



HAL
open science

Étude structurale de l'hélicase répliquative et de l'activation du primosome de *Helicobacter pylori*

Alexandre Bazin

► **To cite this version:**

Alexandre Bazin. Étude structurale de l'hélicase répliquative et de l'activation du primosome de *Helicobacter pylori*. Biochimie [q-bio.BM]. Université Claude Bernard - Lyon I, 2015. Français. NNT : 2015LYO10006 . tel-01252739

HAL Id: tel-01252739

<https://theses.hal.science/tel-01252739>

Submitted on 8 Jan 2016

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.

N° d'ordre : 06-2015

Année universitaire 2014/2015

THÈSE DE L'UNIVERSITÉ DE LYON

Pour l'obtention du
DIPLOME DE DOCTORAT
(Arrêté du 7 Août 2006)
Spécialité Biochimie

Délivré par
L'UNIVERSITÉ CLAUDE BERNARD LYON 1
L'ÉCOLE DOCTORALE INTERDISCIPLINAIRE SCIENCE-SANTÉ

Présentée par
ALEXANDRE BAZIN

ÉTUDE STRUCTURALE DE L'HÉLICASE RÉPLICATIVE ET DE L'ACTIVATION DU PRIMOSOME DE HELICOBACTER PYLORI

Soutenue publiquement le 29 janvier 2015 devant un jury composé de

Pr. Gilbert DELEAGE	Président du Jury
Dr. Hannu MYLLYKALLIO	Rapporteur
Dr. Sébastien FRIBOURG	Rapporteur
Pr. Wim BURMEISTER	Rapporteur
Dr. Joanna TIMMINS	Examinatrice
Dr. Laurent TERRADOT	Directeur de thèse

Thèse préparée au
Laboratoire de biologie structurale des complexes macromoléculaires bactériens
Institut de Biologie et Chimie des Protéines
UMR 5086 : Bases Moléculaires et Structurales des Systèmes Infectieux

UNIVERSITÉ CLAUDE BERNARD - LYON 1

Président de l'Université

Vice-président du Conseil d'Administration

Vice-président du Conseil des Etudes et de la Vie Universitaire

Vice-président du Conseil Scientifique

Directeur Général des Services

M. François-Noël GILLY

M. le Professeur Hamda BEN HADID

M. le Professeur Philippe LALLE

M. le Professeur Germain GILLET

M. Alain HELLEU

COMPOSANTES SANTÉ

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Faculté d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de la Réadaptation

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur J. ETIENNE

Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS

Directeur : Mme la Professeure C. VINCIGUERRA

Directeur : M. le Professeur Y. MATILLON

Directeur : Mme. la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire des Sciences de l'Univers de Lyon

Polytech Lyon

Ecole Supérieure de Chimie Physique Electronique

Institut Universitaire de Technologie de Lyon 1

Ecole Supérieure du Professorat et de l'Education

Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI

Directeur : M. le Professeur F. FLEURY

Directeur : Mme Caroline FELIX

Directeur : M. Hassan HAMMOURI

Directeur : M. le Professeur S. AKKOUCHE

Directeur : M. le Professeur Georges TOMANOV

Directeur : M. le Professeur H. BEN HADID

Directeur : M. Jean-Claude PLENET

Directeur : M. Y. VANPOULLE

Directeur : M. B. GUIDERDONI

Directeur : M. P. FOURNIER

Directeur : M. G. PIGNAULT

Directeur : M. le Professeur C. VITON

Directeur : M. le Professeur A. MOUGNIOTTE

Directeur : M. N. LEBOISNE

Remerciements

Le travail réalisé durant ma thèse et présenté dans ce manuscrit, n'aurait jamais pu être mené à bien sans l'accueil, l'aide et le soutien de plusieurs personnes.

Je tiens tout d'abord à remercier Gilbert Deléage qui m'a accueilli au sein de l'Institut de Biologie et Chimie des Protéines.

Je tiens ensuite à exprimer toute ma gratitude envers mon directeur de thèse Laurent Terradot. Merci de m'avoir accueilli au sein de votre équipe, de m'avoir formé à la cristallographie et supervisé durant ces trois années. Merci également de m'avoir permis de participer au congrès international de cristallographie au Canada. Je vous serai toujours reconnaissant de m'avoir fait confiance alors que vous ne me connaissiez pas car j'arrivais d'une autre université. J'espère pouvoir garder à l'avenir de très bons contacts avec vous.

Je souhaite également remercier les membres du jury Hannu Myllykallio, Sébastien Fribourg, Wim Burmeister, Joanna Timmins et Gilbert Deleage d'avoir accepté de juger ma thèse.

Cette thèse n'aurait pas pu aboutir sans la collaboration de plusieurs personnes :

Je tiens à remercier Irina Gutsche qui a accepté de me recevoir dans son équipe durant ma thèse afin de m'initier à la microscopie électronique à coloration négative.

Je remercie Joanna Timmins pour les expériences d'anisotropie de fluorescence.

Je remercie Anna Zawilak-Pawlik de m'avoir accueilli au sein de son équipe en Pologne ainsi que Rafal Donczew pour m'avoir apporté son expertise sur les tests d'ouverture de l'origine de réplication.

Je voudrais remercier également plusieurs personnes au sein de l'IBCP Virginie Gueguen-Chaignon et la plateforme de cristallogénèse ainsi que Roland Montserret pour sa disponibilité et son aide lors des expériences de titration calorimétrique.

Je remercie Anja Bockman et Carole Gardiennet pour notre collaboration sur l'étude de HpDnaB ainsi que du complexe HpDnaB•HpDnaG^{HBD} en RMN du solide. J'espère que ce travail pourra être mis en valeur.

Je souhaite remercier tous les membres des différentes lignes de lumière de l'ESRF et de SOLEIL sur lesquelles nous avons collecté des données de diffractions (ID29, ID23, ID14, Proxima 1 et Proxima 2). Je remercie plus particulièrement les membres de Proxima 1 et 2 pour leur aide lors de la résolution de la structure de HpDnaB ainsi que les membres de la ligne SWING sur laquelle nous avons collecté les données SAXS présentées dans ce manuscrit.

Je voudrais remercier tous mes camarades thésards : Julien, Jérémie, Paul, Aurore, Sylvie, Romain, tous mes collègues de l'IBCP et plus particulièrement Amandine, Arthur, Guillermo, Sylvain, Cyril.

Je remercie Dorothée Bernard pour son aide pour toutes les questions administratives auxquelles j'ai été confronté durant la thèse.

Je souhaiterais maintenant remercier toutes les personnes que j'ai côtoyées tous les jours au laboratoire pendant trois ans :

Burcu pour sa gentillesse et pour m'avoir initié au traitement des données SAXS. Je te souhaite tout le bonheur possible dans ta vie personnelle et j'espère que ton nouveau poste en Turquie t'apportera toute la satisfaction professionnelle à laquelle tu aspires.

Ali, tu as toujours été à mon écoute quel que soit le sujet. Merci pour tous les conseils que tu as pu m'apporter en cristallographie et en SAXS. Je te souhaite tout le bonheur possible avec ta femme et ton fils, j'espère que ton post-doctorat en Allemagne se déroule de la meilleure façon qui soit.

Célia, la gentillesse incarnée, merci pour toutes ces petites discussions que nous avons pu avoir et pour tous tes supers desserts que tu nous as fait manger. Je souhaite que ta thèse se déroule de la meilleure manière possible.

Thomas, tu ne le sais sans doute pas, mais tu m'as été d'un grand soutien. Je te remercie pour ta gentillesse mais surtout pour toutes les franches rigolades que l'on a pu avoir au quotidien dans le bureau. Cela m'a permis d'une certaine manière de me changer les idées. Je souhaite que tu arrives à décrocher un poste permanent, tu le mérites amplement.

Mickael, tu es le dernier arrivé au laboratoire mais ton soutien n'en est pas des moindres. Je te remercie simplement pour tout ce que tu m'as apporté au quotidien. Merci pour tous tes conseils en cristallographie, SAXS, les expériences du quotidien, les petites astuces informatiques et pour la rédaction de la thèse. J'ai une grande estime pour toi, j'ai appris énormément à ton contact, scientifiquement et humainement parlant. Je te souhaite toute la réussite possible dans ta vie personnelle avec ta famille et dans ta carrière d'enseignant chercheur.

J'espère pouvoir garder de très bons contacts avec vous tous.

Je souhaiterais remercier également Yann Hardivillier. Tu as été mon enseignant de Biochimie pendant toute ma licence. Tu as également été mon premier maître de stage et tu m'as enseigné les bases de la biologie moléculaire et de la biochimie qui m'ont été utiles pendant toute ma thèse.

Je remercie mes amis et plus particulièrement Jean-Yves qui n'a pas cessé de prendre de mes nouvelles et d'être venu me voir sur Lyon. Cela m'a permis de m'aérer l'esprit.

Je remercie ma chère et tendre Laurène pour son amour et son soutien pendant ces trois années qui n'ont pas été faciles tous les jours. Merci d'avoir supporté mon stress, de m'avoir aidé et encouragé surtout pendant la période de la rédaction. Si tu n'avais pas été présente à mes côtés, cette période de ma vie, loin de ma famille, aurait été plus difficile. Merci également à sa famille pour leur soutien.

Je remercie pour finir toute ma famille qui n'a pas pu me voir autant qu'elle le voulait pendant ces trois ans. Je remercie plus particulièrement mon frère, ma sœur et surtout mes parents. Merci maman d'avoir cru en moi. Merci papa, tu m'as toujours dit qu'aller le plus loin possible dans les études était l'une des meilleures choses à faire. Si j'ai réussi à aller jusqu'au doctorat, c'est grâce à vous deux, qui m'avez soutenu de toutes les manières possibles.

Liste des présentations scientifiques

- GTBio, Montpellier, Avril 2012:
 - Poster: Architecture of a Dodecameric Bacterial Replicative Helicase
- ARC Santé 1, Journée scientifique, l'Isle d'Abeau, Septembre 2012:
 - Poster: Architecture of a Dodecameric Bacterial Replicative Helicase
- EDISS Journée Scientifique, Lyon, Novembre 2013:
 - Oral : Structural studies of *Helicobacter pylori* DNA replication
- ARC Santé 1, Journée scientifique, l'Isle d'Abeau, Juin 2013:
 - Poster: Compréhension des mécanismes de la replication chez *Helicobacter pylori*
- GEFH, Lille, Janvier 2014 :
 - Oral: Structure of the *Helicobacter pylori* replicative helicase
- Congrès de l'union internationale de cristallographie (IUCr) 2014 à Montréal
 - Prix de poster (IUCr Mycrolitic) : Structural insights into DNA replication in *Helicobacter pylori*
- GTBio, Grenoble, Octobre 2014:
 - Oral: Structure and activation mechanism of *Helicobacter pylori* primosome

Sommaire

Liste des Figures.....	13
Liste des Tableaux.....	19
Liste des abréviations	21
1. Introduction	23
1.1. Généralités et avant-propos	23
1.2. Initiation de la réplication bactérienne.....	26
1.2.1. Origine de réplication.....	26
1.2.1.1. Origine de réplication chez les bactéries à Gram négatif.....	26
1.2.1.2. Origine de réplication des bactéries à Gram positif.....	27
1.2.1.3. Origine de réplication des eucaryotes.....	28
1.2.1.4. Origine de réplication virale, exemple du virus SV40.....	28
1.2.2. Protéines de l'initiation de la réplication AAA ⁺	29
1.2.3. Structure et mode d'action de l'initiateur de la réplication DnaA.....	30
1.2.3.1. DnaA, protéine initiatrice de la réplication	30
1.2.3.2. Mécanisme d'oligomérisation de DnaA.....	31
1.2.3.3. Régulation de l'initiation par le contrôle de DnaA	33
1.2.4. Structure et mode d'action de l'hélicase répllicative DnaB	36
1.2.4.1. Classification de DnaB dans la famille des hélicases	36
1.2.4.2. Structure générale de l'hélicase DnaB	37
1.2.4.3. Translocation de DnaB le long de l'ADN	39
1.2.4.4. Chargement de l'hélicase sur la fourche de réplication	42
1.2.4.4.1. « Ring opening » : ouverture de l'hélicase et chargement sur l'ADNsb .	43
1.2.4.4.2. « Ring making » : Formation de l'hexamère sur l'origine de réplication	49
1.2.4.4.3. Chargement de l'hélicase sans facteur de chargement	50
1.2.5. Formation du primosome et démarrage de la réplication.....	51
1.3. <i>Helicobacter pylori</i>	53
1.3.1. Maladies associées à l'infection de <i>H. pylori</i>	54
1.3.2. Prise en charge médicale	56
1.4. Réplication de l'ADN chez <i>Helicobacter pylori</i>	58
1.4.1. Une origine de réplication bi-partite.....	58
1.4.2. HobA, régulateur ou composant intégral de l'orisome ?	59
1.4.3. HpDnaB, un prototype d'hélicase dodécamérique bactérienne	60
1.5. Objectifs de la thèse	67
2. Matériels et Méthodes.....	69

2.1. Biologie Moléculaire	69
2.1.1. Souches bactériennes et plasmides	69
2.1.2. Clonage des gènes d'intérêt	70
2.1.3. Milieux et conditions de culture	71
2.1.4. Antibiotiques utilisés	72
2.2. Purification des protéines	72
2.2.1. Méthodes de purification	72
2.2.1.1. Chromatographie d'affinité	72
2.2.1.2. Chromatographie échangeuse d'ions	73
2.2.1.3. Chromatographie d'exclusion stérique	73
2.2.2. Analyse des protéines et dosage	73
2.2.3. Expressions et purifications des protéines étudiées	74
2.2.3.1. Expression et purification de 6His-HpDnaB	74
2.2.3.2. Expression et purification de HpDnaB	75
2.2.3.3. Purification de HpDnaGHBD	76
2.2.3.4. Co-expression et purification du complexe HpDnaB•HpDnaG ^{HBD}	76
2.2.3.5. Expression et purification de 6His-HpDnaB-Se	77
2.2.3.6. Purification de HpDnaA	78
2.2.3.7. Purification de HU	78
2.2.3.8. Purification de HpDnaB ^{NTD}	79
2.3. Cristallographie des protéines	80
2.3.1. Cristallisation des protéines	80
2.3.1.1. Paramètres influençant la cristallogénèse	80
2.3.1.2. Diagramme de phase	81
2.3.1.3. Méthodes de cristallogénèse utilisées	82
2.3.1.4. Cryo-protection des cristaux	83
2.3.2. Diffraction des rayons X par les cristaux de protéines	84
2.3.2.1. Définition et description du cristal	84
2.3.2.2. Systèmes cristallins et groupes d'espace	85
2.3.2.3. Les facteurs de structure	89
2.3.2.4. Enregistrement des données de diffraction	91
2.3.3. Traitement et analyse des données de diffraction	91
2.3.4. Paramètres des cristaux	92
2.3.4.1. Mosaïcité	92
2.3.4.2. Facteur B moyen	92
2.3.4.3. Estimation du nombre de molécules par unité asymétrique	93
2.3.5. Caractéristiques, qualité et validation des données	93

2.3.6. Résolution du problème de phase	94
2.3.7. Diffusion anormale	95
2.3.8. Méthode SAD (Single wavelength anomalous dispersion).....	96
2.3.9. Résolution de la structure de la protéine.....	97
2.3.10. Construction du modèle et affinement.....	97
2.4. Etudes des protéines en solution.....	98
2.4.1. Diffusion des rayons X aux petits angles (SAXS).....	98
2.4.2. Chromatographie d'exclusion de taille couplée à la diffusion de lumière multi- angles et réfractométrie (SEC-MALS).....	101
2.4.3. Titration calorimétrique isotherme (ITC)	102
2.5. Etude de l'interaction ADN-Protéine.....	103
2.5.1. Expériences d'ouverture de la fourche de réplication.....	103
2.5.2. Expériences de chargement de l'hélicase HpDnaB sur l'ADN.....	103
2.5.3. Mesure d'activité ATPase	104
2.6. Microscopie électronique à coloration négative.....	105
2.6.1. Microscopie électronique à coloration négative	106
2.6.2. Préparation de l'échantillon	107
3. Résultats	109
3.1. Préparation et purification des protéines et complexes étudiés.....	109
3.1.1. Purification de 6His-HpDnaB	109
3.1.2. Purification de HpDnaB	110
3.1.3. Purification de HpDnaG ^{HBD}	112
3.1.4. Purification du complexe HpDnaB•HpDnaG ^{HBD}	114
3.1.5. Purification de HpDnaA.....	115
3.1.6. Purification de la protéine HU.....	117
3.1.7. Purification de HpDnaB ^{NTD}	118
3.2. Structure cristallographique de HpDnaB	119
3.2.1. Cristallisation de 6His-HpDnaB.....	119
3.2.2. L'enregistrement et le traitement des données de diffraction	121
3.2.3. Tentatives de phasage par remplacement moléculaire.....	122
3.2.4. Phasage de la structure 6His-HpDnaB par single anomalous dispersion.....	123
3.2.4.1. Localisation des sites, phasage et amélioration de la carte de densité électronique.....	126
3.2.5. Construction du modèle	128
3.2.6. Validation du modèle par l'utilisation des cristaux de 6His-HpDnaB-Se.....	133
3.2.6.1. Purification et cristallisation de 6His-HpDnaB-Se.....	133

3.2.6.2. Collection et traitement des données de diffraction.....	133
3.2.6.3. Remplacement moléculaire.....	134
3.2.7. Structure générale de 6His-HpDnaB.....	135
3.2.7.1. Détails des interactions.....	137
3.2.7.1.1. Anneaux NTD et l'interface de dodécamérisation.....	137
3.2.7.1.2. Anneaux CTD et HPI.....	140
3.2.8. Tentative d'amélioration du modèle.....	143
3.2.9. Trempage des cristaux avec un oligonucléotide.....	144
3.3. Etude de HpDnaB en solution et en microscopie électronique.....	145
3.3.1. Etudes de HpDnaB en solution par SAXS.....	145
3.3.1.1. Influence de l'étiquette histidine sur la structure de HpDnaB.....	147
3.3.1.2. Effet des nucléotides sur la structure de HpDnaB.....	151
3.3.2. Microscopie électronique de HpDnaB.....	156
3.3.3. Etude de l'interaction de HpDnaB avec l'ADN.....	158
3.3.3.1. Etude par SAXS de l'effet de l'ADN sur HpDnaB.....	159
3.4. Etude biochimique de l'interaction HpDnaB•HpDnaG ^{HBD}	164
3.4.1. Etude par SAXS de l'effet de HpDnaG ^{HBD} sur HpDnaB.....	164
3.4.2. Etude par SEC-MALS de l'interaction HpDnaB/HpDnaG ^{HBD}	170
3.4.3. Mesure de l'affinité de l'interaction HpDnaB-HpDnaG ^{HBD} par titration colorimétrique isotherme (ITC).....	172
3.4.4. Mesure de l'activité hélicase de HpDnaB et du complexe HpDnaB•HpDnaG ^{HBD} et de leurs affinités pour l'ADNsb.....	174
3.4.5. Mesure de l'activité ATPase de HpDnaB et des complexes du primosome.....	176
3.4.6. Etude par microscopie électronique de HpDnaB après interaction avec HpDnaG ^{HBD}	177
3.5. Etude fonctionnelle du chargement de HpDnaB sur l'origine de réplication.....	178
3.5.1. Expérience d'ouverture de l'origine de réplication.....	178
3.5.2. Etude du chargement de l'hélicase.....	181
4. Discussion.....	185
Références Bibliographiques.....	195
Article 1.....	217
Article 2.....	231

Liste des Figures

Figure 1 : Dessin dans le sable de la théorie du réplicon	24
Figure 2 : Etapes de la réplication procaryote, de l'initiation de la réplication et de l'assemblage de la polymérase.....	26
Figure 3 : Schéma représentant l'organisation de l'origine de réplication chez <i>E. coli</i> , <i>B. subtilis</i> , <i>S. cerevisiae</i> et du virus SV40	28
Figure 4 : Schéma représentant la structure secondaire du domaine ATPase des protéines de la famille AAA ⁺	29
Figure 5 : (A) Schéma représentant les différents domaines de DnaA.....	31
Figure 6 : (A) A gauche, structure cristallographique d'un dimère de DnaA-AMP-PCP. A droite, détails structuraux du site de liaison à l'AMP-PCP	32
Figure 7 : Schéma du mécanisme de formation de l'orisome chez <i>A. aeolicus</i>	33
Figure 8 : Schéma représentant les différents domaines de DnaB	38
Figure 9 : Structure cristalline (forme B2) de l'hexamère BstDnaB	39
Figure 10 : Structure de BstDnaB•AlF ₄ •dT18	41
Figure 11 : Structure d'AaDnaB en complexe avec l'ADP	42
Figure 12 : Enveloppe de ME du complexe DnaB ₆ DnaC ₆	45
Figure 13 : Schéma de l'initiation de la réplication chez <i>A. aeolicus</i>	45
Figure 14 : (A) Chargement de l'hélicase sur l'ADNsb chez <i>G. kaustrophilus</i> . (B) Schéma du chargement de l'hélicase Mcm 2-7 eucaryote	48
Figure 15 : Schéma illustrant le chargement de l'hélicase LTag	50
Figure 16 : Schéma du chargement de l'hélicase/primase T7gp4.....	51
Figure 17 : Schéma et structures des différents domaines de DnaG	51
Figure 18 : Structure du complexe BstDnaB ₆ •DnaG ^{HBD} ₃	52
Figure 19 : Schéma de l'organisation de l'origine de réplication chez <i>H. pylori</i>	58
Figure 20 : Structure cristallographique du complexe HobA ₄ •DnaA ^{I-II} ₄	59
Figure 21 : (A) Structure du domaine NTD de HpDnaB. (B) Structure du domaine CTD de HpDnaB	62
Figure 22 : Reconstruction en trois dimensions de la structure de HpDnaB d'après l'étude par ME à coloration négative à une résolution de 23 Å	63
Figure 23 : Représentation de la répartition des charges à la surface du CTD de HpDnaB	63
Figure 24 : Enveloppes de ME de HpDnaB (en gris), de Mcm 2-7 (en bleu) et de MtMCM (en jaune).	64
Figure 25 : Représentation schématique de l'initiation de la réplication de l'ADN chez <i>H. pylori</i> , <i>E. coli</i> et <i>B. subtilis</i>	65
Figure 26 : Carte des plasmides pACYCDuet-1, pRSFDuet-1 et pET151/D-Topo.....	69

Figure 27 : Diagramme de phase pour la cristallisation d'une protéine	82
Figure 28 : Schéma du principe de la diffusion de vapeur. (A) Technique de la goutte suspendue. (B) Technique de la goutte assise.....	83
Figure 29 : Schéma représentant la composition d'une unité asymétrique, d'une maille et d'un réseau de cristal de protéine. La flèche représente une molécule de protéine.	85
Figure 30 : Représentation des 14 réseaux de Bravais. Les réseaux sont représentés en fonction de leurs systèmes cristallins et du type de réseau : (P) pour Primitif, (I) pour Centré, (F) pour Face centrée et (C) pour Base centrée.	86
Figure 31 : Représentation schématique de la diffusion d'un rayonnement traversant la matière.	87
Figure 32 : Représentation du facteur de diffusion d'une onde électromagnétique sur un plan complexe.	89
Figure 33 : Représentation vectorielle des facteurs de structure de la protéine (Fp) contenant un atome lourd ($f^o + f' + f''$).....	96
Figure 34 : Intensités diffusées et fonctions de distribution des distances de différentes formes géométriques.....	100
Figure 35 : Représentation schématique des différentes formes d'un plasmide lors d'une migration sur gel d'agarose 1% après chargement de l'hélicase DnaB.....	104
Figure 36 : Représentation schématique des réactions mises en jeu lors de la mesure de l'activité ATPase par le couplage des enzymes pyruvate kinase et lactate deshydrogénase.	105
Figure 37 : Purification de 6His-HpDnaB sur colonne d'affinité au nickel	109
Figure 38 : Purification de 6His-HpDnaB sur colonne d'exclusion stérique	110
Figure 39 : Purification de HpDnaB sur colonne d'héparine	111
Figure 40 : Purification de HpDnaB sur colonne échangeuse d'anions	112
Figure 41 : Purification de HpDnaG ^{HBD} sur colonne d'affinité au nickel	113
Figure 42 : Purification de HpDnaG ^{HBD} sur colonne d'affinité au nickel après coupure à la TEV protéase.....	114
Figure 43 : Purification du complexe HpDnaB•HpDnaG ^{HBD} sur colonne d'affinité au nickel	115
Figure 44 : Purification de HpDnaA.....	116
Figure 45 : Purification de la protéine HU	118
Figure 46 : Analyse des fractions sur gel SDS de la purification de HpDnaB ^{NTD}	119
Figure 47 : A gauche, photographies des premiers cristaux de 6His-HpDnaB obtenus. A droite, cristaux de 6His-HpDnaB obtenus en ajoutant de la spermine tetrahydrochloride comme additif aux conditions de cristallisation.....	120
Figure 48 : Photographies des cristaux de 6His-HpDnaB en trois dimensions de forme hexagonale (gauche) et cocrystallisés avec de l'ADP (droite).....	121
Figure 49 : Cristal de 6His-HpDnaB trempé avec du Ta ₆ Br ₁₂ ²⁺	123

