADN. (2) Le PCNA forme un macrocomplexe avec les protéines ubiquitaires *Mre11* et *Rad50* suggérant un rôle de ce complexe dans la réparation des cassures double brin de l'ADN lors de la réplication. (3) Les protéines *Fen1* et l'ADN primase interagissent physiquement et peuvent collaborer *in vitro* pour résoudre une étape intermédiaire de la voie de réparation par excision de base. Ces résultats enrichissent notre compréhension des processus de réparation de l'ADN chez les archées.

Mots clefs : Archaea, Maintenance Génomique, Réplication de l'ADN, Interaction Protéine-Protéine

Abstract

Archaea are micro-organisms present in all ecosystems, but are majoritary present in extremophilic environments. Hyperthermophilic Archaea, as *Pyrococcus abyssi*, are permanently exposed at temperature which can increase the rate of spontaneous DNA damage. It is probable that hyperthermophilic Archaea possess efficient molecular mechanism to duplicate, maintain and stabilize their genomes. The aim of this project was to investigate the interaction network involved in the process of DNA replication and DNA repair. The methodological approach consists in the coupling of pull-down with mass spectrometry identification of interacting partners. A preliminary interaction network was identified that was composed of new proteins as well as unsuspected interactions between known components of DNA replication machinery.

The main results consist in: (1) the nuclease *Pab2263*, which interacts with PCNA, is a novel player in DNA metabolism. (2) PCNA is likely part of a macrocomplex with the ubiquitary proteins *Mre11* and *Rad50* suggesting a role of this complex in the repair of double strand break in connection with a replication fork. (3) *Fen1* and DNA primase interact physically and can collaborate *in vitro* to resolve intermediate step of the base excision repair pathway. These results enhance our understanding of DNA repair process in Archaea.

Keywords : Archaea, Genomic Maintenance, DNA Replication, Protein-Protein Interaction