Figure 50 : Spectre de fluorescence réalisé sur la ligne de lumière Proxima 2a (en haut). Estimation des valeurs de f'' (bleu) et f' (rouge) (en bas).....	124
Figure 51 : Carte de densité électronique obtenue après phasage par Autosolve. (A) Densité électronique du dodécamère. Vue selon l'axe de symétrie trois, (B) densité électronique correspondant à l'anneau NTD et (C) densité électronique correspondant à l'anneau CTD.	128
Figure 52 : Densité électronique correspondant à l'unité asymétrique avec les structures des domaines NTD (colorés en bleu et cyan) et des domaines CTD (en rouge et rose) placés dans la densité.	129
Figure 53 : Gel SDS des cristaux de 6His-HpDnaB coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.	129
Figure 54 : Carte 2Fo-Fc de 6His-HpDnaB.....	132
Figure 55 : Photographie des cristaux de la protéine 6His-HpDnaB-Se	133
Figure 56 : En haut, carte de densité électronique de différence anormale du sélénium (en rouge) générée après remplacement moléculaire. En bas, tableau comparant les méthionines et la valeur du σ du signal anormal du sélénium.....	135
Figure 57 : (A) Représentation schématique de l'organisation des domaines du monomère de HpDnaB. (B) Structure de 6His-HpDnaB colorée en fonction des différents domaines et représentée en ruban.....	136
Figure 58 : (A) Représentation d'un anneau NTD de 6His-HpDnaB vue du dessus. (B) Superposition des deux anneaux NTD de 6His-HpDnaB vue du dessus.....	137
Figure 59 : Comparaison entre l'anneau NTD issu de la reconstruction en ME (à gauche) et les anneaux NTD de la structure cristallographique (à droite)	138
Figure 60 : Comparaison de l'enveloppe de ME de MtMCM en marron (à gauche) avec l'enveloppe de cryo-ME de YsMcm 2-7 en magenta (à droite) et le modèle de la structure de HpDnaB (au centre).....	139
Figure 61 : Détails des interactions entre les deux anneaux NTD	140
Figure 62 : (A) Arrangement de deux monomères de 6His-HpDnaB.....	141
Figure 63 : Comparaison entre l'anneau CTD de HpDnaB (au milieu) avec l'anneau CTD de BstDnaB (à gauche) et l'anneau CTD d'AaDnaB-ADP (à droite).	141
Figure 64 : Comparaison entre l'anneau NTD de HpDnaB (au milieu) avec l'anneau NTD de BstDnaB (à gauche) et l'anneau NTD d'AaDnaB-ADP (à droite).	142
Figure 65 : Comparaison entre le HPI de la structure de HpDnaB ^{CTD} et les HPI des chaînes A et B de l'unité asymétrique de 6His-HpDnaB.....	143
Figure 66 : Cristaux de 6His-HpDnaB incubés avec un ADNsb polydT18-Cy5	144
Figure 67 : Représentation des différents modèles de HpDnaB générés à partir de notre structure cristallographique pour les formes planaires, à partir de la structure de BstDnaB•GDPAIF ₄ (PDB code : 4ESV) en complexe avec l'ADNsb pour les formes spirales-dilatées et à partir du modèle du complexe EcDnaB ₆ •EcDnaC ₆ obtenu en ME pour les formes spirales-resserrées.....	147

Figure 68 : Chromatogrammes de l'analyse en SEC-SAXS de 6His-HpDnaB et HpDnaB représentant l'intensité diffusée (en bleu) et le rayon de giration (en vert)	148
Figure 69 : Courbes SAXS expérimentales de 6His-HpDnaB et HpDnaB	148
Figure 70 : Calculs des rayons de giration de 6His-HpDnaB et HpDnaB à partir de la région de Guinier des données SAXS expérimentales	149
Figure 71 : Comparaison des données SAXS expérimentales de 6His-HpDnaB avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₆ (rouge) et la courbe du MES (violet)	150
Figure 72 : Chromatogrammes de l'analyse en SEC-SAXS de 6His-HpDnaB + ATP et HpDnaB + ATP ou AMP-PNP représentant l'intensité diffusée (bleu) et le rayon de giration (vert)	151
Figure 73 : Courbes SAXS expérimentales de 6His-HpDnaB + ATP en rouge, HpDnaB + ATP en vert et HpDnaB + AMP-PNP en bleu	152
Figure 74 : Calculs des rayons de giration de 6His-HpDnaB + ATP, HpDnaB + ATP et HpDnaB + AMP-PNP à partir de la région de Guinier des données SAXS expérimentales	153
Figure 75 : Comparaison des données SAXS expérimentales de 6His-HpDnaB + ATP avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₁₂ (vert) et la courbe du MES (violet)	154
Figure 76 : Comparaison des données SAXS expérimentales de HpDnaB + ATP avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₁₂ (vert) et la courbe du MES (violet)	155
Figure 77 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₁₂ (vert) et la courbe du MES (violet)	156
Figure 78 : Classes de HpDnaB obtenues en ME à coloration négative	157
Figure 79 : (A) Expériences d'anisotropie de fluorescence mesurant l'affinité de HpDnaB pour l'ADNdb (noir) et pour l'ADNsb (rouge) en présence ou non d'ATP et d'AMP-PNP. (B) Chromatogrammes et masses (en kDa) obtenus par SEC-MALS pour HpDnaB•ATP ou HpDnaB•ATP•dT20 (à gauche) et pour HpDnaB•AMP-PNP ou HpDnaB•AMP-PNP•dT20 (à droite)	159
Figure 80 : Chromatogrammes de l'analyse en SAXS de HpDnaB + ATP ou AMP-PNP + oligodT20 représentant l'intensité diffusée (bleu) et le rayon de giration (vert)	160
Figure 81 : Courbes SAXS expérimentales de HpDnaB + ATP + oligo-dT20 en vert, HpDnaB + AMP-PNP + oligo-dT20 en bleu	160
Figure 82 : Calculs des R _g de HpDnaB + ATP + oligodT20, HpDnaB + AMP-PNP+ oligodT20 à partir de la région de Guinier des données SAXS expérimentales	161

Figure 83 : Comparaison des données SAXS expérimentales de HpDnaB + ATP + oligodT20 avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₁₂ (vert) et la courbe du MES (violet).....	162
Figure 84 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP + oligodT20 avec les courbes théoriques calculées à partir des modèles sp-di ₆ (jaune), pl-re ₆ (rouge) et la courbe du MES (violet)	163
Figure 85 : Chromatogrammes de l'analyse en SEC-SAXS de HpDnaB + HpDnaG ^{HBD} seules ou en présence d'ATP ou AMP-PNP, représentant l'intensité diffusée (bleu) et le rayon de giration (vert).....	164
Figure 86 : Courbes SAXS expérimentales de HpDnaB + HpDnaG ^{HBD} en rouge, HpDnaB+ ATP+ HpDnaG ^{HBD} en vert et HpDnaB + AMP-PNP + HpDnaG ^{HBD} en bleu	165
Figure 87 : Calculs des rayons de giration de HpDnaB + HpDnaG ^{HBD} , HpDnaB+ ATP+ HpDnaG ^{HBD} et HpDnaB + AMP-PNP + HpDnaG ^{HBD} à partir de la région de Guinier des données SAXS expérimentales.....	166
Figure 88 : Comparaison des données SAXS expérimentales de HpDnaB + HpDnaG ^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₆ + DnaG ₃ (rouge) et la courbe du MES (violet)	168
Figure 89 : Comparaison des données SAXS expérimentales de HpDnaB + ATP + HpDnaG ^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di ₁₂ (bleu), pl-di ₆ (rouge) et la courbe du MES (violet).....	168
Figure 90 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP + HpDnaG ^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di ₆ + DnaG ₃ (jaune), pl-re ₆ (rouge) et la courbe du MES (violet).....	169
Figure 91 : Mesures de la masse (en kDa) de HpDnaB seule, mélangée ou co-exprimée avec HpDnaG ^{HBD} par SEC-MALS (Superdex S200)	170
Figure 92 : Mesures SEC-MALS de HpDnaB, mélangée ou co-exprimée avec HpDnaG ^{HBD} par SEC-MALS (Shodex KW405-4F).....	171
Figure 93 : Titration calorimétrique isotherme de 6His-HpDnaB (à gauche) et de HpDnaB (à droite), titrées par HpDnaG ^{HBD}	173
Figure 94 : (A) Mesures de la masse (en kDa) de 6His-HpDnaB mélangée avec HpDnaG ^{HBD} par SEC-MALS (Superdex S200)	174
Figure 95 : (A) Mesures des activités hélicase représentant le pourcentage d'ADNdb séparé en fonction du temps pour HpDnaB seule (noir) ou du complexe HpDnaB•HpDnaG ^{HBD} (rouge). (B) Expériences d'anisotropie de fluorescence pour la détermination de l'affinité de HpDnaB seule (noir) ou du complexe HpDnaB•HpDnaG ^{HBD} (rouge) pour l'ADNsb en présence d'ATP (à gauche) ou d'AMP-PNP (à droite).....	175

Figure 96 : Mesures de l'activité ATPase de la protéine HpDnaB seule et/ou en présence d'un oligonucléotide-dT20 et/ou de HpDnaG ^{HBD}	176
Figure 97 : Reconstruction en 3D de HpDnaB après interaction avec HpDnaG ^{HBD} avec ou sans purification sur colonne d'exclusion stérique	178
Figure 98 : Expérience d'ouverture de <i>HporiC</i> insérée dans le plasmide <i>pori1ori2</i>	179
Figure 99 : Expérience d'ouverture de <i>HporiC</i> insérée dans le plasmide <i>pori1ori2</i> en présence de HpDnaA et HpDnaB.....	180
Figure 100 : Essai d'ouverture de <i>HporiC</i> insérée dans le plasmide <i>pori1ori2</i> en présence de HpDnaA, HpDnaB et HpDnaG ^{HBD}	181
Figure 101 : (A) Expérience de chargement de l'hélicase DnaB de <i>Pseudomonas aeruginosa</i> (PaDnaB) (Jiang <i>et al.</i> , 2005). (B) Schéma représentant le passage de la forme F1 à F1* par l'action de la gyrase et de SSB une fois le primosome formé	182
Figure 102 : Expérience de chargement de l'hélicase 6His-HpDnaB et HpDnaB en utilisant toutes les protéines nécessaires à la formation du primosome	182
Figure 103 : Expérience de chargement de l'hélicase 6His-HpDnaB ou HpDnaB en testant différentes combinaisons de protéines	183
Figure 104 : Expérience de chargement de l'hélicase 6His-HpDnaB et HpDnaB sans la protéine HobA	184
Figure 105 : Alignement de la séquence de HpDnaB (résidus 1 à 34) avec celles d'autres espèces bactériennes	187
Figure 106 : Alignement de la séquence de HpDnaB (résidus 35 à 98) avec celles d'autres espèces bactériennes	188
Figure 107 : Alignement de la séquence de HpDnaB (résidus 390 à 449) avec celles d'autres espèces bactériennes	189
Figure 108 : Modèles de HpDnaB en conformation spirale-dilatée (A) et planaire-dilatée (B).....	191
Figure 109 : Zone d'interaction (orange) d'une molécule de BstDnaG (vert) avec deux BstDnaBNTD (bleu clair et foncé) (à gauche). Représentation de cette même zone sur deux NTD de HpDnaB (à droite).	192
Figure 110 : Comparaison du modèle de l'initiation de la réplication chez <i>H. pylori</i> (à gauche) avec ceux d' <i>E. coli</i> et de <i>B. subtilis</i> (à droite).	193

Liste des Tableaux

Tableau 1 : Comparaison des protéines impliquées dans l'initiation de la réplication chez les bactéries, les virus, les archées et les eucaryotes.	43
Tableau 2 : Tableau des constructions réalisées.	70
Tableau 3 : Correspondance des différentes familles cristallines, réseaux de Bravais, systèmes cristallins et groupes ponctuels.	87
Tableau 4 : Tableau des statistiques du jeu de données des cristaux natifs de 6His-HpDnaB. Les valeurs entre parenthèses correspondent aux statistiques de la dernière coquille de résolution.	122
Tableau 5 : Tableau de probabilité du nombre de molécules par unité asymétrique.	123
Tableau 6 : Statistiques des données de diffraction des deux jeux Tant 1 et Tant 2. Les valeurs entre parenthèses correspondent aux statistiques de la dernière coquille de résolution.	125
Tableau 7 : Tableau de statistiques après phasage par SAD en utilisant le groupe d'espace $I2_13$	126
Tableau 8 : Statistiques d'affinement de la structure de 6His-HpDnaB à une résolution de 6,7 Å.	131
Tableau 9 : Statistiques des données de diffraction d'un cristal de 6His-HpDnaB-Se au seuil du sélénium. Les valeurs entre parenthèses correspondent à la dernière coquille de résolution.	134
Tableau 10 : Statistiques des données de diffraction d'un cristal de 6His-HpDnaB trempé avec un ADNsb polydT25-Cy5. Les valeurs entre parenthèses correspondent à la dernière coquille de résolution.	145
Tableau 11 : Tableau représentant l'ensemble des χ^2 des différents modèles testés ainsi que les résultats de l'expérience de MES. Les calculs ont été réalisés sur le serveur web FoXs.	149
Tableau 12 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de 6His-HpDnaB + ATP, HpDnaB + ATP et HpDnaB + AMP-PNP. Les calculs ont été réalisés sur le serveur web FoXs.	154
Tableau 13 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de HpDnaB + ATP + oligo dT20, HpDnaB + AMP-PNP+ oligo dT20. Les calculs ont été réalisés sur le serveur web FoXs.	162
Tableau 14 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de HpDnaB + HpDnaG ^{HBD} , HpDnaB + ATP + HpDnaG ^{HBD} et HpDnaB + AMP-PNP+ HpDnaG ^{HBD} . Les calculs ont été réalisés sur le serveur web FoXs.	167

Liste des abréviations

°C/K: Degré Celsius/Kelvin	MCM: Mini chromosome maintenance
A, T, G, C: Adénine, Thymine, Guanine, Cytosine	MES: Recherche des ensembles minimaux
Å: Ångström	ME: Microscopie électronique à transmission
AAA⁺: ATPase associated with diverse cellular Activities	LB: lysogeny broth
ADN: Acide désoxyribonucléique	NAD/NADH: Nicotinamide adénine dinucléotide réduit/oxydé
ADNdb: Acide Désoxyribonucléique double brin	NTD: Domaine N-terminal
ADNsb: Acide Désoxyribonucléique simple brin	NTP: Nucléotide Triphosphate
ADP: Adénosine Diphosphate	OBD: Origin Binding Domain
AMP-PCP: 5'-adenylyl(β,γ -methylene) diphosphonate	ORC: Origin Replication Complex
AMP-PNP: Adénylyl-imidodiphosphate	pb: paire de bases
ARN: Acide ribonucléique	PCR: Réaction en chaîne par polymérase
ARS: Autonomously replicating sequence	PDB: Protein Data Bank
ASCE: Additional Strand Conserved Glutamate	PEG: Poly-éthylène glycol
ATP: Adénosine Triphosphate	PEP: Phosphoénolpyruvate
CTD: Domaine C-terminal	pH: potentiel hydrogène
Dmax: Distance maximale	pI: point isoélectrique
DO: Densité Optique	PK/LDH: Pyruvate kinase/Lactate déshydrogénase
DTT: Dithiothréitol	R/Rfree: coefficients d'accord entre les facteurs de structure observés et calculés
DUE: DNA unwinding élément	Rg: Rayon de giration
EDTA: Ethylène Diamine Tétra Acétique	RMN: Résonance magnétique nucléaire
ESRF: European Synchrotron Radiation Facility	Rsym / Rmeas: Coefficient d'accord entre réflexions symétriques / pondéré
EXAFS: Extended X-Ray Absorption Fine Structure	SAD: Singlewavelength Anomalous Diffraction
GDP: Guanosine Diphosphate	SAXS: Small Angle X-ray Scatering
GTP: Guanosine Triphosphate	SDS: Sodium Dodecyl-Sulfate
HBD: Helicase Binding Domain	SEC-MALS: Size exclusion chromatographie-Multi Angle Laser light Scatering
HEPES: Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique	SV40: Simian virus 40
HPI: Helicobacter pylori insertion	TEV: Tobacco Etch Virus
HTH: Hélice-Tour-Hélice	Tris: 2-amino-2-hydroxyméthyl-1,3-propanediol
HU : Protéine histone-like	Ts: Thermosensible
I/σ(I): Rapport signal sur bruit	UV: Ultraviolet
IPTG: Isopropyl β -D-1-thiogalactopyranoside	Vm: Volume de Matthews
ITC: Titration calorimétrique isotherme	βME: Beta-mercaptoéthanol
K_d : Constante de dissociation	
MALT: Lymphome des tissus lymphoïdes à muqueuses associées	

1. Introduction

1.1. Généralités et avant-propos

En 1953, James D. Watson et Francis H. C. Crick publient dans la revue *Nature* (Watson et Crick, 1953), une proposition de structure tridimensionnelle de la molécule d'ADN. Ils ont montré en utilisant les clichés de diffraction aux rayons X de Rosalind Franklin que la molécule d'ADN est une hélice droite formée de deux chaînes dont le squelette est composé de la succession des désoxyriboses reliés entre eux par un phosphate. La base azotée quant à elle se projette au centre de l'hélice et se lie à la base azotée du brin complémentaire. Ces deux polymères appelés « brins d'ADN » sont le reflet de l'un par rapport à l'autre par complémentarité des bases. Cela signifie que lorsqu'une adénine est présente sur un brin, une thymine est présente sur le deuxième brin et lorsque l'on trouve une guanine, on retrouve une cytosine. Cette hypothèse était en accord avec Edwin Chargaff qui observait que le rapport $A/T=C/G=1$ (Kornberg et Baker, 1992; Voet et Voet, 2005). Ces découvertes ont permis de suggérer que l'ADN (et non les protéines comme le pensait jusqu'à la communauté scientifique) était la molécule support des gènes. La détermination de la structure de l'ADN a également permis de comprendre comment un mécanisme de réplication permettait de transmettre l'information génétique aux générations suivantes.

Une cellule, qu'elle soit procaryote ou eucaryote, a un cycle de vie : elle naît, croît, se multiplie et meurt. La division cellulaire est l'étape de la vie cellulaire durant laquelle la cellule produit deux cellules filles identiques, ayant le même patrimoine génétique. Pour cela, la cellule mère duplique son ADN par un mécanisme appelé la réplication. Ce processus est l'un des plus conservés parmi les mécanismes biologiques fondamentaux du vivant.

En 1963, François Jacob et Sydney Brenner élaborent la théorie du réplicon (Jacob et Brenner, 1963). C'est en dessinant leurs idées sur le sable d'une plage lors d'un congrès international qu'ils en sont arrivés à cette hypothèse (Figure 1). Ils pensaient que la bactérie avait besoin d'une extrême précision et de spécificité pour dupliquer son ADN à un moment précis de la vie cellulaire et donc que la réplication ne pouvait pas être régulée seulement par un ensemble de précurseurs disponibles dans la cellule. Ils ont ainsi postulé que la réplication d'un plasmide bactérien ou d'un chromosome pouvait être contrôlée par une séquence dite en

« cis » et d'un gène codant une protéine initiateur qui agirait en « trans ». En venant se fixer sur la séquence en « cis », cette protéine permettrait de contrôler le début de la réplication.



Figure 1 : Dessin dans le sable de la théorie du réplicon (Jacob et Brenner, 1963). Le cercle représente le réplicon (unité de réplication) et la boîte correspond à la séquence dite en « cis ». La flèche indique l'endroit du génome où est codée la protéine initiateur « trans » et où elle se fixe une fois synthétisée.

Depuis cette hypothèse, les recherches dans le domaine de la réplication de l'ADN ont été extrêmement abondantes. On distingue maintenant les protéines permettant l'initiation de la réplication de celles impliquées dans la synthèse de l'ADN proprement dite. Bien que les enzymes impliquées dans ces deux étapes existent chez tous les êtres vivants, elles présentent des caractéristiques différentes selon les organismes.

La réplication bactérienne peut être divisée en trois étapes : l'initiation, l'élongation et la terminaison (Figure 2A). Chez la plupart des bactéries, l'initiation de la réplication se déroule sur un locus précis du chromosome, l'origine de réplication ou *oriC*. La protéine DnaA liée à l'ATP se fixe sur des sites spécifiques d'*oriC*, appelés boîtes à DnaA. La présence d'ATP va permettre la formation d'un complexe nucléoprotéique appelé orisome, basé sur l'oligomérisation de DnaA sur les boîtes à DnaA. La formation de l'orisome entraîne alors l'ouverture d'une zone d'*oriC* riche en base AT appelée DUE pour « DNA unwinding element » (Figure 2B). Cette région d'ADN sert de point d'ancrage au réplisome. Il est constitué de l'ensemble des protéines nécessaires à la synthèse des brins complémentaires : le complexe DnaB•DnaG ou primosome et l'ADN polymérase III holoenzyme (revue dans Schaeffer *et al.*, 2005). La première protéine chargée sur l'orisome est l'hélicase DnaB qui

permet de désappairier l'ADNdb lors de l'étape d'élongation. La primase DnaG se lie à DnaB et synthétise de courts fragments d'ARN qui serviront d'amorces à la polymérase. Par ailleurs, la protéine SSB (« single strand binding protein ») se lie à l'ADNsb afin de l'empêcher de se réappairier.

L'ADN polymérase III holoenzyme d'*Escherichia coli* est un complexe protéique composé de 17 sous-unités (Figure 2B). Il y a deux unités polymérase/exonucléase (Pol III) sur chaque brin parental. Chaque domaine Pol III est lié à une enzyme dimérique entourant l'ADN : la pince β , qui augmente la processivité de Pol III. Les deux complexes Pol III/pince β sont reliés au primosome (qui désapparie l'ADNdb) ainsi qu'au facteur de chargement de la pince β par la sous-unité τ . Cette sous-unité assure le couplage de la réplication des deux brins (Figure 2B). Les deux brins d'ADN étant antiparallèles et les domaines Pol III ne se déplaçant que dans le sens 5' en 3', la synthèse du néo brin doit se faire de façon continue sur le premier brin (brin avancé) et de façon discontinue sur le second brin (brin retard). Une fois la synthèse des brins complémentaires achevée, les deux chromosomes entiers sont séparés (Kornberg et Baker, 1992; Voet et Voet, 2005). Chez *E. coli* la terminaison de la réplication fait appel à des sites de terminaison (*ter*) sur le chromosome et à la protéine Tus (Mulcair *et al.*, 2006). Cependant, cette protéine ne semble pas être conservée chez toutes les bactéries et le mécanisme de terminaison reste encore mal défini.

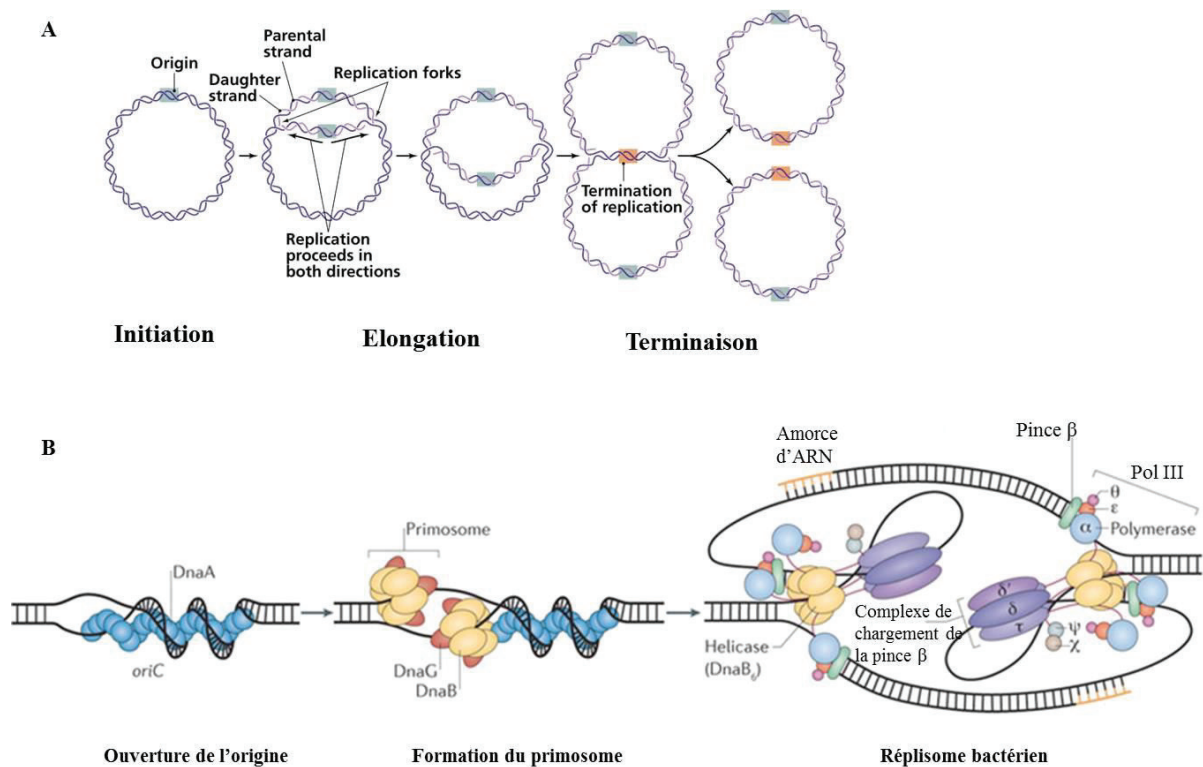


Figure 2 : (A) Etapes de la réplication procaryote. (B) Etapes de l'initiation de la réplication et de l'assemblage de la polymérase (Robinson et Van Oijen, 2013).

Les développements de la génétique, de la biologie moléculaire et de la biochimie ont permis l'identification des enzymes dédiées qui initient la réplication de l'ADN. L'avènement de la biologie structurale et en particulier de la cristallographie des protéines, a permis de réaliser des progrès spectaculaires dans le domaine durant les quinze dernières années. Dans cette introduction nous nous focaliserons sur les acteurs et les mécanismes moléculaires de l'initiation de la réplication chez les bactéries en comparant parfois ces derniers à ceux trouvés chez d'autres organismes.

1.2. Initiation de la réplication bactérienne

1.2.1. Origine de réplication

1.2.1.1. Origine de réplication chez les bactéries à Gram négatif

E. coli est considéré comme un modèle pour l'étude de la réplication chez les bactéries à Gram négatif. Son origine de réplication *oriC* est une séquence unique d'environ 260 pb présente sur son chromosome. Elle est composée d'une région DUE et d'une séquence où se

fixe la protéine initiatrice DnaA composée de boîtes DnaA (appelées cluster). La région DUE contient trois séquences de 13 pb riches en AT (Kornberg et Baker, 1992). Les sites principaux d'attachement de DnaA sont appelés sites R1 à R5 et contiennent 9 pb dont la séquence consensus est TTATNCACA (Fuller *et al.*, 1984). Des études biochimiques ont permis de classer ces boîtes en fonction de leur affinité pour DnaA : R1=R4>R2>R3=R5 (revue dans Ozaki et Katayama, 2009). On distingue également des boîtes ayant une affinité variable selon que DnaA soit liée à l'ATP ou à l'ADP. Ainsi, les sites R1 et R4 ont une affinité similaire pour DnaA-ATP ou DnaA-ADP alors que les sites R2, R3 et R5 fixent plus efficacement DnaA-ATP. Par ailleurs, il existe des sites supplémentaires sur lesquels se fixe DnaA, appelés sites I1 à I3 (Grinwade *et al.*, 2000 ; Ryan *et al.*, 2002) et sites τ 1 et τ 2 (revue dans Katayama *et al.*, 2010 ; revue dans Leonard et Grimwade, 2011 ; McGarry *et al.*, 2004). Les sites I1 à I3 sont des séquences de 9 pb situées entre les sites de hautes affinités R1 et R4. Ces sites (comme les sites τ 1 et τ 2) ont une meilleure affinité pour DnaA-ATP en comparaison avec DnaA-ADP. Cependant, cette affinité n'est pas aussi élevée comparée à celles des sites R1 et R4 (Figure 3) (revue dans Katayama *et al.*, 2010).

Si l'origine de réplication d'*E. coli* a été étudiée en détail, on ne trouve que peu d'informations sur les origines de réplication des autres bactéries à Gram négatif. Il semble cependant que la séquence des boîtes à DnaA soit assez conservée mais leurs nombres et leurs dispositions varient considérablement (Zawilak-Pawlik *et al.*, 2005).

1.2.1.2. Origine de réplication des bactéries à Gram positif

Chez quelques bactéries à Gram positif telles que *Bacillus subtilis* (bactérie modèle pour ce type de bactéries), on trouve des origines de réplication divisées en plusieurs clusters de boîtes à DnaA. Chaque cluster contient des séquences pouvant fixer DnaA comme les sites R et les sites I pour *E. coli*. Les différents monomères de DnaA, liés à l'ATP, vont interagir entre eux à l'intérieur de chaque cluster mais également permettre l'interaction entre les différents clusters et conduire l'ADN à former une boucle. On note aussi chez les bactéries à Gram positif la présence de la région DUE. Néanmoins, les séquences riches en AT sont de tailles différentes par rapport à celle d'*E. coli* : 27 pb chez *B. subtilis*, 19 pb chez *Mycobacterium tuberculosis* (Figure 3).

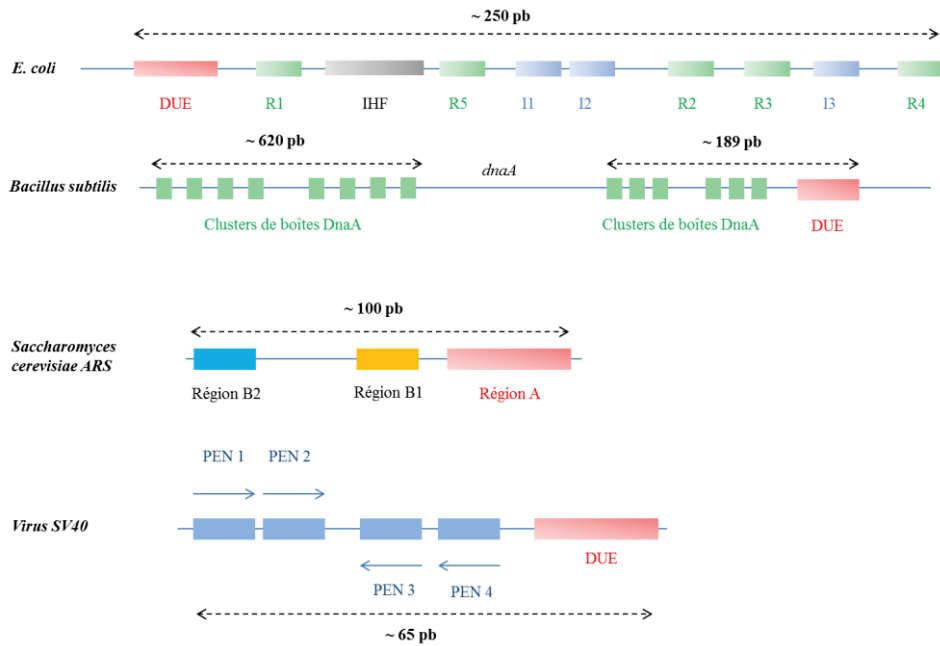


Figure 3 : Schéma représentant l'organisation de l'origine de réplication chez *E. coli*, *B. subtilis*, *S. cerevisiae* (adapté de la revue Costa *et al.*, 2013) et du virus SV40 (Chang *et al.*, 2013).

1.2.1.3. Origine de réplication des eucaryotes

L'organisme le mieux étudié est la levure *Saccharomyces cerevisiae* qui sert de modèle pour les organismes eucaryotes. Chez ces derniers, les origines de réplication sont appelées ARS (Autonomously replicating sequence) (Stinchcomb *et al.*, 1979). Ces séquences ARS mesurent environ 100 pb et sont composées de deux régions A et B (Figure 3). La région A est une séquence de 11 pb (5' (A/T)TTTA(T/C)(A/G)TTT(A/T)3'). La séquence de la région B n'est pas conservée mais elle est tout aussi importante pour l'initiation de la réplication que la région A (Marahrens et Stillman, 1992). En effet, la région A et la première partie de la région B (région B1) sont reconnues par la protéine initiateur eucaryote ORC (Origin Replication Complex) (Bell et Stillman, 1992 ; Diffley *et al.*, 1994). Le site de formation des fourches se situe au niveau de la région B. L'interaction entre ORC et la région A requiert la présence d'ATP mais contrairement à DnaA, une seule molécule d'ORC se fixe sur une ARS (Figure 3).

1.2.1.4. Origine de réplication virale, exemple du virus SV40

La majorité de nos connaissances sur la réplication de l'ADN eucaryote proviennent des études sur la levure mais également du virus SV40 (Voet et Voet, 2005). En effet, comme

la plupart des virus, SV40 dépend de sa cellule hôte pour répliquer son ADNdb. A l'exception de la protéine Large T antigen (LTag), toutes les protéines nécessaires pour la réplication de l'ADN sont des protéines de la machinerie eucaryote. La réplication du génome de SV40 peut être étudiée *in vitro* à partir d'un plasmide contenant l'origine de réplication SV40, en présence d'extraits de cellules infectées et de la protéine LTag. L'origine de réplication du virus SV40 est une séquence de 64 pb, divisée en trois parties. Une partie centrale contient quatre séquences pentanucléotidiques 5'GAGGC3' (PEN1 à PEN4). Cette partie est bordée d'un côté par un palindrome et de l'autre par une région riche en AT (Deb *et al.*, 1986 ; Chang *et al.*, 2014 ; Soultanas, 2012) (Figure 3). La protéine LTag reconnaît les séquences PEN de l'origine de réplication virale et s'oligomérisent pour former des dodécamères. La protéine LTag joue à la fois le rôle de protéine initiateur et d'hélicase répllicative.

1.2.2. Protéines de l'initiation de la réplication AAA+

Les protéines impliquées dans l'initiation se lient à l'ATP et utilisent son hydrolyse pour leurs fonctions. C'est le cas des protéines DnaA et DnaB qui, de part cette fonction, vont contenir des motifs structuraux communs. DnaA fait partie de la famille des protéines AAA⁺ (ATPase associated with diverse cellular activities). Ces dernières sont des membres de la branche ASCE (additional strand conserved glutamate) des protéines de type phosphate-binding loop (P-loop). Le domaine de ces protéines qui lie le NTP a une structure conservée (Figure 4).

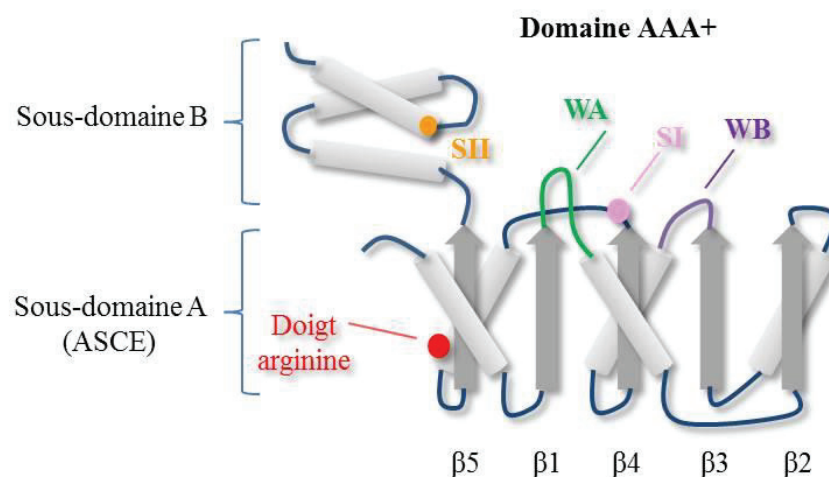


Figure 4 : Schéma représentant la structure secondaire du domaine ATPase des protéines de la famille AAA⁺. Les cylindres représentent des hélices α et les flèches des brins β . Les boucles en vert et en violet représentent les motifs Walker A et le Walker B, les motifs en rose, orange et rouge sont respectivement le senseur I, le senseur II et le doigt arginine.

Ce domaine est constitué d'un feuillet β (comprenant cinq brins β parallèles) pris en « sandwich » par trois hélices α d'un côté et deux hélices α de l'autre, adoptant un repliement $\alpha\beta\alpha$. Les protéines AAA⁺ possèdent un domaine $\alpha\beta\alpha$ (sous-domaine A) et un motif structural supplémentaire : trois hélices α regroupées (sous-domaine B) venant recouvrir le domaine $\alpha\beta\alpha$. Ces domaines contiennent les motifs essentiels, communs à toutes les protéines AAA⁺ (Figure 4) :

- Le Walker A est une séquence consensus GxxGxGKT/S localisée entre le brin β 1 et l'hélice α 1 du domaine ASCE. C'est à cet endroit que se lie le nucléotide.
- Le Walker B est une séquence consensus D (D/E) située sur l'apex du brin β 3
- Le senseur I est généralement un résidu polaire (souvent N mais on peut également trouver D, T, H). Ce résidu est situé sur l'apex du brin β 4.
- Le senseur II est une arginine à la base de l'hélice α 7 au niveau du sous-domaine B. Elle interagit avec le phosphate γ du nucléotide. Cette arginine est très importante puisqu'elle va permettre de discriminer si le Walker A est lié à l'ATP ou à l'ADP, entraînant ainsi des changements de conformation du domaine AAA⁺.
- Le doigt arginine correspond à une arginine qui va également se lier au phosphate γ du nucléotide.

1.2.3. Structure et mode d'action de l'initiateur de la réplication DnaA

1.2.3.1. DnaA, protéine initiatrice de la réplication

La protéine initiatrice DnaA est requise pour la formation d'un complexe nucléoprotéique sur l'origine de réplication. Elle est présente chez les eubactéries à l'exception des bactéries endosymbiotiques *Blochmannia floridanus* et *Wigglesworthia glossinidia* (Gil *et al.*, 2003). DnaA est constituée de quatre domaines distincts qui ont des fonctions spécifiques et dont les structures ont été déterminées séparément (Figure 5). Le domaine I est un domaine globulaire dont la séquence est peu conservée (revue dans Kaguni, 1997 ; Messer, 2002 ; Sutton et Kaguni, 1997). Ce domaine est impliqué dans plusieurs interactions protéine-protéine qui sont importantes pour la régulation de la réplication (Ishida *et al.*, 2004 ; Noiro-Gros *et al.*, 2006 ; Wagner *et al.*, 2009). Chez *E. coli*, le domaine I interagit aussi avec DnaB et cette interaction est nécessaire à l'initiation de la réplication (Seitz *et al.*, 2000 ; Sutton *et al.*, 1998). Le domaine I interagit également avec DiaA

(Keyamura *et al.*, 2007) et Hda (Su'etsugu *et al.*, 2013) chez *E. coli*, SirA chez *B. subtilis* (Jameson *et al.*, 2014) et HobA chez *Helicobacter pylori* (Natrajan *et al.*, 2009). L'interaction du domaine I avec ces protéines joue un rôle de régulation de l'initiation de la réplication. Le domaine II est un domaine non structuré dont la longueur varie également selon les espèces bactériennes. Son rôle est mal défini bien que sa taille soit importante dans le contrôle de la réplication (Molt *et al.*, 2009). Le domaine III contient le domaine AAA⁺, permettant la liaison et l'hydrolyse de l'ATP. Le domaine III est divisé en deux sous-domaines IIIA et IIIB qui correspondent aux sous-domaines structuraux A et B des protéines AAA⁺. Le sous-domaine IIIB, également appelé « lid », est constitué d'hélices α recouvrant le sous-domaine IIIA (Erzberger *et al.*, 2002). Le domaine IV se fixe sur les boîtes DnaA via son motif hélice-tour-hélice (Fujikawa *et al.*, 2003) (Figure 5).

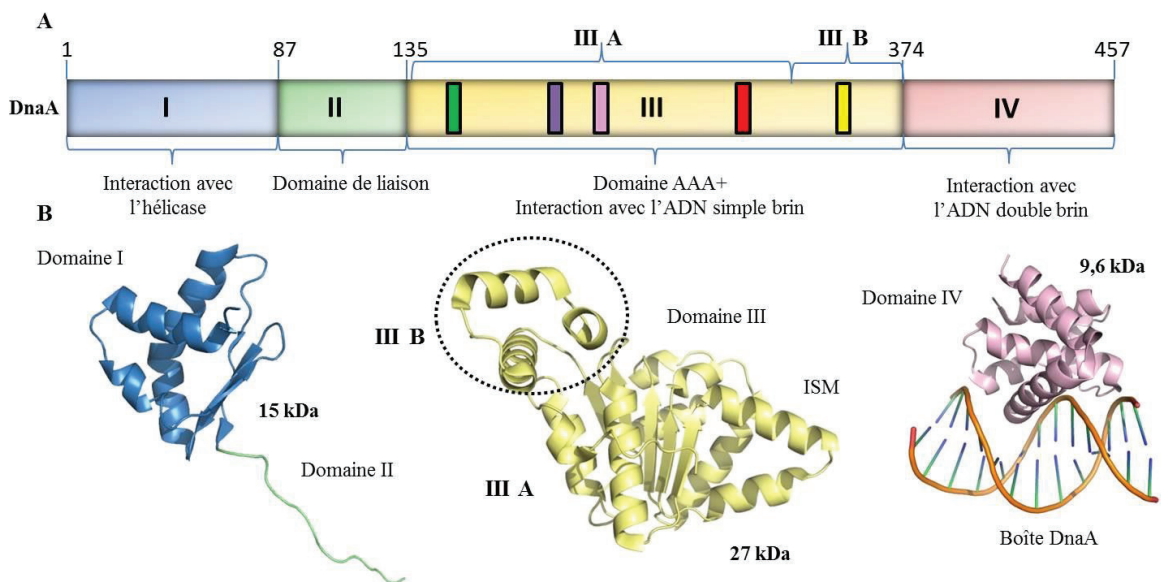


Figure 5 : (A) Schéma représentant les différents domaines de DnaA. Les motifs Walker A et B sont représentés en vert et en violet, respectivement. Le senseur I est coloré en rose, le doigt arginine en rouge et le senseur II en jaune. (B) Structures RMN des domaines I et II de DnaA d'*E. coli* (PDB : 2E0G; Abe *et al.*, 2007). Structures cristallographiques du domaine III de DnaA d'*A. aeolicus* (PDB : 2HCB, Erzberger *et al.*, 2006), et du domaine IV de DnaA d'*E. coli* en complexe avec une boîte DnaA (PDB : 1J1V, Fujikawa *et al.*, 2003).

1.2.3.2. Mécanisme d'oligomérisation de DnaA

Une des fonctions principales de DnaA est de s'oligomériser sur *oriC* pour former l'orisome en présence d'ATP. Cette propriété a été observée très tôt en microscopie électronique à transmission (ME) (Fuller *et al.*, 1984). Si la véritable structure tridimensionnelle de l'orisome est encore inconnue, la structure cristallographique du domaine III-IV de DnaA d'*Aquifex aeolicus* (AaDnaA^{III-IV}) en complexe avec l'AMP-PCP, un

analogue non hydrolysable de l'ATP, a permis d'obtenir des détails importants sur le mode d'oligomérisation de la protéine. Ainsi, le senseur II (arginine) du sous-domaine IIIB de DnaA joue un rôle crucial dans l'assemblage des molécules de DnaA. En effet, l'arginine interagit avec le phosphate γ de l'AMP-PCP et positionne le site actif dans un état dit « ouvert ». Cet état permet au doigt arginine (à ne pas confondre avec le senseur II) d'une seconde molécule de DnaA d'interagir avec l'ATP de la première. Cet arrangement structural est répété dans le cristal ce qui conduit à la formation d'un oligomère (Erzberger *et al.*, 2006) (Figure 6). Au sein du sous-domaine IIIA est insérée une petite hélice α appelée ISM pour « initiator specific motif ». Cette insertion dépasse du domaine ASCE et empêche la formation d'un anneau fermé de DnaA. La conséquence est un décalage des sous-unités de DnaA, ce qui conduit à une conformation en hélice droite (Erzberger *et al.*, 2006 ; Iyer *et al.*, 2004). Ce mode d'oligomérisation explique que l'ATP est nécessaire à l'oligomérisation de DnaA. En effet, dans le cas de la présence d'ADP, le doigt arginine de la seconde molécule de DnaA ne peut plus interagir avec le phosphate γ : il n'y a donc plus d'oligomérisation.

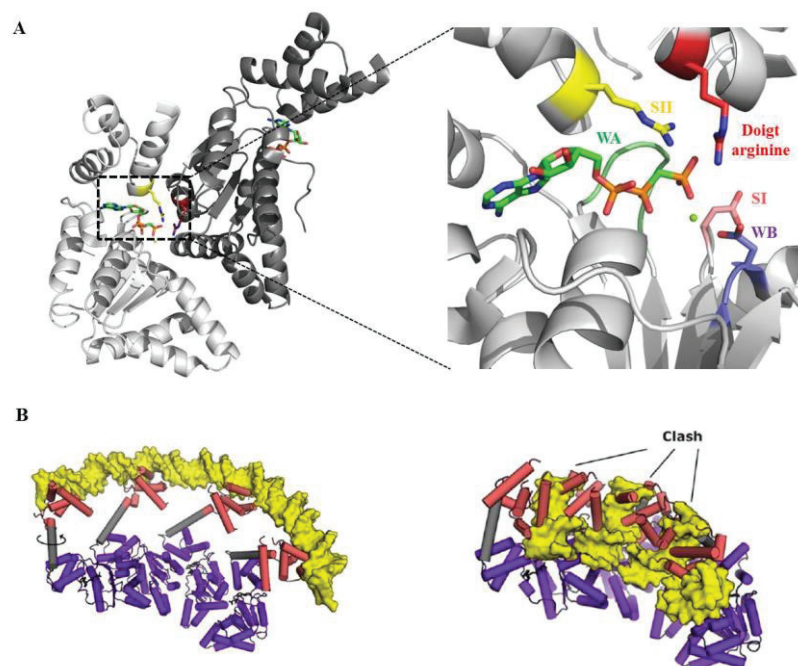


Figure 6 : (A) A gauche, structure cristallographique d'un dimère de DnaA-AMP-PCP. A droite, détails structuraux du site de liaison à l'AMP-PCP permettant l'oligomérisation de DnaA (PDB : 2HCB ; Erzberger *et al.*, 2006). (B) A gauche, modèle de DnaA^{III-IV} lié à l'ADNdb obtenu par rotation du domaine IV. A droite, filament de AaDnaA (PDB : 2HCB) superposé avec la structure de EcDnaA_{IV} en complexe avec une boîte DnaA (PDB 1J1V), montrant que les domaines IV ne peuvent plus se lier à l'ADNdb dans cette conformation.

La structure d'AaDnaA^{III-IV} en complexe avec l'ATP et un oligonucléotide polyA de 12 pb (PDB : 3R8F) a montré que l'ADNs_b s'attache à l'intérieur de l'hélice formée par DnaA de manière similaire aux filaments de RecA, une autre protéine AAA⁺ impliquée dans la recombinaison homologue (Duderstadt *et al.*, 2011). Cependant, dans l'hélice d'AaDnaA^{III-IV}, le domaine IV est orienté de telle manière qu'il ne peut lier les boîtes à DnaA. Ainsi, l'hélice d'AaDnaA^{III-IV} n'est pas compatible avec la liaison à l'ADN_{db} (boîtes à DnaA) et l'oligomérisation sur l'*oriC*. Les auteurs proposent donc que deux modes d'assemblage de DnaA existent : un sur le simple brin dit compacte tel qu'observé dans la structure cristallographique et un sur le double brin, dit étiré qui reste à déterminer (Figure 6 et Figure 7).

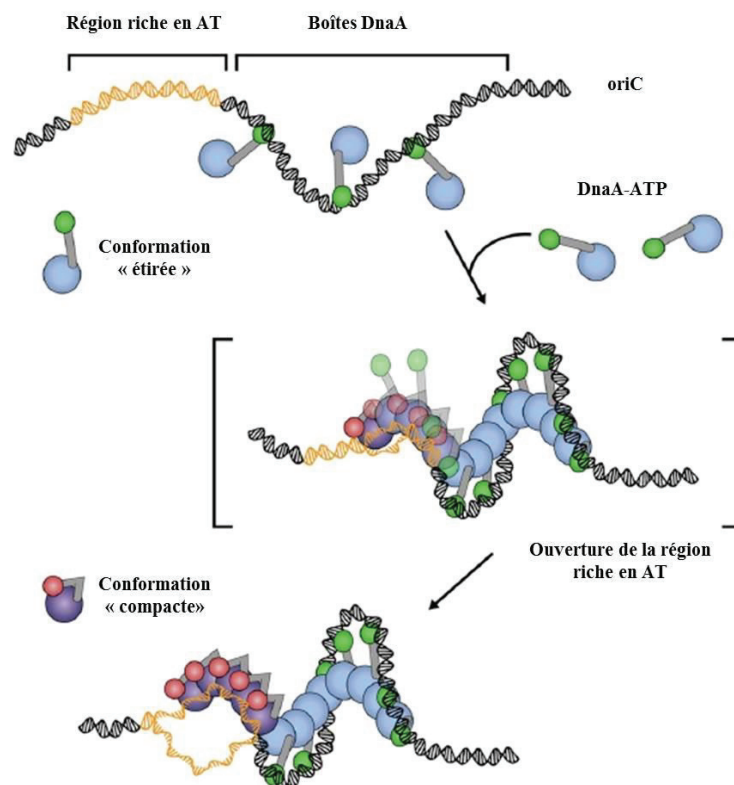


Figure 7 : Schéma du mécanisme de formation de l'orisome chez *A. aeolicus*. La protéine initiateurice DnaA-ATP se fixe sur les boîtes DnaA. La présence d'ATP va permettre l'oligomérisation de DnaA qui va former un filament et faciliter l'ouverture de la région riche en AT (Duderstadt *et al.*, 2011).

1.2.3.3. Régulation de l'initiation par le contrôle de DnaA

La réplication de l'ADN doit être synchronisée avec la division cellulaire. Dans les colonies bactériennes en croissance exponentielle, l'initiation de la réplication ne doit avoir lieu qu'une seule fois par génération. Pour cela, l'action de la protéine initiateurice DnaA est

finement régulée que ce soit au niveau de sa disponibilité dans la cellule, de son activité ou encore de sa fixation à l'origine. L'expression de la protéine DnaA est auto régulée car DnaA peut se fixer sur le promoteur situé en amont de son gène, entraînant ainsi une inhibition de sa transcription (Atlung *et al.*, 1985 ; Braun *et al.*, 1985 ; Kucherer *et al.*, 1986 ; Wang et Kaguni, 1987). DnaA peut également être régulée par titration. Le génome d'*E. coli* contient une trentaine de séquences *datA* réparties sur tout le chromosome. DnaA se fixe sur ces séquences durant la croissance bactérienne ce qui empêche le démarrage de la réplication (Kitagawa *et al.*, 1996 ; Kitagawa *et al.*, 1998). Cependant, le système de répression prédominant chez *E. coli* reste le système RIDA (Regulatory Inactivation of DnaA) (Camara *et al.*, 2005). Le complexe RIDA est formé de la protéine HdA (protéine homologue au domaine III de DnaA), de l'ADP (qui permet d'activer HdA) et de la pince β (Katayama *et al.*, 1998 ; Kato *et al.*, 2001). Ce système stimule l'hydrolyse de l'ATP fixé à DnaA provoquant ainsi la dissociation de l'oligomère de DnaA d'*oriC* et empêche ainsi un nouveau cycle réplcatif (Kato *et al.*, 2001 ; Su'etsugu *et al.*, 2005 ; Xu *et al.*, 2009).

Un autre mécanisme, appelé « séquestration » fait intervenir la protéine SeqA et des séquences spécifiques GATC. L'origine de réplication d'*E. coli* contient onze copies de ces séquences et lors de la réplication ces sites sont méthylés sur les adénines par la méthylase Dam (revue dans Kaguni, 2006 ; Kornberg et Backer, 1992 ; revue dans Messer, 2002). La réplication étant semi-conservative, il en résulte un ADN méthylé sur un seul brin (hémi-méthylé). Les séquences GATC hémi-méthylées sont des sites de fixation de la protéine SeqA. Une fois fixée, SeqA empêche l'attachement de DnaA sur les boîtes de faibles affinités de l'origine (Lu *et al.*, 1994 ; Slater *et al.*, 1995). Des études ont montré que la délétion du gène *dam* ou des séquences GATC dans l'*oriC* entraîne une réplication de l'ADN asynchrone et des initiations répétées (revue dans Kaguni, 2006).

Il existe également des protéines permettant de stimuler l'initiation de la réplication chez *E. coli*. Ce sont les protéines IHF, Fis et DiaA. IHF (Integration Host Factor) et Fis (Factor for inversion of stimulation) sont des protéines qui reconnaissent des séquences d'ADN distinctes au niveau de l'origine. IHF est une protéine Histone-like qui intervient dans l'ouverture d'*oriC* chez *E. coli* (revue dans Katayama *et al.*, 2010 ; Ryan *et al.*, 2002). A l'inverse de IHF, Fis inhibe l'ouverture de l'origine en s'attachant sur *oriC*. Cependant, des études récentes suggèrent que le rôle de Fis dépendrait du cycle cellulaire (Flatten et Skarstad, 2013). Un autre mécanisme fait appel à la protéine DiaA. Cette protéine est essentielle chez *E. coli* pour la synchronisation du cycle cellulaire et de l'initiation de la réplication. DiaA est un régulateur

positif qui se lie sur le domaine I de DnaA (Ishida *et al.*, 2004). La fonction de DiaA n'est pas claire car elle peut à la fois stimuler l'attachement de DnaA et empêcher le chargement de DnaB (Keyamura *et al.*, 2007).

Chez *B. subtilis*, la réplication peut être régulée par les protéines SirA, YabA et Soj. D'une manière générale, l'action de ces protéines inhibe l'initiation de la réplication. SirA est un facteur de répression de l'initiation de la réplication intervenant lors de la sporulation. La bactérie privée du gène *sirA* présente des phénotypes d'extra-initiation de la réplication durant la sporulation. Des analyses de fluorescence utilisant DnaA-GFP suggèrent que la liaison de SirA à DnaA dissocie DnaA de l'*oriC*. La liaison directe de SirA à DnaA a été montrée par double hybride (Wagner *et al.*, 2009).

YabA interagit avec DnaA et la pince β de la polymérase (Noirot-Gros *et al.*, 2002). YabA purifiée forme des homotétramères qui se lient directement à DnaA et la pince β (Noirot-Gros *et al.*, 2006). La délétion du gène codant pour YabA ou l'incorporation de mutations dans la protéine provoque des phénomènes d'extra-initiation. Ceci indiquerait que YabA réprime l'activité de DnaA en présence de la pince β . YabA a donc une fonction homologue à Hda chez *E. coli* bien qu'elle ne possède pas de similarité de séquence (Hayashi *et al.*, 2005 ; Noirot-Gros *et al.*, 2002 ; Noirot-Gros *et al.*, 2006).

La protéine Soj joue un rôle dans la sporulation et dans la partition des chromosomes lors de la division cellulaire (Murray *et al.*, 2008). Soj possède une activité ATPase et peut se lier à l'ADN. La protéine SpoOJ stimule l'activité ATPase de Soj et les deux protéines se lient sur le locus *parS* du chromosome bactérien à proximité de l'origine de réplication. La fixation du complexe sur *parS* provoque la séparation des deux *oriC* lors de la ségrégation des chromosomes (Lee et Grossman, 2006). Les analyses *in vivo* et *in vitro* suggèrent que Soj lié à l'ATP mais pas à l'ADN peut se lier à DnaA sur l'*oriC* indépendamment de SpoOJ. Le mécanisme proposé est le suivant : lors de l'initiation de la réplication, Soj se fixe sur l'ADN dans la région *parS-oriC*. Soj permettrait notamment de réduire la compaction de l'ADN dans la région *parS-oriC* facilitant l'action de DnaA. Après initiation, SpoOJ interagirait avec Soj en stimulant son activité ATPase et en réduisant sa capacité à lier l'ADN. La dissociation de Soj de l'ADN empêcherait un nouveau cycle d'initiation de la réplication par DnaA (Murray *et al.*, 2008 ; Ogura *et al.*, 2003 ; Veening *et al.*, 2009).

Néanmoins, ces mécanismes ne sont pas extrapolables à l'ensemble des bactéries car peu de ces protéines sont conservées. Il reste sans doute un grand nombre de régulateurs de la

réplication à découvrir qui sont peut-être spécifiques à des espèces bactériennes. C'est le cas par exemple de HobA qui forme un complexe avec le domaine I de DnaA dans la bactérie *H. pylori* (voir partie 1.4.2).

1.2.4. Structure et mode d'action de l'hélicase répllicative DnaB

Lors de la réplication de l'ADN, celui-ci doit être désapparié en simples brins pour servir de matrice à l'ADN polymérase qui va synthétiser les néo-brins. Les enzymes dédiées à cette fonction sont des hélicases répliquatives. Ces enzymes essentielles sont en général hexamériques. Chez les archées et les eucaryotes, les hélicases répliquatives appartiennent à la famille des protéines de maintenance des mini-chromosomes (MCM, voir partie 1.2.4.3.1). Chez les bactéries, cette fonction est assurée par DnaB qui utilise l'hydrolyse de l'ATP pour séparer les brins d'ADN. DnaB est une protéine strictement conservée chez les bactéries et certaines espèces ont même plusieurs allèles du gène.

1.2.4.1. Classification de DnaB dans la famille des hélicases

La liaison et l'hydrolyse de l'ATP se font via le domaine ATPase qui appartient à la classe des ASCE des protéines P-loop. On dénombre six superfamilles (SF) d'hélicases. Les superfamilles 3 à 6 (SF3 et SF6) regroupent les hélicases hexamériques alors que les superfamilles 1 et 2 (SF1 et SF2) regroupent les hélicases monomériques (Itsathitphaisarn *et al.*, 2012 ; Singleton *et al.*, 2007). Les superfamilles 3 à 6 sont subdivisées en deux groupes : les hélicases de type AAA+ qui se déplacent de 3' en 5' (SF3 et SF6) et celles de type RecA qui se déplacent (comme les hélicases des superfamilles 1 et 2) de 5' en 3' (SF4 et SF5). L'hélicase répllicative DnaB chez les bactéries à Gram négatif est une protéine d'environ 55 kDa qui est composée d'un domaine N-terminal, d'un domaine de liaison et d'un domaine C-terminal ATPase du type RecA. Elle est membre de la SF4 qui inclue également l'hélicase G40P du bactériophage SPP1 et l'hélicase du bactériophage T7 (T7gp4) (Kornberg et Baker, 1992).

1.2.4.2. Structure générale de l'hélicase DnaB

DnaB est composée de deux grands domaines, le domaine N-terminal (NTD) et le domaine C-terminal (CTD). Le NTD varie considérablement au niveau de la séquence primaire (Weigelt *et al.*, 1999). Il est divisé en deux sous-domaines : une tête globulaire composée de six hélices α , et un sous-domaine dans lequel deux hélices forment une épingle à cheveux α (Bailey *et al.*, 2007a). La délétion du NTD de DnaB d'*E. coli* (EcDnaB) et de *Bacillus stearothermophilus* (BstDnaB) provoque une perte d'activité hélicase. Cependant, ces mutants conservent leur activité ATPase (Bird *et al.*, 2000 ; Biswas *et al.*, 1994 ; Biswas et Biswas, 1999). Ce constat ne semble pas s'appliquer à toutes les hélicases de la famille puisqu'un mutant équivalent de la protéine G40P du phage SS1 conserve les deux activités (Mesa *et al.*, 2005). La structure de GkDnaC (DnaB chez les bactéries à Gram positif) en complexe avec un oligonucléotide simple brin montre que le NTD de l'hélicase possède également un site de fixation à l'ADNsb (Lo *et al.*, 2009).

Le domaine C-terminal (CTD) dont la séquence est beaucoup plus conservée, adopte un repliement typique du sous-groupe RecA de la famille ASCE des protéines P-loop et contient les cinq motifs conservés (Bailey *et al.*, 2007a). Quatre de ces motifs forment le cœur du feuillet β (H1=Walker A, H1a, H2=Walker B, H3). Le cinquième motif (H4) (Ilyina *et al.*, 1992) et les boucles majoritaires du CTD (boucles 1, 2, 3) sont constituées d'acides aminés impliqués dans la liaison à l'ADNsb. Le domaine CTD contient enfin trois hélices formant une crevasse appelée poche d'hexamérisation. Ce site est conservé dans toutes les hélicases de la famille SF4 (Figure 8).

Le domaine de liaison est une région de la protéine située entre le domaine NTD et le CTD. Il est constitué d'une boucle flexible avec en son centre une hélice α .

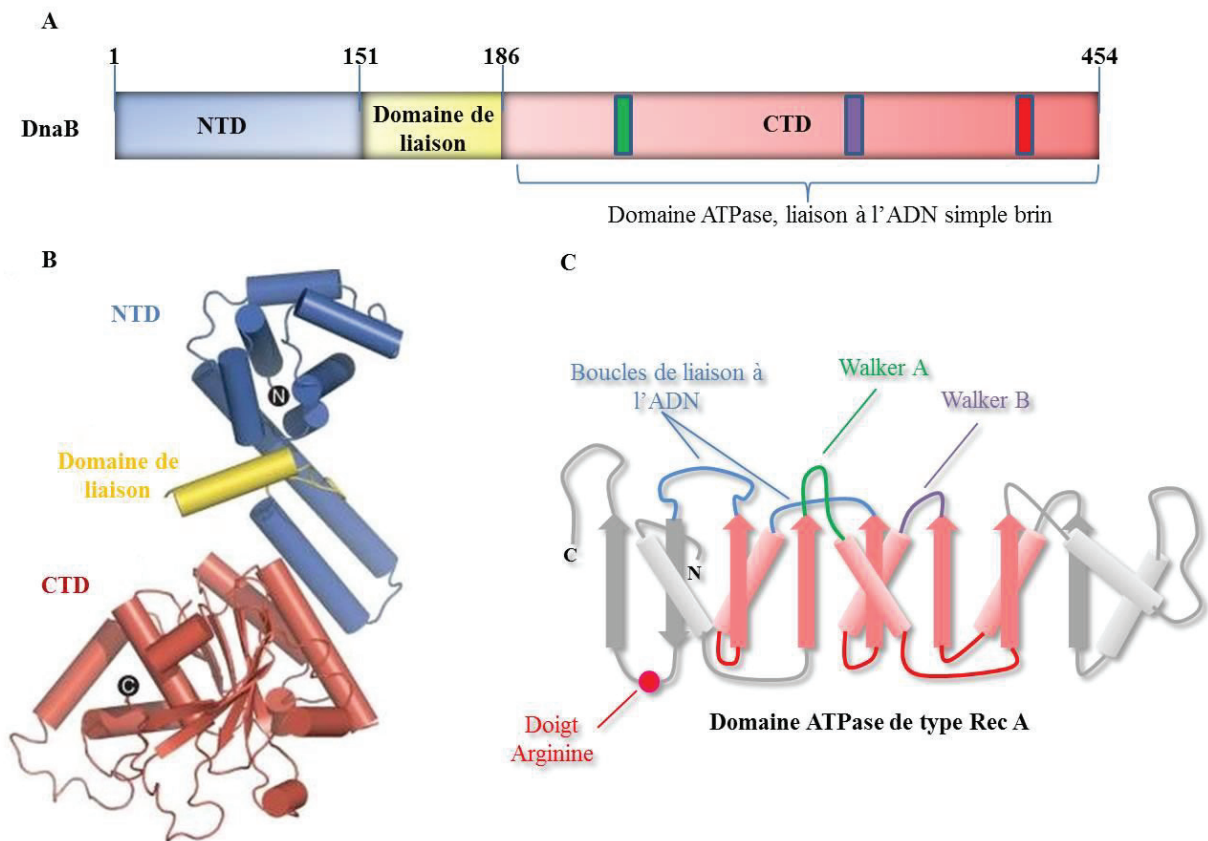
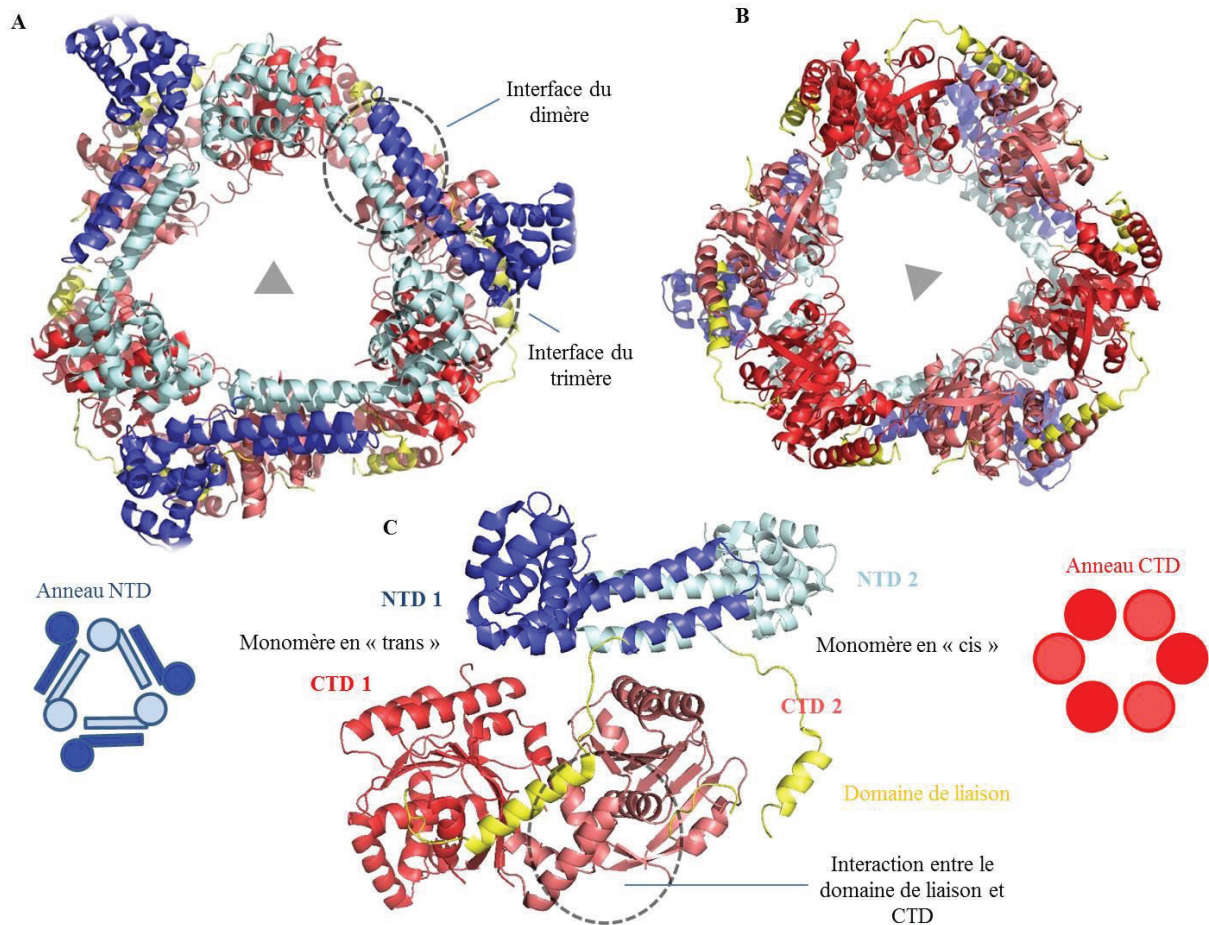


Figure 8 : (A) Schéma représentant les différents domaines de DnaB. (B) Structure cristalline du monomère de DnaB chez *T. aquaticus* (PDB : 2Q6T ; Bailey *et al.*, 2007). (C) Schéma représentant l'organisation du domaine ATPase de type RecA. Les motifs Walker A et B sont colorés en vert et en violet respectivement et le doigt arginine en rose. Les boucles de liaison à l'ADN sont colorées en bleu, et en rouge est représenté le repliement structural commun du domaine ATPase des protéines de la famille ASCE.

Les études en ME d'EcDnaB ont montré que les hexamères forment des anneaux qui peuvent adopter une symétrie C3, C6 ou un intermédiaire C6-C3 (Nunez-Ramirez *et al.*, 2006 ; Yang *et al.* 2002). Les structures cristallines des hélicases G40p du bactériophage SSP1, BstDnaB et DnaC de *Geobacillus kaustophilus* (GkDnaC), ont permis d'élucider les différentes organisations observées en ME. En 2007, l'équipe du Professeur Steitz a publié deux structures cristallographiques (B1 et B2) de BstDnaB sous forme hexamérique. Ces structures montrent que DnaB est composée de deux anneaux constitués respectivement des domaines NTD et CTD qui ont des symétries différentes. Dans l'anneau NTD, deux domaines NTD s'assemblent en dimère, et trois dimères s'associent ensuite en trimère générant ainsi une symétrie C3. Cette conformation de l'anneau NTD est appelée « dilatée ». L'anneau CTD est irrégulier dans la forme cristalline B1 mais adopte une symétrie C3 dans la forme B2, illustrant la flexibilité de l'enzyme (Bailey *et al.*, 2007b) (Figure 9). On remarque que l'interaction entre les deux anneaux est faible avec une interface de moins de 1500 Å².

L'association entre les anneaux NTD et CTD se fait par l'intermédiaire des domaines de liaison de chaque sous-unité. L'hélice α du domaine de liaison d'une sous-unité s'insère dans la poche d'hexamérisation du CTD voisin et permet de maintenir et de coordonner les sous-unités entre elles (Bailey *et al.*, 2007b).



1.2.4.3. Translocation de DnaB le long de l'ADN

Récemment, l'équipe du Professeur Steitz a publié la structure de BstDnaB en complexe avec le GDP-AlF₄ et un oligonucléotide de 18 pb (BstDnaB•AlF₄•dT18) (Itsathitphaisarn *et al.*, 2012). Le groupement AlF₄ permet de mimer l'état de transition de GTP en GDP lors de son hydrolyse par l'enzyme. Dans cette structure, la planarité des anneaux NTD et CTD est rompue et BstDnaB forme une spirale droite. Le NTD conserve une conformation dilatée mais forme un escalier en colimaçon avec un dimère par « marche ».

L'organisation du CTD est également modifiée et suit la spirale formée par les NTD mais avec un domaine ATPase par « marche ». La présence d'ADNsb induit la rotation des CTD et permet la fixation du GDP-AlF₄ entre deux sous-unités de la protéine. La liaison de l'ADN entraîne une réorganisation des boucles de l'anneau central qui interagissent principalement avec le squelette phosphate. Le domaine de liaison maintient les anneaux NTD et CTD ensemble car très peu de contacts sont observés entre les deux domaines d'un même monomère. Une conséquence importante de cette organisation est la présence d'une ouverture dans l'anneau qui ressemble à un écrou de verrouillage (« lock-washer »). Cette ouverture suggère qu'une forme spiralée de l'anneau permettrait le passage de l'ADNsb de l'extérieur vers l'intérieur de l'enzyme (Figure 10).

L'analyse de la structure suggère un mécanisme de translocation le long de l'ADN. Dans la structure du complexe BstDnaB•AlF₄•dT18, chaque sous-unité de DnaB est liée à deux nucléotides. DnaB se déplacerait donc de deux nucléotides lors de l'hydrolyse de l'ATP du domaine CTD en 5' de l'ADN. En effet, lors de l'hydrolyse de cet ATP, le domaine CTD, situé à l'extrémité 5' de l'ADN, se dissocierait de celui-ci et migrerait vers l'extrémité 3'. Une nouvelle molécule d'ATP viendrait alors se fixer, lui permettant ainsi d'interagir de nouveau avec l'ADN. Au niveau de l'anneau NTD, les dimères ne peuvent pas se dissocier. L'anneau NTD tournerait donc tous les quatre nucléotides, soit quand deux CTD ont effectué une rotation de 5' en 3' (Itsathitphaisarn *et al.*, 2012).

Ce modèle de translocation contraste avec les hélicases hexamériques de la famille SF3 et SF5 qui restent planaires durant la translocation. En effet chez ces dernières, les boucles de liaison à l'ADN forment une spirale au centre qui se lie à l'ADN et accompagne ses mouvements.

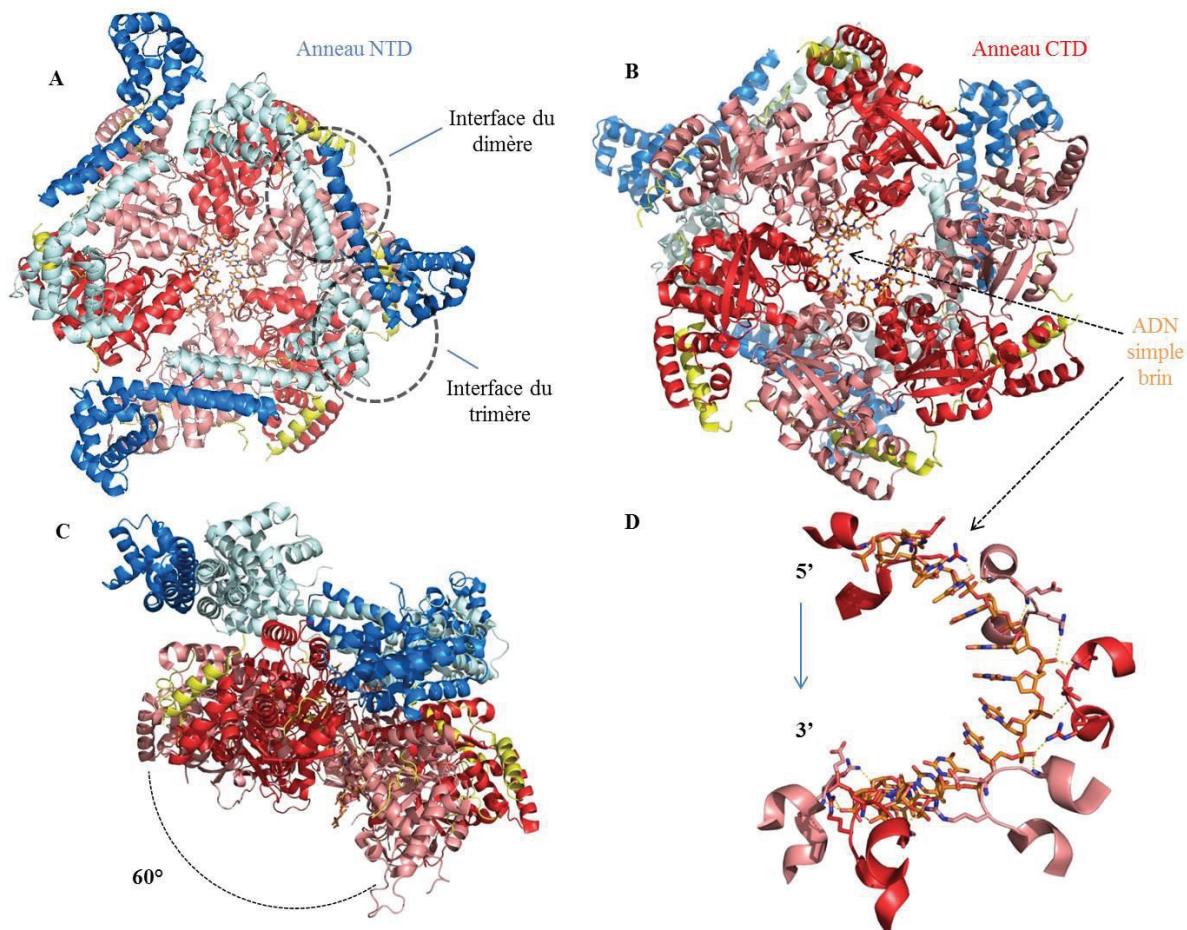


Figure 10 : (A) Structure de BstDnaB•AlF₄•dT18 vue du dessus, les domaines NTD formant un trimère de dimères sont représentés en bleu. (B) Vue de dessous, les domaines CTD sont représentés en rouge et l'ADNs en orange. (C) Vue de côté, montrant la forme spiralée prise par les anneaux NTD et CTD indiquée par la ligne en pointillés noirs (D) Agrandissement de la zone d'interaction entre l'ADNs et la boucle de liaison des domaines CTD (PDB : 4ESV ; Itsathitphaisarn *et al.*, 2012).

La structure d'AaDnaB en complexe avec l'ADP résolue récemment (Strycharska *et al.*, 2013) montre à quel point la structure de l'hexamère est flexible en fonction de l'élément avec lequel il interagit. Dans cette structure, l'anneau NTD adopte une conformation complètement différente. Chaque dimère du NTD a subi une rotation de 40° projetant une tête globulaire de chaque dimère de NTD vers l'extérieur et l'autre vers l'intérieur de l'hexamère dans une conformation dite « resserrée ». Comparée à celle dilatée, la conformation resserrée réduit considérablement le diamètre du canal central au niveau de l'anneau NTD. L'anneau CTD possède une véritable symétrie C₆ planaire dans laquelle chaque sous-unité est orientée comme dans la structure du complexe BstDnaB•AlF₄•dT18 (Figure 11). Des études en ME à coloration négative et par diffusion des rayons X aux petits angles (SAXS) montrent qu'EcDnaB adopte seulement une conformation dilatée. En revanche, lorsqu'EcDnaB est incubée avec de l'ATP ou des analogues non hydrolysables, la moitié de la protéine prend une

conformation resserrée. L'ajout d'ADNsb en absence d'ATP n'a aucun effet sur la conformation de l'anneau NTD (Strycharska *et al.*, 2013). Ce réarrangement structural en présence de nucléotides a été également observé avec les protéines AaDnaB et BstDnaB.

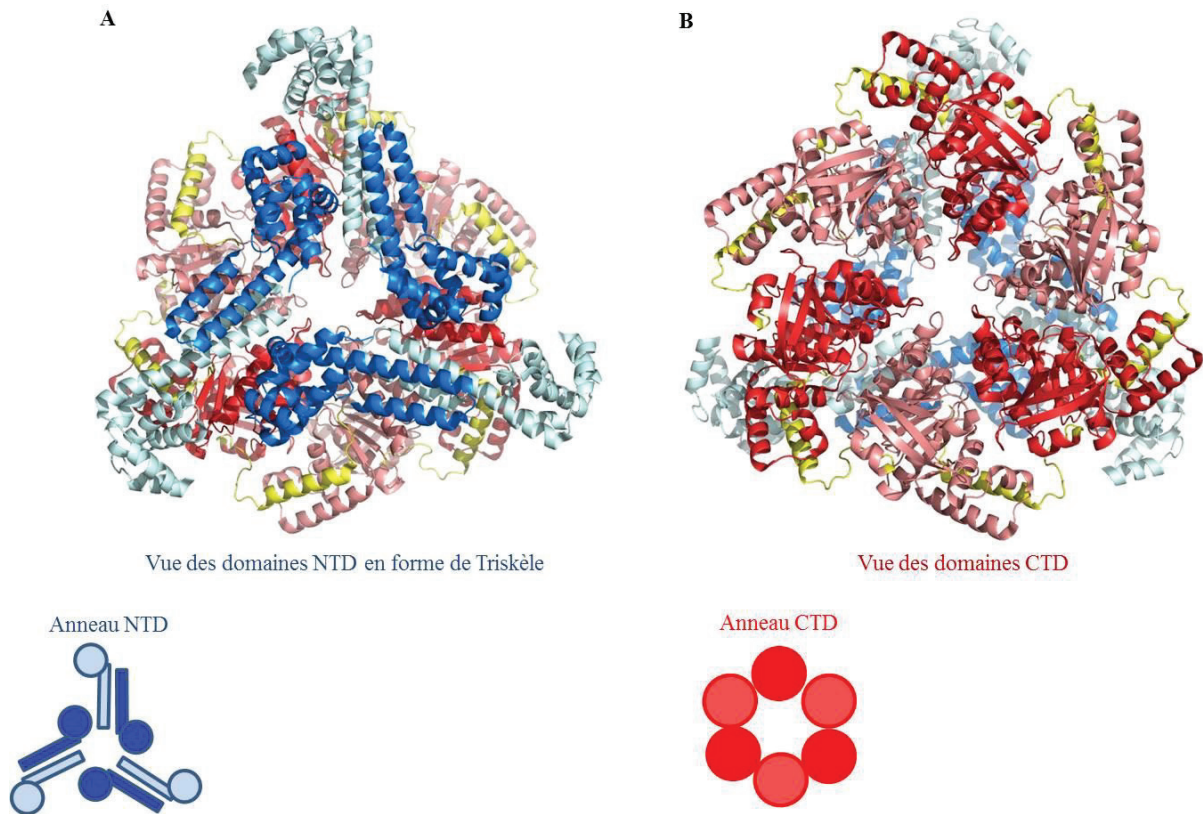


Figure 11 : Structure d'AaDnaB en complexe avec l'ADP. (A) A gauche, vue de dessus, les six domaines NTD s'organisent en « triskèle ». (B) A droite, vue des six domaines CTD qui adoptent une symétrie C6 (PDB : 4NMN, Strycharska *et al.*, 2013).

1.2.4.4. Chargement de l'hélicase sur la fourche de réplication

Tous les organismes vivants utilisent une hélicase pour la réplication, accompagnée ou non (dans le cas des virus) d'une protéine initiatrice et de facteurs de chargement de l'hélicase (Tableau 1). Néanmoins, ces protéines ne présentent que peu d'homologie de séquences et ont des mécanismes de chargement différents. Les exemples que nous allons donner ici illustrent cette diversité.

Protéine	Bactérie	Virus SV40	Eucaryote	Archée
Initiatrice	DnaA	LTag	ORC	Orc1/Cdc6
Hélicase	DnaB (DnaC chez Gram +)	LTag	Mcm 2-7	MCM
Facteur de chargement	DnaC (DnaI chez Gram +)	LTag	Cdc6/Cdt1	Orc1/Cdc6 ?

Tableau 1 : Comparaison des protéines impliquées dans l'initiation de la réplication chez les bactéries, les virus, les archées et les eucaryotes.

1.2.4.4.1. « Ring opening » : ouverture de l'hélicase et chargement sur l'ADNsb

- *Structure et rôle du complexe DnaB•DnaC des bactéries A. aeolicus et E. coli*

Si la structure de l'hexamère de DnaB est flexible, son architecture circulaire crée un problème topologique durant le chargement sur l'ADNsb. L'anneau formé par l'hexamère doit être ouvert pour permettre le passage de l'ADNsb dans le canal central alors que le second brin restera à l'extérieur de l'anneau. De plus, DnaB ne se déplaçant que dans le sens 5' vers 3', il est nécessaire d'orienter correctement l'hélicase sur chaque brin de la fourche de réplication.

Après l'ouverture de l'origine de réplication, quelques bactéries utilisent donc des protéines spécialisées appelées facteurs de chargement pour placer l'hélicase sur l'ADNsb. Les bactéries à Gram négatifs telles qu'*A. aeolicus* et *E. coli* codent pour le facteur de chargement DnaC (Koonin, 1992). Certaines bactéries à Gram positif (comme *B. subtilis*) codent pour une protéine homologue nommée DnaI (Soulтанas 2002). DnaC et DnaI sont composées de deux domaines : un domaine N-terminal, différent chez les deux protéines qui se lie à l'hélicase et un domaine C-terminal conservé, qui a un repliement typique des protéines AAA⁺ (Ioannou *et al.*, 2006 ; Loscha *et al.*, 2009 ; Mott *et al.*, 2008 ; Tsai *et al.*, 2009). La structure de DnaC d'*A. aeolicus* montre que le domaine AAA⁺ adopte une structure similaire à celle du domaine IIIB de DnaA (Mott *et al.*, 2008). Comme DnaA, DnaC se lie, peut hydrolyser l'ATP et s'oligomériser en formant un mini filament hélicial.

La structure et le rôle du complexe DnaB•DnaC sont longtemps restés inconnus. En 2013, l'équipe du Professeur Berger a résolu la structure à une résolution de 25 Å du complexe DnaB•DnaC de la bactérie *E. coli* par cryo-ME (Arias-Palomo *et al.*, 2013). L'enveloppe montre que l'assemblage est un dodécamère DnaB₆•DnaC₆ et comprend trois parties distinctes (Figure 12). La première partie correspond à l'anneau NTD de DnaB et la seconde à l'anneau CTD de DnaB, l'ensemble formant une hélice similaire à celle de BstDnaB•AlF₄•dT18. La troisième partie est composée d'un mini filament hélical de six monomères de DnaC. En replaçant les différentes structures cristallines individuelles de BstDnaB et AaDnaC, un modèle quasi atomique a pu être reconstruit, à l'exception du domaine N-terminal de DnaC pour lequel il n'existe aucune structure cristalline. Il est intéressant de noter que l'anneau NTD adopte la conformation resserrée mais cette fois sous forme spiralée. Dans le complexe, la continuité des anneaux de DnaB est brisée comme dans la structure de BstDnaB•AlF₄•dT18. Cette observation suggère que l'interaction avec DnaC induirait la forme spiralée de DnaB et permettrait l'ouverture des anneaux. D'un point de vue fonctionnel, ce changement structural permettrait le passage de l'ADNsb à l'intérieur de DnaB lors du chargement de DnaB (Arias-Palomo *et al.*, 2013) (Figure 12). De manière intéressante, une structure en spirale similaire a été observée avec le facteur de chargement de la pince β (« β-clamp loader ») (Kelch *et al.*, 2011, 2012). Cette protéine est pentamérique et possède également un domaine AAA+. En présence d'ATP, le pentamère prend une forme spiralée et permet le chargement de la pince β sur l'ADNdb. (Bowman *et al.*, 2004; Kelch *et al.*, 2011; Miyata *et al.*, 2005; Seybert *et al.*, 2006; Simonetta *et al.*, 2009). L'ATP fixé sur le facteur de chargement est ensuite hydrolysé, ce qui libère la pince β du complexe (revue dans Pomerantz et O'Donnell, 2007). On peut donc penser qu'un mécanisme similaire d'hydrolyse de l'ATP permettrait de séparer DnaC de DnaB et d'achever le chargement de cette dernière.

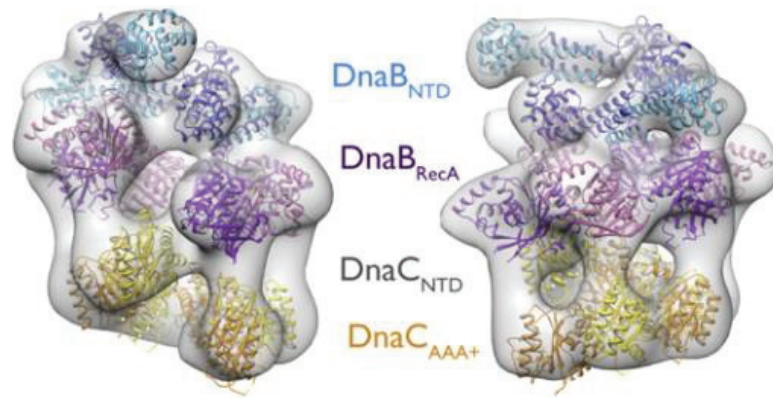


Figure 12 : Enveloppe de ME du complexe DnaB₆DnaC₆. Les structures des différents domaines ont été placées à l'intérieur de l'enveloppe pour obtenir un modèle quasi atomique (Arias-Palomo *et al.*, 2013).

Chez *A. aeolicus*, DnaA chargée en ATP, forme un filament sur l'ADN capable d'ouvrir la région riche en AT de l'*oriC* en simple brin. Les études récentes montrent que le filament de DnaA vient s'envelopper autour du simple brin supérieur (brin 5') par l'intermédiaire de son domaine III ($\alpha3/\alpha4$) et de deux hélices du domaine IV ($\alpha5/\alpha6$) (Duderstadt *et al.*, 2011). Dans le modèle proposé par Mott *et al.*, les deux hexamères de l'hélicase seraient sous forme de complexe DnaB₆•DnaC₆. Le premier serait positionné par interaction entre DnaC et DnaA enroulé autour du brin 5' du DUE. Plus précisément, le doigt arginine du domaine AAA⁺ de DnaC viendrait interagir avec l'ATP du domaine III de DnaA. Sur le brin 3' du DUE, l'hélicase doit être chargée dans le sens opposé. Le domaine I de DnaA interagirait avec le NTD de DnaB qui reste accessible dans le complexe DnaB₆•DnaC₆ (Ludlam *et al.*, 2001 ; Seitz *et al.*, 2000 ; Sutton *et al.*, 1998 ; Sultanas, 2012) (Figure 13).

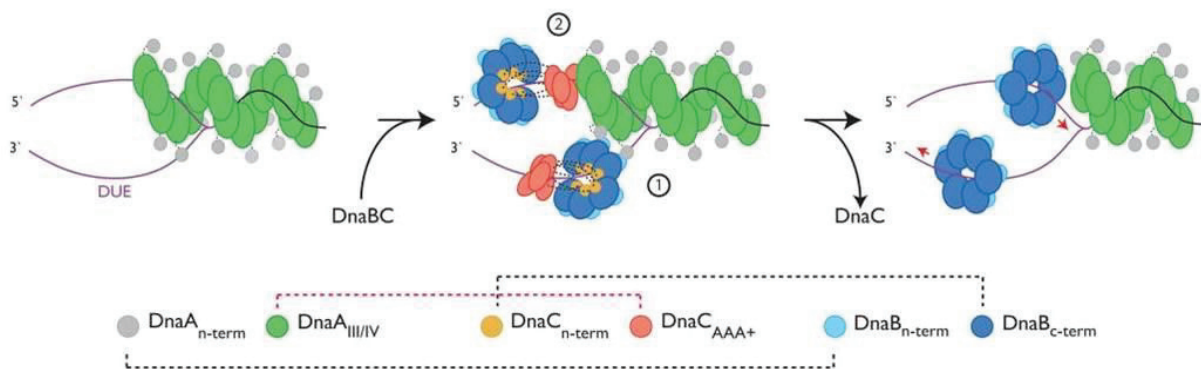


Figure 13 : Schéma de l'initiation de la réplication chez *A. aeolicus* (Mott *et al.*, 2008). L'oligomérisation de DnaA autour de l'ADNs_b provoque l'ouverture de la région riche en AT (DUE). L'interaction entre DnaC et les domaines III-IV de DnaA place le premier hexamère de DnaB sur le premier brin et l'interaction entre le NTD de DnaB et les domaines I-II de DnaA place le second hexamère sur le deuxième brin.

Chez *E. coli*, le mécanisme est plus complexe du fait de l'intervention de la protéine IHF qui se fixe entre les boîtes R1 et R5 et interrompt le filament de DnaA. Un autre modèle de chargement de l'hélicase a été proposé ; il prend en compte le rôle d'IHF et n'implique pas d'interaction directe entre DnaA et DnaC (Ozaki et Katayama, 2012). Dans ce modèle, la fixation d'IHF provoquerait un renversement de 160° de l'ADN créant une courbure d'*oriC* (Rice *et al.*, 1996). La fixation de DnaA sur les boîtes R1 et R5 et la présence de IHF suffiraient à ouvrir la région AT et à permettre l'enroulement de DnaA autour du brin riche en T (Hwang et Kornberg, 1992 ; revue dans Ozaki et Katayama, 2012 ; Roth *et al.*, 1994). Le premier hexamère de DnaB serait alors chargé sur l'ADNsb par interaction entre le domaine I de DnaA fixé sur le site R1 et le NTD de DnaB du complexe DnaB₆•DnaC₆ (Marszalek et Kaguni, 1994 ; Ozaki et Katayama, 2012 ; Seitz *et al.*, 2000). Ce premier hexamère commencerait à dérouler l'ADNdb de 5' vers 3' et créerait ainsi une plus grande ouverture de la région DUE. Le second hexamère pourrait alors se charger sur le second brin d'ADN, toujours par l'intermédiaire de DnaB₆•DnaC₆ (Ozaki et Katayama, 2012 ; Soutanas, 2012).

Cependant DnaC et DnaI n'ont pas été identifiées chez toutes les bactéries. DnaC est présente chez les génomes des entérobactéries et des *Aquificiae* et DnaI uniquement chez les firmicutes (Robinson *et al.*, 2012). Ceci implique que les bactéries n'utilisent pas toutes un facteur de chargement de l'hélicase de type AAA+ pour l'initiation de la réplication. D'autres protéines, qui restent à identifier, pourraient accomplir cette fonction. Cependant, chez *Pseudomonas putida* et *P. aeruginosa*, DnaB peut être chargée *in vitro* sur une origine de réplication d'un plasmide par interaction avec DnaA en l'absence de DnaC (Caspi *et al.*, 2001). Ceci suggère que la présence d'un facteur de chargement n'est pas forcément nécessaire. Enfin, l'utilisation de complexes analogues à celui du redémarrage de la réplication (PriA/PriB/DnaT) d'*E. coli* est encore une autre possibilité (Lopper *et al.*, 2007).

- *Geobacillus kautrophilus*

Chez certains Bacilles comme *G. kautrophilus*, le facteur de chargement DnaI lie l'ATP sous forme de monomère. La protéine s'oligomériserait en hexamère autour de l'ADNsb formant un anneau fermé de DnaI. Les monomères se positionneraient autour du simple brin avec leur domaine CTD et NTD en 3' et 5' de l'ADN respectivement (Tsai *et al.*, 2009). Le recrutement de l'hélicase se produirait par interaction entre le CTD de l'hélicase et le NTD de DnaI. Cette interaction provoque l'ouverture de l'hélicase et permet le passage de l'ADNsb dans le canal central de l'enzyme. L'interaction avec l'ADN activerait le site de

liaison à l'ATP par rotation des domaines CTD de l'hélicase et induirait la fermeture de l'hexamère (Tsai *et al.*, 2009) (Figure 14A).

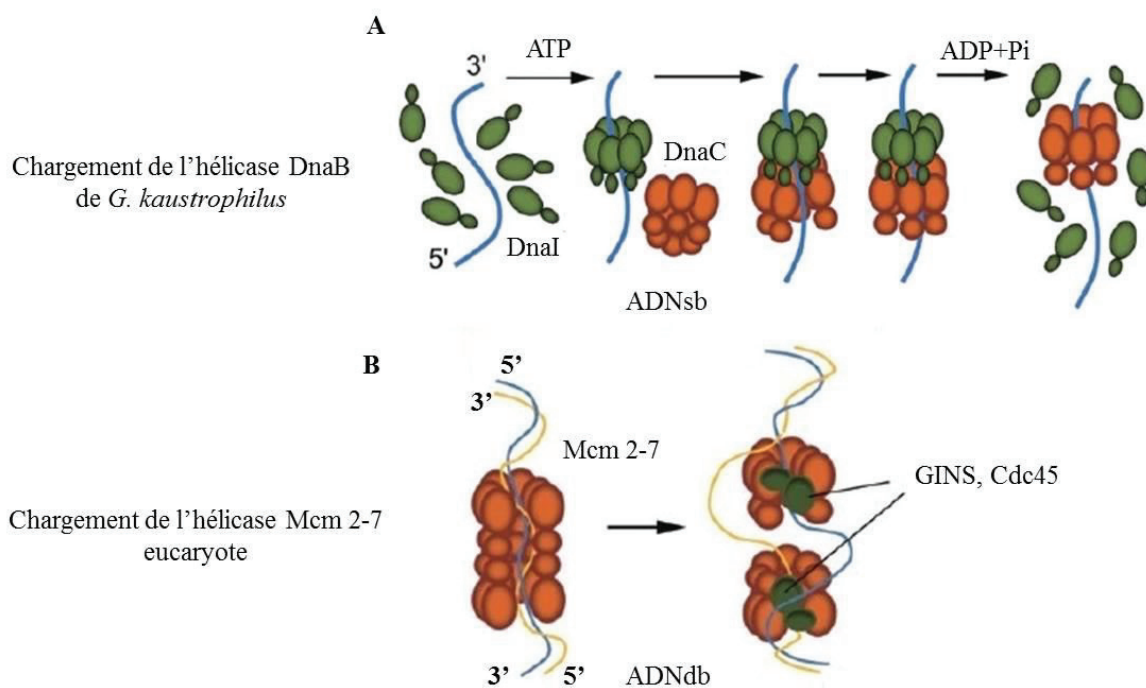
- Archées

Les hélicases des archées appartiennent à la famille des protéines dites « mini chromosome maintenance » (ou MCM). Les MCM possèdent, elles aussi, un domaine NTD et CTD. La structure du domaine NTD montre que ce dernier forme un anneau avec une symétrie C6 qui peut se lier à l'ADN simple ou double brin (Fletcher *et al.*, 2003). Le domaine CTD contient le domaine AAA⁺ qui lie et hydrolyse l'ATP. Il forme également un anneau avec une symétrie C6 (Brewster *et al.*, 2008 ; Pape *et al.*, 2003). Différentes études par cristallographie aux rayons X et par ME montrent que les MCM forment des hexamères (Pape *et al.*, 2003), des dodécamères ou encore des doubles heptamères (Costa *et al.*, 2006a). Cependant, les rôles précis de ces trois formes restent encore inconnus à ce jour. Le chargement des hélicases MCM sur l'origine de réplication des archées n'a pas encore pu être démontré. Cependant le facteur de chargement des MCM serait la protéine Orc1. En effet, plusieurs d'origines de réplication d'archées possèdent deux sites de liaison à Orc1 inversés. Ceux-ci étant séparés d'une distance suffisante, le chargement de MCM en double hexamère serait possible (revue dans Costa *et al.*, 2013 ; Gaudier *et al.*, 2007 ; Robinson *et al.*, 2004).

- Eucaryotes

Chez les eucaryotes, l'hélicase appartient également à la famille des MCM et forme des hexamères analogues aux MCM des archées (Costa *et al.*, 2011 ; Pape *et al.*, 2003). Cependant, six sous-unités différentes, appelées Mcm-2 à 7 (Mcm 2-7), s'assemblent en hétéro-hexamère qui se déplace de 3' en 5' sur le brin avancé. L'hélicase Mcm 2-7 diffère également de l'hélicase bactérienne en plusieurs points. Les NTD ont des structures différentes de ceux de DnaB. Les domaines ATPase de chaque sous-unité ne possèdent pas de repliement de type RecA mais sont des domaines AAA⁺. Une structure en ME montre que l'hexamère de Mcm 2-7 de *Drosophila melanogaster* adopte une conformation ouverte en spirale en l'absence de substrat et forme une conformation fermée et planaire quand elle se lie à l'ATP et à ses partenaires protéiques (GINS/Cdc45) (Costa *et al.*, 2011). Le même phénomène a été observé chez l'hélicase de l'eucaryote minimaliste *Encephalitozoon cuniculi* (EcMcm 2-7). Des études en ME et en SAXS de EcMcm 2-7 indiquent que l'hélicase adopte spontanément une structure d'hexamère en spirale ouverte. Les auteurs ont également montré que l'ajout d'ATP entraînait une conformation plus planaire de l'hexamère mais qui

n'atteignait pas une fermeture complète (Lyubimov *et al.*, 2012). Le chargement de l'hélicase sur l'ADN est possible grâce à l'action conjuguée des protéines ORC, Cdc6 et Cdt1 qui forment le complexe de pré-initiation (pre-RC). Cdt1 interagissait avec Mcm 6 et l'interaction du complexe entre ORC et Cdc6 en présence d'ATP permettrait le chargement de deux hexamères tête-tête sur l'ADNdb (Evrin *et al.*, 2009 ; Remus *et al.*, 2009). L'autre modèle possible serait que Mcm 2-7 soit chargée sur l'ADNdb sous forme de dodécamère par interaction avec le pre-RC (Bochman *et al.*, 2008). Dans les étapes suivantes, les deux hypothèses convergent en proposant que les deux protéines GINS et Cdc45 interagissent avec Mcm 2-7 ce qui provoque la fermeture du dodécamère (Figure 14B).



Ouverture de l'hélicase sur l'ADN simple ou double brin

Figure 14 : (A) Chargement de l'hélicase sur l'ADNsb chez *G. kaustrophilus*. Le facteur de chargement DnaI en complexe avec l'ATP s'hexamérise autour de l'ADNsb. Puis, le domaine NTD de DnaI interagit avec le domaine CTD de l'hélicase DnaC ce qui provoque l'ouverture de l'hexamère et permet le passage de l'ADNsb dans l'hélicase. (B) Schéma du chargement de l'hélicase Mcm 2-7 eucaryote. Le dodécamère Mcm 2-7 est chargé sur l'ADNsb par le complexe ORC•Cdc6 (non représenté) puis dissocié par les protéines GINS et Cdc45 (Soultanas, 2012).

1.2.4.4.2. « Ring making » : Formation de l'hexamère sur l'origine de réplication

Le « ring making » est un mécanisme dans lequel l'hexamère est directement assemblé *de novo* sur l'origine. C'est le cas pour *B. subtilis* ou encore chez certains virus tels que le SV40. Il existe également des hélicases qui se chargent sur l'origine de réplication sans facteur de chargement.

- *Bacillus subtilis*

La bactérie *B. subtilis* semble utiliser un autre mécanisme que *G. kaustrophilus*. L'hélicase appelée DnaC, peut former des hexamères en solution mais ceux-ci sont inactifs (Velten *et al.*, 2003). L'hexamère ne serait actif qu'une fois positionné sur l'ADN par son facteur de chargement DnaI. Pour ce faire, des complexes DnaI₁•DnaC₁ s'assembleraient autour de l'ADNsb après ouverture de l'origine de réplication (Soultanas, 2012 ; Velten *et al.*, 2003). On note également l'existence d'une autre protéine, appelée DnaB (à ne pas confondre avec l'hélicase) qui s'associe au complexe DnaI₁•DnaC₁ et participe à la formation du primosome.

- SV40 LTag

Une fois dans la cellule eucaryote le virus SV40 produit la protéine LTag qui joue les rôles de protéine initiatrice et d'hélicase. LTag est composée des domaines NTD, CTD (ATPase), et d'un domaine supplémentaire appelé OBD (Origin Binding Domain). OBD permet de reconnaître spécifiquement l'origine de réplication virale grâce à quatre séquences pentanucléotidiques (PEN) (Arthur *et al.*, 1988 ; Luo *et al.*, 1996). Deux monomères de LTag vont se fixer sur les deux séquences de chaque côté de l'origine. Il y a alors formation de deux dimères qui se font face et s'oligomérisent en deux hexamères de chaque côté. Chaque moitié de l'origine de réplication permet donc la formation d'un hexamère. Les deux hexamères interagissent tête-tête au niveau des domaines OBD ; on parle alors de dodécamère (Gai *et al.*, 2004 ; Li *et al.*, 2003 ; Valle *et al.*, 2006). Néanmoins, la structure cristalline du dimère de LTag en complexe avec l'origine virale a montré que seul le domaine OBD du premier monomère était lié à la séquence PEN1 (Chang *et al.* 2013). De plus, des études biochimiques ont établi que seules PEN1 et PEN3 étaient indispensables pour l'assemblage du dodécamère (Joo *et al.*, 1998 ; Sreekumar *et al.*, 2000). D'autres études sont donc nécessaires pour connaître les rôles exacts de PEN2 et PEN4. Une fois le dodécamère assemblé, la torsion

appliquée sur l'ADN par les CTD de LTag provoque la séparation du double brin. Il semblerait que l'ADNsb ne reste pas dans le canal central de l'hélicase mais passe par des pores latéraux chargés positivement (Li *et al.*, 2003 ; Souttanas, 2012) (Figure 15).

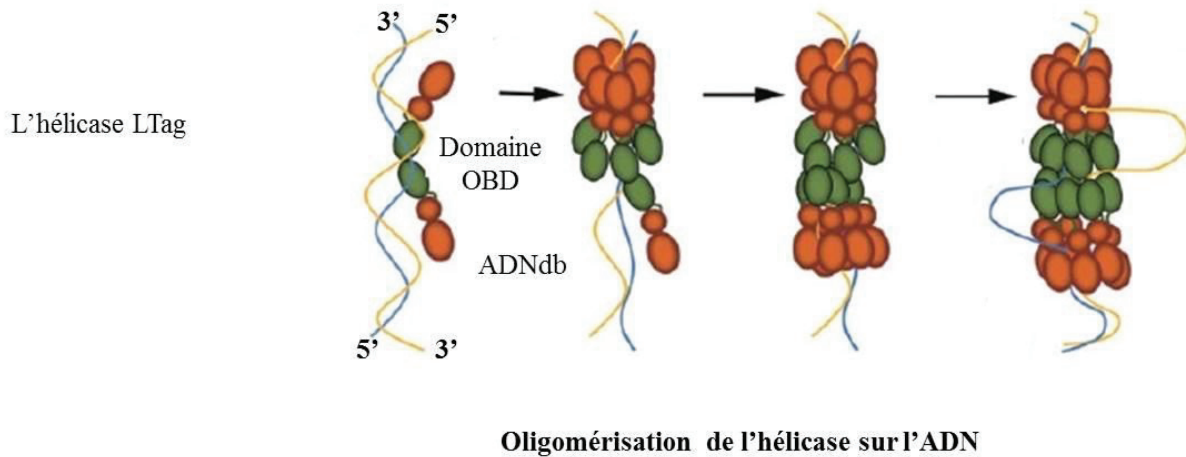


Figure 15 : Schéma illustrant le chargement de l'hélicase LTag. Les monomères se fixent sur l'origine de réplication et s'oligomérisent en deux hexamères tête-tête interagissant par leurs NTD. Cette oligomérisation se faisant autour de l'ADNdb, celui-ci est ensuite séparé en deux ADNsb par la torsion créée par le dodécamère sur l'ADN et par l'hydrolyse de l'ATP réalisée par les CTD qui « tirent » l'ADN dans des directions opposées (Souttanas, 2012).

1.2.4.4.3. Chargement de l'hélicase sans facteur de chargement

- *Les hélicases T7gp4 et Twinkle*

Les hélicases Twinkle et T7gp4 hexamériques sont produites respectivement par la mitochondrie et par le bactériophage T7. Leur particularité est de regrouper les activités hélicase (assurée par le domaine CTD) et primase (domaine NTD). Le domaine NTD (primase) de T7gp4 est responsable de l'interaction avec l'ADNsb et permet de positionner la protéine sur celui-ci (Farge *et al.*, 2008 ; Toth *et al.*, 2003). La fixation du domaine NTD (primase) sur l'ADN induit un changement de conformation général de la protéine et ouvre l'hexamère de T7gp4. L'ADN peut alors glisser à l'intérieur de la cavité centrale formée par les domaines CTD. Twinkle est capable de se charger seule sur l'ADN simple circulaire et pourrait utiliser un mécanisme similaire de celui de T7gp4 (Jemt *et al.*, 2011)

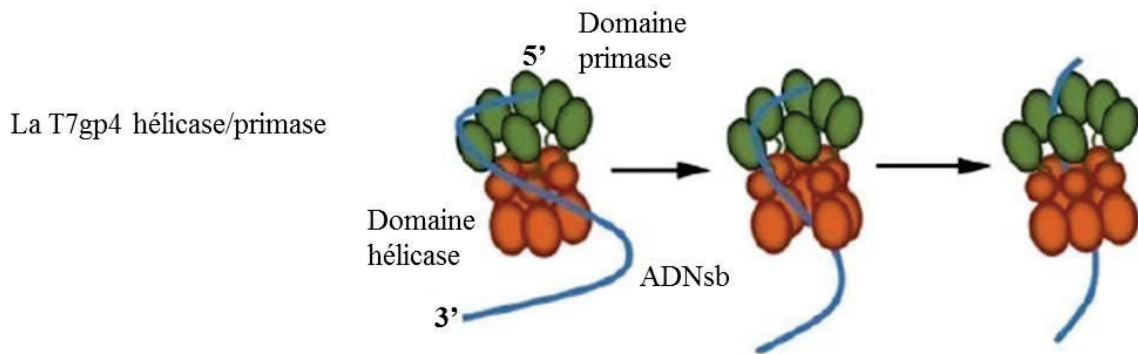


Figure 16 : Schéma du chargement de l'hélicase/primase T7gp4. L'ADNsb interagit avec le domaine primase provoquant l'ouverture du domaine hélicase, permettant ainsi le passage de l'ADN dans le canal central. La liaison entre l'ADN et le domaine hélicase induit la fermeture de l'anneau (Soultanas, 2012).

1.2.5. Formation du primosome et démarrage de la réplication

Chez les bactéries, la fonction de synthèse des amorces d'ARN nécessaire à la polymérase est assurée par la primase DnaG (Kornberg et Baker, 1992). Cette protéine agit en concert avec DnaB formant un complexe DnaB•DnaG appelé primosome. Pendant que DnaB déroule l'ADNdb, DnaG utilise l'ADNsb pour la synthèse des amorces d'ARN sur le brin retard. DnaG est une protéine qui possède trois domaines : un domaine ZBD (zinc binding domain) liant un atome de zinc, un domaine impliqué dans la synthèse des amorces d'ARN ou RPD (RNA polymerase domain) et un domaine se liant à l'hélicase ou HBD (helicase binding domain) localisé à l'extrémité C-terminale de la protéine (Tougu *et al.*, 1994) (Figure 17).

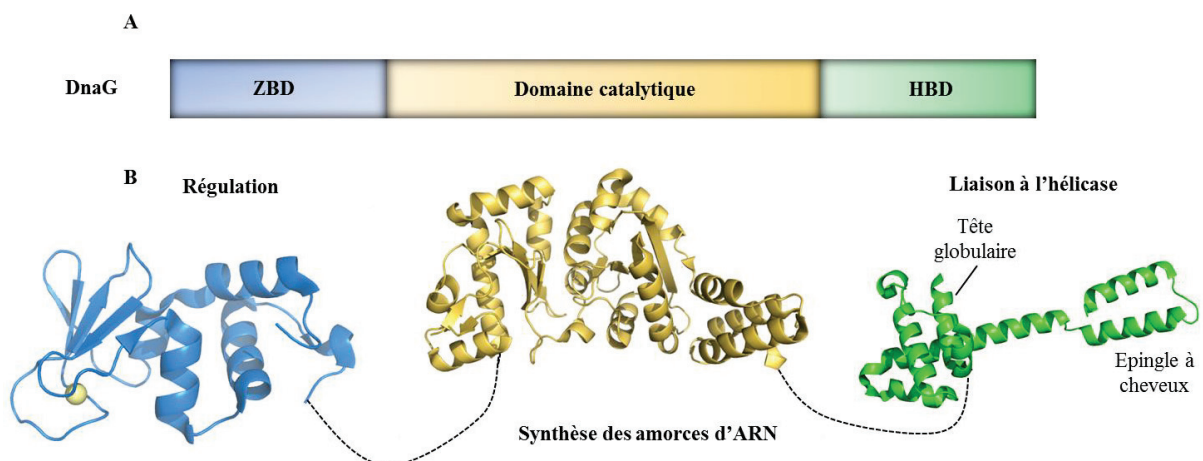


Figure 17 : (A) Schéma représentant les domaines de DnaG. (B) Structure du domaine ZBD de BstDnaG (PDB : 1D0Q, Pan et Wigley, 2000), du domaine RPD de EcDnaG (PDB : 1DDE, Keck *et al.*, 2000) et du domaine HBD d'EcDnaG (PDB : 1T3W, Oakley *et al.*, 2005).

Le domaine HBD adopte un repliement similaire à celui du NTD de DnaB : une tête globulaire, composée de six hélices α (sous-domaine C1), suivie de deux hélices formant une épingle à cheveux (sous-domaine C2). Le domaine HBD est suffisant pour lier DnaB et stimuler ses activités ATPase et hélicase. La stabilité de l'interaction entre DnaB et DnaG varie en fonction des espèces bactériennes. L'interaction entre EcDnaB et EcDnaG est relativement faible, de l'ordre du micromolaire. En effet, en résonance plasmodique de surface (SPR), elle varie entre 1,4 μM , pour EcDnaG entière, et 4,9 μM pour EcDnaG^{HBD} (Oakley *et al.*, 2005). L'affinité est sans doute plus forte dans le cas du complexe BstDnaB•BstDnaG puisque le complexe résiste au passage sur colonne de gel filtration (Bird *et al.*, 2000). La structure cristalline du complexe BstDnaB•BstDnaG^{HBD} a été résolue par l'équipe du Professeur Steitz (Bailey *et al.*, 2007b) (Figure 18).

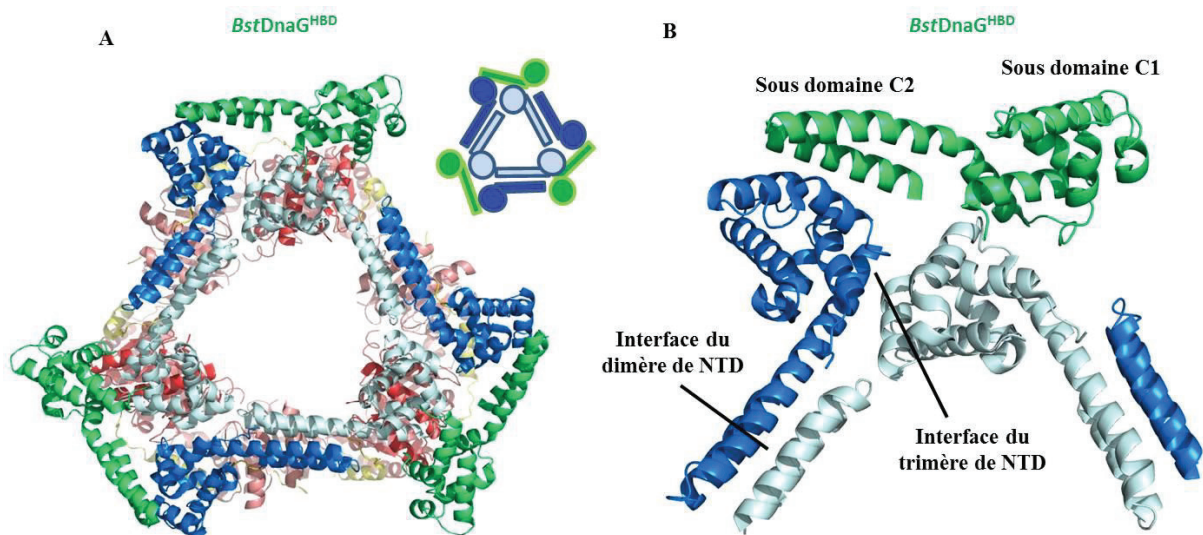


Figure 18 : (A) Structure du complexe BstDnaB₆•DnaG^{HBD}₃ (vert) vue de dessus. (B) Agrandissement de la zone d'interaction entre un domaine HBD de DnaG en vert et l'anneau NTD de DnaB en bleu (Bailey *et al.*, 2007b).

Dans celle-ci, BstDnaB adopte une conformation similaire à la forme B2 de la protéine seule. Trois molécules de BstDnaG^{HBD} sont liées à l'anneau NTD de BstDnaB. La stœchiométrie est donc d'une molécule de DnaG pour deux molécules de DnaB (BstDnaB₆•BstDnaG^{HBD}₃), ce qui est en accord avec les données obtenues par ultracentrifugation analytique et gel filtration (Bird *et al.*, 2000). Le domaine HBD interagit avec les NTD au niveau des sous-domaines C1 et C2. Cependant, ce complexe planaire n'est pas en accord avec la structure spiralée de DnaB sur l'ADNsb. La structure du complexe BstDnaB•BstDnaG durant la translocation le long de l'ADN et la synthèse des amorces est donc encore mal comprise.

Chez *E. coli*, l'interaction d'EcDnaG avec EcDnaB semble dissocier le complexe EcDnaB•EcDnaC après son chargement sur *oriC* (Makowska-Grzyska et Kaguni, 2010). Cette étape permettrait de libérer l'hélicase et de démarrer la réplication. Cependant, il semble que cette étape pourrait être différente chez les bactéries *B. steraothermophilus* et *B. subtilis*. En effet, la structure d'un complexe ternaire entre BstDnaB₆•BstDnaG^{HBD}₃•BsuDnaI₆ a récemment été résolue (Liu *et al.*, 2013). Ce complexe, formé d'un hexamère de l'hélicase, de trois BstDnaG^{HBD} et de six sous-unités de BsuDnaI, indique qu'une étape supplémentaire serait nécessaire pour démarrer la réplication. Il a donc été proposé que la libération du facteur de chargement DnaI serait partielle lors du chargement du complexe sur l'ADNsb puis totale lors de la translocation de BstDnaB et de la synthèse des amorces d'ARN par BstDnaG.

1.3. *Helicobacter pylori*

H. pylori est une bactérie à Gram négatif qui fait partie de la famille des epsilon-proteobactéries. Son nom lui vient de sa forme incurvée et spiralée. Elle possède deux à six flagelles qui lui confèrent sa mobilité (Goodwin *et al.*, 1990). C'est une bactérie qui mesure 2,5 à 5,0 µm de long et 0,5 à 1 µm de large (Brown, 2000). Sa niche écologique est exclusivement l'estomac humain, pourtant longtemps admis comme un organe stérile à cause de son pH acide. En effet, cette bactérie isolée pour la première fois par Warren et Marshall (Warren et Marshall, 1983) a la capacité de neutraliser l'acidité de l'estomac de son environnement proche. Elle a ensuite rapidement été associée aux gastrites antrales chroniques et aux ulcères (Marshall et Warren, 1984). On estime aujourd'hui que près de 50% de la population mondiale est infectée. La prévalence de l'infection à *H. pylori* est d'environ 20% dans les pays industrialisés, mais passe à plus de 90% dans les pays en voie de développement. Il s'agit de l'infection bactérienne chronique la plus répandue dans le monde (Mégraud, 2003).

La transmission de *H. pylori* semble être essentiellement interhumaine. Les études ont montré que la transmission de la bactérie se produit durant l'enfance au sein de la famille (Bastos *et al.*, 2013). La bactérie persiste durant toute la vie de l'individu en absence de traitement antibiotique. Dans les pays en voie de développement, la contamination par voie fécale-orale et orale-orale sont les sources de contamination. En effet, *H. pylori* est retrouvée dans les vomissures, la salive et les selles. Dans ces pays, les diarrhées fréquentes, l'hygiène mal contrôlée et l'eau qui est souvent non traitée sont des facteurs qui favorisent une

contamination à *H. pylori*. Dans les pays développés, les facteurs de contamination par voie fécales-orales semblent avoir disparus (Mégraud, 2003).

1.3.1. Maladies associées à l'infection de *H. pylori*

En 1994, l'Organisation Mondiale de la Santé a classé *H. pylori* comme la première bactérie carcinogène de classe I. En effet, la présence d'une infection stomacale à *H. pylori* est un facteur de risque pour le développement d'un cancer gastrique. Depuis la fin du XXème siècle, *H. pylori* a été associée à de nombreuses maladies digestives et extra-digestives (Roesler, 2014). Dans la majorité des cas (environ 80%), l'infection reste asymptomatique (Watari *et al.*, 2014), néanmoins une infection à *H. pylori* peut conduire à divers troubles gastro-intestinaux. L'infection de l'estomac par *H. pylori* provoque des gastrites aiguës qui évoluent en général en gastrites chroniques. Chez certains individus, l'infection progresse vers la formation d'ulcères duodénaux dans le cas d'une colonisation antrale ou vers des ulcères gastriques s'il s'agit d'une colonisation du corps et du fundus ; on parle alors de pangastrite (Suerbaum et Josenhans, 2007). L'infection à *H. pylori* peut également provoquer des lymphomes des tissus lymphoïdes à muqueuses associées (MALT) (Covacci *et al.* 1999 ; Kusters *et al.*, 2006). Les statistiques montrent que 72 à 98% des patients atteints d'une MALT sont infectés par *H. pylori* (Suerbaum et Michetti, 2002).

- *Gastrite aiguë et chronique*

La gastrite correspond à une inflammation de la muqueuse de la paroi de l'estomac. On distingue la gastrite aiguë, qui est une inflammation transitoire et qui disparaît après traitement, de la gastrite chronique, qui est une inflammation à plus long terme. L'une des causes peut être une infection chronique par *H. pylori* mais elle peut également apparaître lors d'une maladie auto-immune. Les symptômes de la gastrite aiguë ou chronique sont identiques, c'est pour cela qu'il est difficile dans un premier temps de penser à une infection à *H. pylori*. Les patients vont se plaindre de brûlures d'estomac après les repas qui peuvent être associées à des douleurs au niveau de l'abdomen. Ces deux principaux symptômes peuvent également s'accompagner de ballonnements, de nausées ou encore de vomissements.

- *Ulcères duodénaux et gastriques*

Les ulcères sont une inflammation des couches profondes de la paroi de l'estomac (paroi musculaire de l'estomac). Ces inflammations surviennent lorsque la couche protectrice (la muqueuse) se détériore à certains endroits de l'estomac. Une détérioration de la muqueuse signifie qu'à ces endroits, il n'y a plus de protection physique de la paroi musculaire et le mucus qui la protège contre l'acidité de l'estomac n'est plus sécrété. Les acides peuvent donc pénétrer, entraînant ainsi une inflammation. On parle d'ulcère duodéal lorsque celui-ci est localisé au niveau du duodénum et d'ulcère gastrique lorsque c'est la paroi qui tapisse l'estomac qui est touchée.

- *Lymphomes du MALT (Mucosa-associated lymphoid tissue)*

L'infection à *H. pylori* provoque une réaction immunitaire de la part de l'organisme. Cela se traduit notamment par la production de lymphocytes B, globules blancs circulant dans les vaisseaux lymphatiques. Lorsque l'un d'eux devient anormal du fait d'infections chroniques et qu'il s'infiltré dans la muqueuse de l'estomac, cela provoque l'apparition de proliférations cellulaires malignes qui conduisent au cancer du système lymphatique au niveau de la muqueuse associée.

- *Cancer (adénocarcinome) de l'estomac*

On parle de cancer de l'estomac lorsqu'une tumeur maligne apparaît au niveau de la muqueuse de l'estomac (on parle alors d'adénocarcinome). Si une souche de *H. pylori* est détectée lors d'un cancer de l'estomac, cette souche est très fréquemment CagA positive. Cela veut dire qu'elle possède l'îlot de pathogénicité *cagPAI* qui code pour les protéines formant le système de sécrétion de type IV. Le dernier gène de cet îlot code pour la protéine CagA, l'un des facteurs de virulence les plus importants de *H. pylori* (revue dans Roesler *et al.*, 2014). CagA est une protéine d'environ 130 kDa dont la structure du fragment 1 à 184 a été élucidée récemment par cristallographie aux rayons X (Kaplan-Türköz *et al.*, 2012). La bactérie injecte CagA grâce au système de sécrétion de type IV qui s'apparente à une seringue moléculaire. La protéine se retrouve ainsi à la surface des cellules où elle interagit avec son récepteur : l'intégrine $\alpha 5\beta 1$. Après interaction, elle reste à la surface interne de la membrane plasmique et elle perturbe de nombreuses voies de signalisation de la cellule (revue dans Backert *et al.*, 2011). Même si les détails moléculaires de ses interactions dans la cellule sont encore mal connus, la structure de la protéine est une avancée majeure. Elle rend possible le travail sur

une potentielle inhibition de sa translocation ou de son interaction avec l'intégrine (Kaplan-Türköz et Terradot, 2013).

1.3.2. Prise en charge médicale

En France, le diagnostic d'une infection à *H. pylori*, dans le cas de pathologies gastriques, a été déterminé dans le consensus européen Maastricht III en mars 2005. Il a été repris par la Haute Autorité de santé en avril 2010 dans le cadre des recommandations en santé publique pour le dépistage de l'infection à *H. pylori*. Deux types de diagnostics peuvent être proposés, les tests invasifs ou non invasifs :

- Tests invasifs (ce sont des tests qui nécessitent une endoscopie) :
 - L'examen histologique détecte l'infection et évalue les lésions de la muqueuse.
 - Le test à l'uréase : ce test repose sur la forte activité uréasique de *H. pylori* qui hydrolyse l'urée en ammoniacque.
 - L'amplification génique de l'ADN par PCR de *H. pylori* à partir de biopsies gastriques, voire d'échantillons de selles : cette méthode très sensible permet de détecter des facteurs de virulence de la bactérie et permet de détecter d'éventuelles résistances à certains antibiotiques.
- Test non invasifs
 - Le test respiratoire : le principe est de mettre en évidence l'activité uréasique de *H. pylori*. Pour cela, le patient ingère une urée marquée au ^{13}C . L'uréase bactérienne hydrolyse l'urée en ammonium et en bicarbonate. Le bicarbonate est ensuite converti en CO_2 portant le ^{13}C . L'enrichissement en ^{13}C de l'air expiré permet de déterminer une infection à *H. pylori*.
 - La sérologie : elle permet de détecter dans le sang les IgG dirigées contre la bactérie.
 - La détection dans les selles des antigènes par des anticorps monoclonaux.

Les techniques citées proviennent des recommandations en santé publique sur le dépistage de l'infection à *H. pylori* (Haute Autorité de Santé, avril 2010) ainsi que du Groupe d'Etude Français des Helicobacter (<http://www.gefh.fr>).

- Traitements

Il n'existe pas de traitements spécifiques dirigés contre *H. pylori*. Les médecins utilisent des combinaisons d'antibiotiques : la clarithromycin, l'amoxicillin, le metronidazole et d'autres. Ces antibiotiques sont souvent utilisés en combinaison avec des inhibiteurs de la pompe à proton ou encore des sels de bismuth. Cependant, l'utilisation abusive des antibiotiques durant les soixante dernières années a provoqué l'émergence de souches résistantes. Les souches résistantes à la clarithromycin et au metronidazole réduisent considérablement l'efficacité de la thérapie de 100% à 80% et parfois moins dans plusieurs zones du globe. Cette résistance à l'un ou à l'autre des antibiotiques varie considérablement en fonction de la région et de l'usage des antibiotiques (Mégraud, 2013 ; Garza-Gonzalez *et al.*, 2014).

Il est urgent de développer de nouvelles molécules avec de nouveaux modes d'action pour toutes ces souches bactériennes. Une des difficultés est de trouver de nouvelles cibles thérapeutiques. La plupart des nouveaux antibiotiques ne sont que des améliorations des anciens. La réplication de l'ADN est l'un des mécanismes essentiels les plus conservés mais étrangement, n'a été que peu exploitée pour la recherche d'antibiotiques. En effet, peu de produits naturels efficaces contre la réplication de l'ADN ont été découverts. Seules deux classes d'antibiotiques ont pour cible une protéine de la réplication de l'ADN. Il s'agit des quinolones et des aminocoumarines qui ont pour cible l'ADN gyrase. Les nouvelles techniques de criblage à haut débit ainsi que l'abondance de données structurales des protéines de la réplication et de leurs interactions donnent l'espoir de pouvoir dans un futur proche développer de nouveaux inhibiteurs (Robinson *et al.*, 2012). En effet, les protéines de la réplication sont des cibles de choix pour de nouveaux antibiotiques, car essentielles au développement bactérien. De plus, si les principes gouvernant la réplication sont conservés chez tous les êtres vivants, les détails moléculaires et les structures des protéines sont différents entre eucaryotes, virus et bactéries. Ainsi, la réplication constitue une cible prometteuse pour de nouvelles stratégies de lutte contre les bactéries. Dans ce cadre, la connaissance des mécanismes de réplication particuliers à chaque bactérie, s'ils existent, permettrait une recherche d'antibiotiques spécifiques. La détermination des structures des protéines de la réplication serait une étape déterminante dans cette perspective.

1.4. Réplication de l'ADN chez *Helicobacter pylori*

On trouve chez *H. pylori* les éléments principaux de l'initiation de la réplication : une origine de réplication *oriC*, la protéine initiatrice DnaA et l'hélicase DnaB. Néanmoins, de nombreuses protéines régulatrices de l'initiation présentes chez *E. coli* ou *B. subtilis* sont absentes chez *H. pylori*. De plus, le génome de la bactérie ne code pas pour une protéine homologue de DnaC ou DnaI. Si les principes de bases semblent conservés, les travaux réalisés durant les dix dernières années indiquent que *H. pylori* représente un modèle nouveau de réplication chez les bactéries à Gram négatif.

1.4.1. Une origine de réplication bi-partite

L'étude *in silico* de l'origine de réplication de *H. pylori* a identifié des boîtes de liaison à DnaA en aval et en amont du gène *dnaA*. Cette analyse a également montré que la région riche en AT (DUE) est située en aval du gène *dnaA*, alors qu'elle est située en amont chez les autres bactéries à Gram négatif. La région en amont du gène a donc été appelée *oriC1* et la région en aval *oriC2* (Donczew *et al.*, 2012) (Figure 19). *OriC1* contient cinq boîtes de liaison à HpDnaA (boîtes c1-c5) composées de neuf nucléotides dont la séquence consensus (5'T(C/T)ATTCACN3') correspond à celle des boîtes R de l'origine d'*E. coli*. *OriC2* contient trois boîtes de liaison (boîtes c6-c8) à HpDnaA constituées de neuf nucléotides. La boîte c6 a la même séquence que les boîtes d'*oriC1* alors que les boîtes c7-c8 ont une séquence qui diffère d'un ou deux nucléotides. Il a donc été proposé que la séquence consensus des boîtes de liaison à HpDnaA soit 5'TCATTCACN3' (Donczew *et al.*, 2014a, 2014b). Bien qu'*oriC1* soit essentielle à l'initiation de la réplication, il a été démontré que l'ouverture de l'origine est majoritairement dûe à l'attachement de HpDnaA sur *oriC2*. Une topologie super enroulée d'*oriC2* au niveau des boîtes c7 et c8 serait nécessaire à la fixation de HpDnaA mais aucune topologie particulière n'est requise pour *oriC1* (Donczew *et al.*, 2012).



Figure 19 : Schéma de l'organisation de l'origine de réplication chez *H. pylori*

1.4.2. HobA, régulateur ou composant intégral de l'orisome ?

H. pylori possède une protéine unique nommée HobA (Helicobacter Orisome Binding protein A) qui interagit avec DnaA (Zawilak-Pawlik *et al.*, 2007). La délétion du gène codant pour HobA entraîne un arrêt de la croissance bactérienne (Zawilak-Pawlik *et al.*, 2007). La structure cristalline du complexe $HobA_4 \cdot DnaA^{I-II}_4$ montre que HobA forme des tétramères sur lequel s'attachent quatre molécules de DnaA (Figure 20). Malgré une homologie de séquence très faible, HobA possède une homologie structurale avec DiaA d'*E. coli* (Natrajan *et al.*, 2009). DiaA est une protéine régulatrice de l'initiation de la réplication permettant la synchronisation des événements de réplication avec le cycle cellulaire (Ishida *et al.*, 2004 ; Keyamura *et al.*, 2007). Si le mode d'interaction avec DnaA est le même pour HobA et DiaA, les deux protéines n'ont pas les mêmes fonctions et ne sont pas interchangeables (Zawilak-Pawlik *et al.*, 2011). Enfin des mutations de HobA abolissant l'interaction avec DnaA ne peuvent pas être introduites dans la bactérie. Ceci indique que l'interaction entre HobA et DnaA est essentielle à la survie de la bactérie, ce qui n'est pas le cas avec DiaA. On sait cependant que HobA favorise l'attachement de DnaA sur *oriCI* et pourrait ainsi jouer un rôle structural important dans la formation de l'orisome (Natrajan *et al.*, 2009). Cependant, la fonction de HobA n'est pas encore clairement définie, elle pourrait être associée à d'autres événements que la réplication.

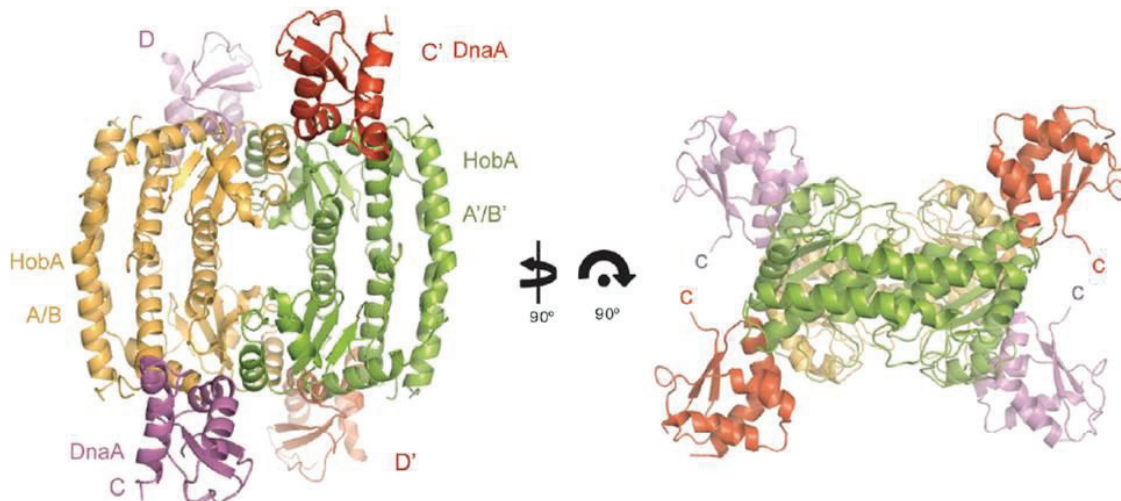


Figure 20 : Structure cristallographique du complexe $HobA_4 \cdot DnaA^{I-II}_4$. Le complexe est constitué de deux dimères de HobA qui interagissent pour former un tétramère. Chaque monomère de HobA s'associe avec le domaine I-II de DnaA (PDB : 2WP0 ; Natrajan *et al.*, 2009).

1.4.3. HpDnaB, un prototype d'hélicase dodécamérique bactérienne

- Structures et propriétés de HpDnaB

La séquence de HpDnaB présente 32% de similarité avec celle d'EcDnaB avec une conservation plus importante du CTD que du NTD (Soni *et al.*, 2003). La structure cristallographique du domaine NTD est composée d'une tête globulaire de six hélices α suivie d'une épingle à cheveux α (Kashav *et al.*, 2009) (Figure 21A). Cependant, la structure est incomplète car la seconde hélice de l'épingle à cheveux est absente dans la structure.

Le rôle du NTD de HpDnaB a également été étudié par mutagenèse dirigée. Chez *E. coli*, une protéine mutante d'EcDnaB privée de son NTD forme des hexamères en solution, possède une activité ATPase mais aucune activité hélicase (Biswas et Biswas, 1999). La protéine mutante équivalente de HpDnaB conserve une activité ATPase, une activité hélicase et forme des hexamères (Nitharwal *et al.*, 2007 ; Soni *et al.*, 2003). Des expériences de complémentation génétique ont été réalisées à partir d'une souche d'*E. coli* thermosensible (ts) pour l'expression d'EcDnaB. A température permissive, le mutant NTD de HpDnaB permet de compléter EcDnaB (Soni *et al.*, 2003). Néanmoins, les activités ATPase et hélicase de la protéine mutante sont plus faibles que celles de la protéine sauvage (Nitharwal *et al.*, 2007). Ceci indique que le NTD de HpDnaB possède des caractéristiques particulières.

Des études biochimiques ont également été réalisées pour étudier la liaison de HpDnaB à l'ADN. HpDnaB est capable de se lier à l'ADN simple brin et double brin correspondant à ce qui est observé pour les autres hélicases. Contrairement à EcDnaB (Arai et Kornberg, 1981 ; Biswas et Biswas, 1999), HpDnaB ne requiert pas la présence d'ATP pour pouvoir lier l'ADNsb. Le CTD de HpDnaB contient un motif RLQLRK (RSRARR pour EcDnaB) qui joue un rôle dans l'interaction avec l'ADN. Ce motif permettrait de coupler l'interaction avec l'ADN et l'activité ATPase, bien que le mécanisme ne soit pas encore compris (Nitharwal *et al.*, 2007). L'interaction de HpDnaB avec le domaine C-terminal de HpDnaG stimule également la liaison de l'hélicase à l'ADNsb. Ces résultats suggèrent que l'affinité de HpDnaB pour l'ADN varie en fonction des interactions avec ses partenaires protéiques. En général, les hélicases n'ont pas de forte affinité pour l'ADN puisqu'elles migrent le long de l'ADNsb. L'anneau NTD masquerait donc le motif RLQLRK afin d'éviter à des interactions non spécifiques avec l'ADNsb. L'interaction avec HpDnaG exposerait le motif RLQLRK ce qui augmenterait l'affinité pour l'ADNsb. La translocation de DnaB serait alors ralentie afin que HpDnaG puisse synthétiser les amorces d'ARN du brin retard (Nitharwal *et al.*, 2011).

- Architecture de HpDnaB

A mon arrivée au laboratoire, la structure du domaine CTD de HpDnaB avait été résolue par le Dr. Meike Stelter, post-doctorante dans l'équipe (Figure 21B) (Article 1, Stelter *et al.*, 2012). On retrouve au niveau du domaine ATPase de type RecA tous les motifs conservés de la famille des hélicases (Ilyina *et al.*, 1992), notamment les motifs Walker A et B. Les éléments de liaison à l'ADN sont cependant différents des autres structures de DnaB. La boucle 1 du domaine RecA de HpDnaB contient un petit brin β et une hélice α supplémentaires. La plus grande différence entre HpDnaB et les autres hélicases est la présence d'une insertion de 34 acides aminés formant un motif hélice-tour-hélice (HTH) appelé *Helicobacter pylori* insertion (HPI). Cette insertion contient dix lysines et dix arginines disposées de manière à créer deux surfaces, l'une positivement chargée et l'autre négativement chargée (Stelter *et al.*, 2012). Le HPI remplace la boucle 3 du domaine RecA, et pourrait participer à la liaison à l'ADN (Figure 21B). Cette insertion est également très importante pour le fonctionnement de HpDnaB puisqu'un mutant de délétion de cette insertion ne montre aucune activité ATPase et hélicase. De plus, cette mutation entraîne un état monomérique de la protéine (Nitharwal *et al.*, 2007).

La structure du domaine C-terminal contenait également l'hélice du domaine de liaison bien que la protéine ne formait pas d'hexamères en solution. Une étude précédente avait montré que la délétion du domaine de liaison entraînait une réduction de l'activité ATPase (en présence ou non d'ADN) et de l'activité hélicase. L'étude sur colonne d'exclusion stérique de ce même mutant montre qu'il forme des dimères. Cela implique donc que le domaine de liaison et son hélice sont importants pour l'assemblage de l'hexamère de HpDnaB (Nitharwal *et al.*, 2007).

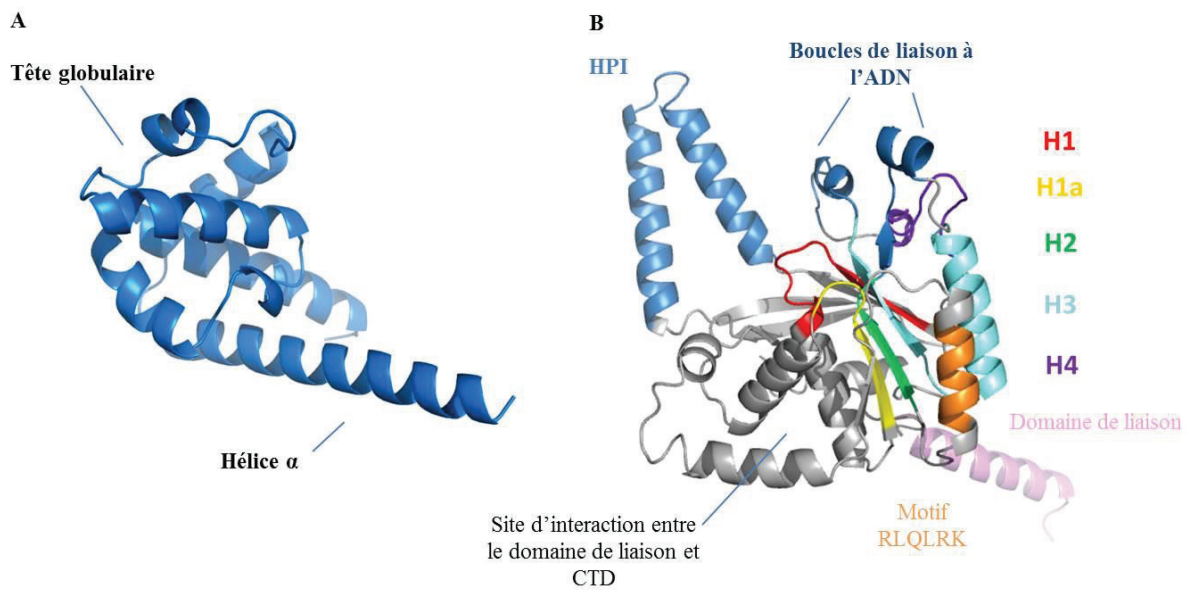


Figure 21 : (A) Structure du domaine NTD de HpDnaB. Il est composé de six hélices formant la tête globulaire et d'une épingle à cheveux α (la deuxième hélice est manquante dans la structure). (B) Structure du domaine CTD de HpDnaB. Le domaine ATPase de type RecA contient tous les motifs conservés des protéines P-loop et le HPI forme un motif HTH.

Les premières études de HpDnaB avaient montré que la protéine possédait une masse de 360 kDa en chromatographie d'exclusion stérique (Soni *et al.*, 2003). A mon arrivée, notre équipe a montré par chromatographie d'exclusion stérique couplée à la diffusion de lumière multi-angles (SEC-MALS) que la protéine a une masse de 680 kDa, soit deux fois la taille d'un hexamère (Stelter *et al.*, 2012). De plus, la présence de nucléotide n'influence pas l'état oligomérique, bien que la protéine soit active et hydrolyse l'ATP. Une étude en ME à coloration négative réalisée par le Dr. Irina Gutsche (UVHCI Grenoble) montre que HpDnaB forme des docécamères qui consistent en deux hexamères interagissant tête-tête au niveau des domaines NTD. Ces travaux ont montré pour la première fois l'existence chez une bactérie d'une hélicase répliquative dodécamérique (Stelter *et al.*, 2012) (Figure 22).

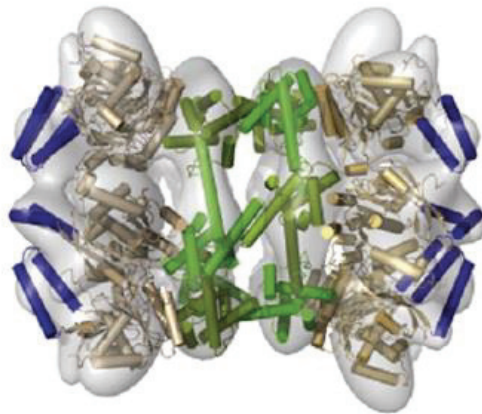


Figure 22 : Reconstruction en trois dimensions de la structure de HpDnaB d'après l'étude par ME à coloration négative à une résolution de 23 Å (Stelter *et al.*, 2012).

La reconstruction de la particule de HpDnaB montre que les CTD sont organisés selon une symétrie C6 avec une organisation en hélice. De manière surprenante, les domaines NTD sont aussi organisés selon une symétrie C6. Dans la carte de ME, on observe un anneau supplémentaire sur l'extrémité du CTD. Le modèle de HpDnaB, construit en remplaçant les structures cristallographiques du CTD et NTD obtenues individuellement dans la carte, montre que l'anneau supplémentaire correspond à l'agencement des HPI. L'analyse de la distribution des charges de ce domaine suggère que l'anneau HPI a un rôle dans la liaison à l'ADN (Figure 23).

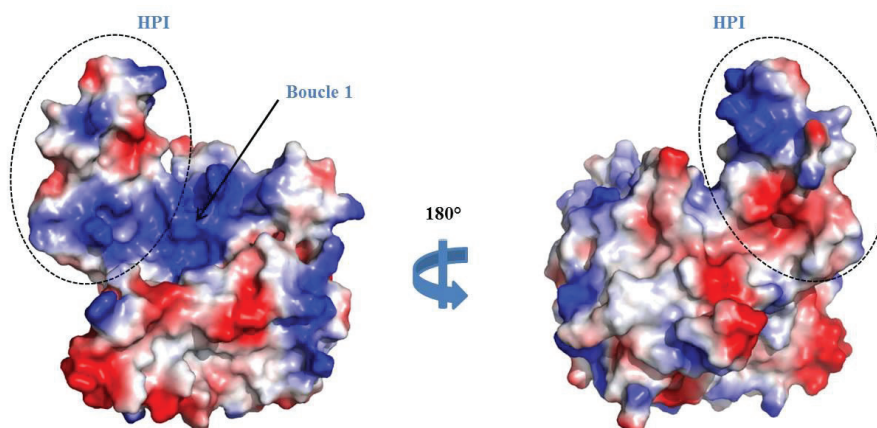


Figure 23 : Représentation de la répartition des charges à la surface du CTD de HpDnaB, la couleur bleue représente les régions positivement chargées et la couleur rouge représente les régions négativement chargées.

Comme nous l'avons vu précédemment, l'architecture en double anneau est en fait répandue chez de nombreux organismes. Des dodécamères ont été identifiés chez les MCM d'archées, les Mcm 2-7 eucaryotes et les virus (Figure 24).

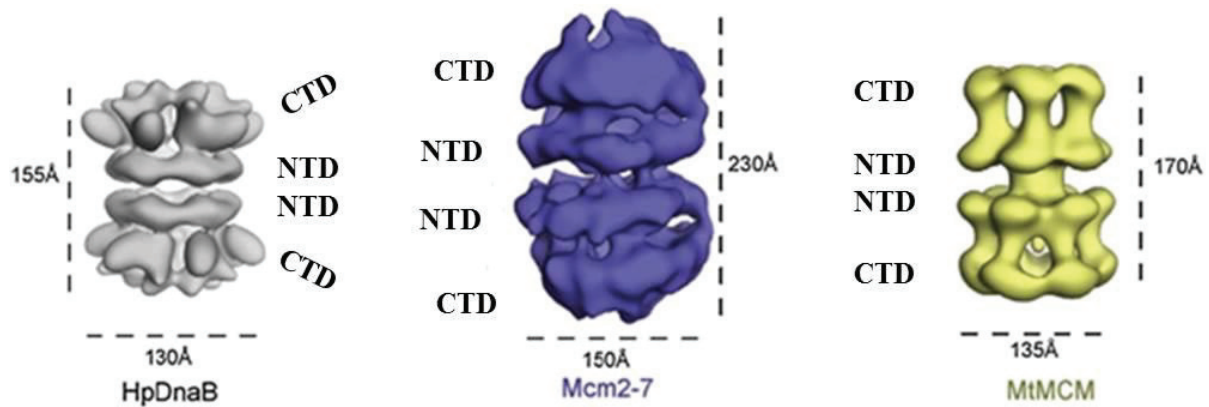


Figure 24 : Enveloppes de ME de HpDnaB (en gris), de Mcm 2-7 (en bleu) et de MtMCM (en jaune).

La découverte du double anneau de HpDnaB est importante compte tenu des propriétés uniques de cette protéine. En effet, l'expression de HpDnaB chez une souche mutante d'*E. coli dnaC^{ts}* compense la perte d'expression de DnaC. Cette souche est capable de survivre et de se répliquer à des températures permissives. Cela suggère que HpDnaB est capable de se passer de facteur de chargement pour initier la réplication chez *E. coli* (Soni *et al.*, 2005). Nous avons donc proposé que HpDnaB pourrait se charger sans facteur de chargement sur la fourche de réplication. La dodécamère apporte en effet une solution topologique évidente pour placer deux hexamères dans les deux directions opposées au niveau de l'origine de réplication en l'absence de facteur de chargement (Figure 25). De plus, un changement de la symétrie C6 en symétrie C3 pourrait permettre un mouvement des NTD qui laisserait le passage pour l'ADNsb. De manière intéressante, un tel changement de symétrie avait été observé en ME à force atomique pour BsuDnaB en présence de BsuDnaG^{HBD} (Thirlway *et al.*, 2004). Un tel rôle pourrait être joué par HpDnaG. Enfin, le HPI pourrait, durant ces mouvements du NTD, maintenir les CTDs sous forme d'hexamères.

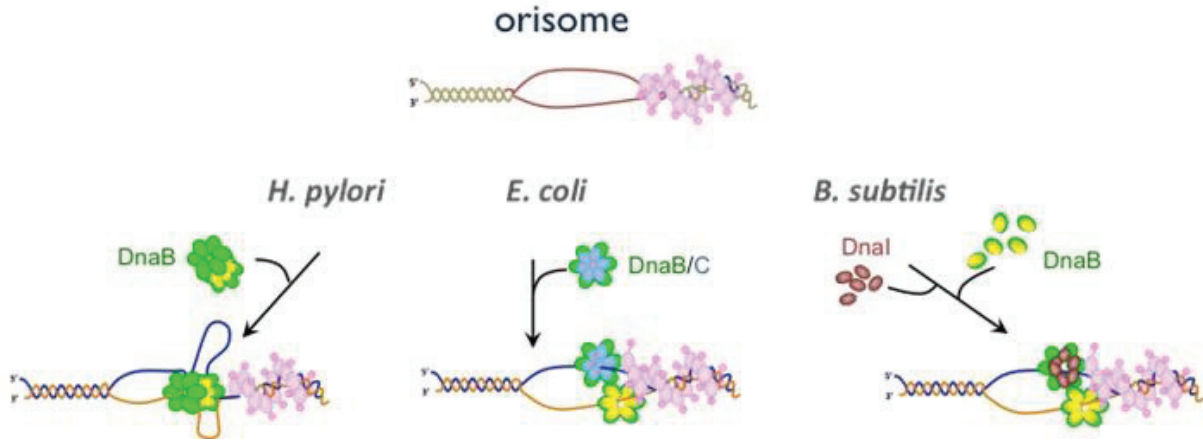


Figure 25 : Représentation schématique de l'initiation de la réplication de l'ADN chez *H. pylori*, *E. coli* et *B. subtilis*

- Primase DnaG de *H. pylori* (HpDnaG)

Bien que les trois domaines composant DnaG entière soient communs aux bactéries, il n'y a que très peu de similarité de séquence entre HpDnaG, cDnaG et BstDnaG. Cependant, la structure de HpDnaG^{HBD} résolue à 1,8 Å de résolution est similaire à celle d'EcDnaG et de BstDnaG (Abdul Rehman *et al.*, 2013). Celle-ci est composée d'une tête globulaire suivie d'une épingle à cheveux α . En revanche, le nombre total d'hélices α diffère : six pour HpDnaG^{HBD}, sept pour EcDnaG^{HBD} et huit pour BstDnaG^{HBD}. La superposition des trois structures au niveau des têtes globulaires montre que l'épingle à cheveux α est très flexible en fonction des espèces. Cette flexibilité laisse penser que l'épingle à cheveux α peut adopter différentes orientations selon que la protéine soit liée ou non à l'hélicase. L'analyse de la structure et la modélisation d'un complexe HpDnaG^{HBD}•HpDnaB^{NTD} *in silico* ont mis en évidence un acide aminé non conservé qui serait important pour l'ancrage et la stabilisation de l'épingle à cheveux α , la Phénylalanine 534 (Phe 534). Des expériences de SPR ont montré une diminution de l'affinité primase-hélicase si Phe 534 est mutée en alanine. Néanmoins, l'interaction entre les deux protéines reste toujours de l'ordre du nanomolaire (Abdul Rehman *et al.*, 2013). On sait que la primase stimule les activités ATPase et hélicase de DnaB, et inversement que DnaB stimule l'activité primase de DnaG (Bird *et al.*, 2000 ; Johnson *et al.*, 2000). Dans le cas du mutant F534A, l'activité hélicase est diminuée de 50% par rapport à la protéine sauvage alors que l'activité ATPase n'est pas affectée. De plus, la liaison de HpDnaB à l'ADNsb est stimulée pour HpDnaG^{HBD} mais aussi par le mutant F534A. Les résultats de dynamique moléculaire et d'activité ATPase de cette étude suggèrent que ce résidu est

important pour la formation du primosome de *H. pylori*. Cependant, cet acide aminé n'étant pas conservé, il serait spécifique à *H. pylori* (Abdul Rehman *et al.*, 2013).

DnaG n'est pas la seule protéine modulant l'activité de DnaB. Il a été montré chez *E. coli* que SSB stimule l'activité hélicase d'EcDnaB et inhibe l'activité ATPase (Biswas *et al.*, 2002). La protéine SSB dont le rôle principal serait de stabiliser l'ADNsb, régulerait donc chez *E. coli* les activités ATPase et hélicase d'EcDnaB. Une même étude chez *H. pylori* a montré une interaction par co-élution entre HpSSB et HpDnaB. HpSSB a une action inhibitrice sur l'activité ATPase de HpDnaB en présence d'ADNsb. HpSSB aurait, par ailleurs, la capacité de stimuler l'activité hélicase de HpDnaB à faible concentration et au contraire de l'inhiber à haute concentration (Sharma *et al.*, 2008).

1.5. Objectifs de la thèse

Mon travail s'est inscrit dans la problématique générale du laboratoire qui est de mieux comprendre les étapes clés de l'infection à *H. pylori*. Les infections à *H. pylori* provoquent de nombreuses maladies gastriques et en particulier le cancer de l'estomac, responsable de plus de 800 000 morts par an dans le monde. Comme nous venons de le voir, le système de l'initiation de la réplication chez *H. pylori* présente de nombreuses divergences par rapport aux systèmes modèles. Le chargement de l'hélicase réplicative HpDnaB s'opère par un mécanisme particulier, ne faisant pas appel à un facteur de chargement de type DnaC/I. L'enveloppe de HpDnaB obtenue par ME suggère que la formation du primosome doit également faire appel à des mécanismes différents. Comprendre les détails moléculaires de ces événements pourrait apporter des informations importantes sur l'initiation de la réplication chez *H. pylori* mais également chez d'autres bactéries. La réplication étant une étape essentielle du cycle bactérien, une meilleure compréhension de ces mécanismes ouvrirait de nouvelles pistes pour la recherche d'antibiotiques.

Mes études se sont donc focalisées sur la structure de l'hélicase HpDnaB puis sur la formation du primosome par interaction avec la primase HpDnaG et l'ADN. Pour cela, une étude biochimique et structurale a été entreprise :

- Etude structurale de HpDnaB

A mon arrivée au laboratoire, la structure cristallographique du domaine ATPase du monomère de HpDnaB et une enveloppe de ME de la protéine entière sous forme de dodécamère avaient pu être déterminées. Cependant, la résolution obtenue (23Å) en ME ne permettait ni de déterminer les positions des domaines de la protéine ni sa structure avec précision. J'ai donc entrepris de résoudre la structure tridimensionnelle de la protéine entière par cristallographie aux rayons X afin d'obtenir un meilleur modèle. J'ai également utilisé le SAXS pour déterminer la structure et la dynamique de la protéine en solution.

- Etude du complexe HpDnaB•HpDnaG^{HBD}

Je me suis ensuite intéressé au complexe HpDnaB•HpDnaG^{HBD}. Pour ce faire, j'ai tout d'abord essayé de cristalliser le domaine HBD de HpDnaG. Cependant, la structure de cette dernière a été déposée dans la banque de données à ce qui nous a conduit à abandonner ce projet. Mes travaux se sont donc focalisés sur le complexe HpDnaB•HpDnaG^{HBD}. Ce travail a fait appel à de nombreuses techniques (micro calorimétrie par titration, SEC-MALS, SAXS

et ME et avait pour objectif de définir les conditions de formation du complexe et les conséquences sur la structure et l'activité de l'hélicase. Les résultats obtenus nous ont amené à étudier les interactions des protéines et du complexe avec l'ADN.

- Etude du chargement de HpDnaB sur l'origine

Dans un troisième temps, je me suis intéressé à la formation *in vitro* de l'orisome et du primosome de *H. pylori*. Des expériences ont été réalisées à partir des protéines purifiées (HpDnaA, HpDnaB, HobA et SSB) sur la région *HporiC* afin de reconstituer les différentes étapes de l'initiation et ainsi de mieux comprendre le mécanisme de chargement de HpDnaB.

2. Matériels et Méthodes

2.1. Biologie Moléculaire

2.1.1. Souches bactériennes et plasmides

Le clonage des différents gènes a été réalisé dans la souche d'*E. coli* Top10 (Invitrogen) chimiquement compétente (F^- *mcrA*, $\Delta(mrr-hsdRMS-mcrBC)$, $\Phi80lacZ\Delta M15$, $\Delta lacX74$, *recA1*, *araD139*, $\Delta(araleu)$ 7697, *galU*, *galK*, *rpsL(Str^R)* *endA1*, *nupG*). Les souches bactériennes utilisées ensuite pour la surexpression des protéines sont les souches Bl21 F^- *ompT hsdS_B* (rB^- mB^-) *gal dcm* (DE3) ou Bl21* F^- *ompT hsdS_B* (rB^- mB^-) *gal dcm rne131* (DE3) (Invitrogen). Ces souches sont déficientes pour les gènes *Lon* et *ompT* qui codent des protéases pouvant dégrader la protéine. Les souches Bl21* portent également une mutation du gène de la RNase E réduisant le taux de RNase E intracellulaire. Cette souche permet de réduire la dégradation des ARNm codant pour la protéine à surproduire. Les gènes sont insérés dans un plasmide pET151/D-Topo (résistance à l'ampicilline, Invitrogen), pACYCDuet-1 (résistance au chloramphénicol, Novagen) ou pRSFDuet-1 (résistance à la kanamycine, Novagen) (Figure 26).

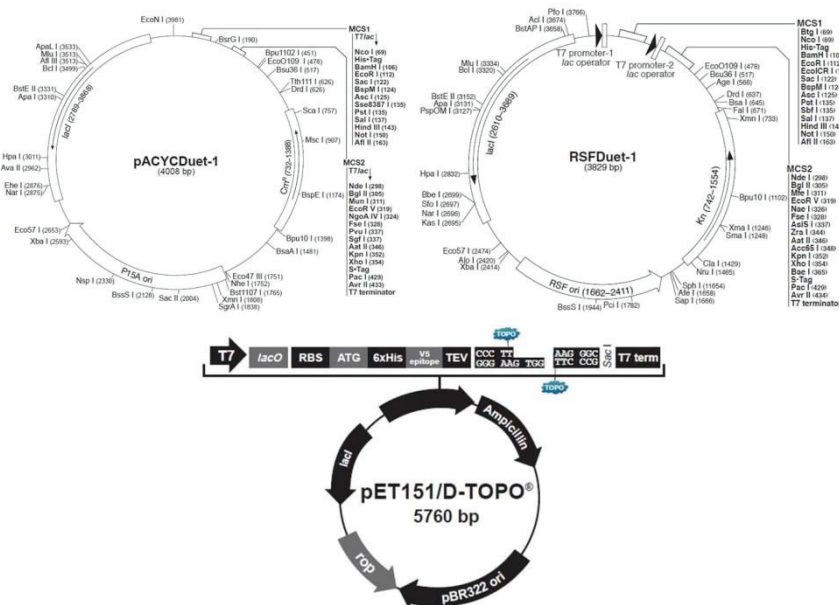


Figure 26 : Carte des plasmides pACYCDuet-1, pRSFDuet-1 et pET151/D-Topo

2.1.2. Clonage des gènes d'intérêt

Les vecteurs exprimant les protéines d'intérêt ont été créés en utilisant le kit pET151/D-Topo commercialisé par Invitrogen. Le clonage en pET151 permet de fusionner l'extrémité N-terminale de la protéine avec une étiquette composée de six histidines suivie d'un site de clivage à la TEV. Les gènes codant pour HpDnaB^{NTD} (résidus 1 à 153) et HpDnaG^{HBD} (résidus 415 à 559) ont été amplifiés par réaction de polymérisation en chaîne (PCR) en utilisant la HF (High Fidelity) polymérase en présence d'oligonucléotides spécifiques (Tableau 2). Le produit PCR est mis en présence du plasmide pET151/D-Topo et incubé 10 min à température ambiante ; afin que la Topoisomérase 1 du kit effectue la ligation. L'échantillon est ensuite placé dans la glace pendant 5 min. 40 µL de bactéries compétentes Top10 sont ajoutés au mélange réactionnel. Les cellules sont ensuite transformées par choc thermique (20 min dans la glace, 40 sec à 42°C, 5 min dans la glace). 200 µL de milieu SOC sont ajoutés et les cellules sont incubées 1h à 37°C sous agitation. Les cultures sont étalées sur boîtes de Pétri en présence d'ampicilline et incubées toute la nuit à 37°C.

Pour les protéines 6His-HpDnaB, HpDnaB, HpDnaA, les trois constructions étaient disponibles à mon arrivée au laboratoire. Le gène codant pour HpDnaB avait été inséré dans le vecteur pET151 (pour la production de 6His-HpDnaB) ou dans les sites de restriction *NdeI* et *XhoI* du vecteur pACYCDuet-1. Le gène codant pour HpDnaA avait été inséré dans les sites de restriction *NcoI* et *SacI* du vecteur pRSFDuet-1.

Protéines	Résidus	Séquences des amorces (5'-3')	
		sens	antisens
HpDnaB ^{NTD}	1-153	sens	caccatggatcatttaaagcatttgcag
		antisens	ctaattcctaaagccttctatgggtgc
HpDnaG ^{HBD}	415 à 559	sens	caccgtctcttttcagcctttttacc
		antisens	tcatatggcgactaattctccttg

Tableau 2 : Tableau des constructions réalisées.

2.1.3. Milieux et conditions de culture

Le milieu LB peut être préparé sous forme liquide pour les cultures et sous forme solide pour les boîtes de Pétri. Dans les deux cas, le milieu LB contient 10 g/L de Tryptone, 5 g/L d'extrait de levure et 5 g/L de NaCl. La poudre est ensuite dissoute dans l'eau distillée et stérilisée par autoclavage. Pour les milieux solides, 20 g/L d'agar bactériologique sont ajoutés au LB liquide. Les bactéries transformées sont cultivées à 37°C dans du milieu LB liquide contenant l'antibiotique approprié jusqu'à ce que les bactéries atteignent la phase exponentielle de croissance ($DO_{600} = 0,6$ environ). La production de protéine est ensuite induite par ajout de l'isopropyl β -D-1-thiogalactopyranoside (IPTG) à 1 mM (concentration finale). Les cultures sont arrêtées par centrifugation.

En cristallographie, il est possible de résoudre le problème d'obtention des phases d'une protéine en utilisant des cristaux de cette même protéine sélénométhionylée. En contrôlant de manière stricte les sources d'acides aminés, il est possible de produire une protéine ayant intégrée, à la place de ces méthionines, des sélénométhionines (l'atome de soufre étant remplacé par un atome de sélénium). Les souches bactériennes non-auxotrophiques pour la méthionine peuvent être utilisées pour produire des protéines sélénométhionylées. En effet, une haute concentration d'isoleucine, lysine et thréonine dans le milieu de culture permet de bloquer la voie de biosynthèse de la méthionine. De plus, la phénylalanine et la lysine agissent en synergie avec la lysine (Doublé, 1997). La bactérie utilise la sélénométhionine présente dans le milieu de culture pour la synthèse protéique. Pour réaliser cela, les bactéries sont cultivées dans un milieu minimum (M9). Ce milieu est appauvri en métabolites comparé au milieu LB et entraîne une baisse de la croissance bactérienne. On y ajoute donc des sels comportant les éléments chimiques strictement nécessaires à la croissance bactérienne ainsi que les acides aminés.

Les sels pour le milieu M9 (sels M9) sont préparés de la manière suivante : à 800 mL d'eau distillée sont ajoutés : 34 g de Na_2HPO_4 , 15 g de KH_2PO_4 , 2,5 g NaCl et 5 g de NH_4Cl . Une fois les sels dissous, le volume est ajusté à 1 litre et ces derniers sont autoclavés.

Pour 1 litre de milieu M9, on ajoute à 778 mL d'eau stérile : 200 mL de la solution de sels M9, 2 mL de $MgSO_4$ 1 M stérile, 20 mL de glucose 20% (g/v) filtré et stérile, 100 μ L de $CaCl_2$ 1 M stérile.

2.1.4. Antibiotiques utilisés

Chaque antibiotique est préparé sous forme de solution mère conservée à -20°C aux concentrations suivantes : ampicilline à 100 mg/mL (dans de l'eau), chloramphénicol à 34 mg/mL (dans l'éthanol absolu) et kanamycine à 50 mg/mL (dans de l'eau). Lors des cultures bactériennes, chaque solution d'antibiotique est diluée 1000 fois dans le milieu de culture. Les antibiotiques permettent de sélectionner et de maintenir une pression de sélection des bactéries portant le ou les plasmides utilisés.

2.2. Purification des protéines

2.2.1. Méthodes de purification

2.2.1.1. Chromatographie d'affinité

Cette technique utilise une colonne contenant une résine d'agarose sur laquelle sont liés des ions nickel par l'intermédiaire de l'acide nitrilotriacétique. Les protéines produites avec une étiquette six histidines ont alors la capacité de se fixer au nickel grâce à des liaisons de coordination qui peuvent se former entre le noyau imidazole de l'histidine et le nickel. Les protéines ne portant pas d'étiquette six histidines ne sont pas retenues et sont alors éliminées par lavage. L'élution de la protéine se fait en utilisant un gradient d'imidazole, qui va entrer en compétition avec les histidines de l'étiquette pour la liaison avec le nickel. L'attachement de l'imidazole provoque le détachement de la protéine d'intérêt qui peut alors être récupérée en sortie de colonne.

L'autre technique de chromatographie d'affinité utilisée fait appel à une colonne héparine. La colonne est constituée d'une matrice d'agarose couplée de façon covalente à de l'héparine (molécule composée d'acide uronique et D-glucosamine) qui possède la particularité d'imiter la structure et la charge de l'ADN. Les protéines liant l'ADN peuvent donc théoriquement se fixer à cette phase stationnaire, alors que la plupart des autres protéines ne le font pas et sont éliminées.

2.2.1.2. Chromatographie échangeuse d'ions

Cette technique sépare les protéines selon leur point isoélectrique (pI, qui représente la charge globale de la protéine). On utilise une colonne composée d'une matrice d'agarose sur laquelle sont greffés soit des groupements anioniques soit des groupements cationiques. En fonction du pH du tampon et du pI de la protéine d'intérêt, cette dernière possédera soit une charge globale négative, soit une charge globale positive. Dans le cas d'une colonne échangeuse d'anions, les protéines chargées négativement se fixeront sur la colonne alors que les protéines chargées positivement se retrouvent dans la fraction non retenue. C'est la situation inverse qui se passe avec une colonne échangeuse de cations. Les protéines sont éluées avec un gradient de NaCl (de 0 à 1M). Le pourcentage de NaCl augmentant dans le tampon d'éluion, la force ionique va augmenter et les protéines sont décrochées en fonction du nombre de groupes chargés à leur surface. Moins elles sont chargées, plus elles seront éluées tôt durant le gradient de NaCl.

2.2.1.3. Chromatographie d'exclusion stérique

La chromatographie d'exclusion stérique permet de séparer les protéines en fonction de leur taille et de leur forme. Pour cela, la solution protéique est injectée sur une colonne composée de granules de gel poreux. Si une protéine a une taille supérieure à celui du diamètre des pores, elle ne peut pas entrer dans les pores et sera éluee en première (on parle de volume mort de la colonne). Les protéines de tailles intermédiaires et les petites protéines entrent dans les pores, parcourent un trajet plus long, et en conséquence leur temps d'éluion est augmenté. Plus une protéine est petite, plus elle sera retenue dans les pores de la colonne.

2.2.2. Analyse des protéines et dosage

Afin de vérifier la pureté des fractions protéiques à l'issue d'une chromatographie, on utilise la technique du gel de polyacrylamide SDS (Dodécylsulfate de sodium). Le gel est placé dans une cuve d'électrophorèse. Le principe de l'électrophorèse repose sur le déplacement des ions sous l'effet d'un champ électrique. La migration des échantillons sur gel est réalisée en condition dénaturante : ils sont chauffés 3 min à 100°C en présence d'un agent réducteur, le β -mercaptoéthanol et de dodécylsulfate de sodium. Le SDS va se fixer aux protéines et leur conférer une charge globale négative. Sous l'effet d'un champ électrique, les

protéines ainsi chargées vont donc se déplacer vers l'anode de la cuve électrophorétique et se séparer les unes des autres en fonction de leur poids. En effet, les gels utilisés contiennent du polyacrylamide qui en polymérisant forme un réseau. Plus le poids moléculaire d'une protéine est élevé, moins elle migrera dans le gel. Les protéines sont ensuite révélées en colorant le gel par du bleu de Coomassie. Les protéines étant séparées en fonction de leur poids, on utilise un marqueur de poids moléculaire (PageRuler Unstained Broad Range Protein ladder, Fermentas). L'estimation de la concentration des protéines s'effectue par spectrophotométrie (Nanodrop 2000, Thermo Scientific) en mesurant les densités optiques à 280 nm et 260 nm. Grâce à l'absorbance à 280 nm de nos échantillons et en utilisant le coefficient d'absorption (ϵ) de la protéine, il est possible d'obtenir leurs concentrations grâce à la loi de Beer-Lambert ($A_{280} = \epsilon \cdot C \cdot l$).

2.2.3. Expressions et purifications des protéines étudiées

2.2.3.1. Expression et purification de 6His-HpDnaB

Une préculture de 20 ml de LB inoculée avec la souche bactérienne BL21* portant le plasmide pET151-6His-HpDnaB est réalisée à 37°C pendant 16h. Un litre de culture LB dans des flasques de 2 litres à encoche est inoculé avec la préculture et les cellules sont placées à 37°C sous agitation jusqu'à une valeur de $DO_{600}=0,6$ (ces étapes seront identiques pour toutes les productions de protéines). La température est diminuée à 20°C et l'expression de la protéine est induite par l'ajout d'IPTG dans le milieu toute une nuit. Les cellules sont centrifugées à 6000 g (JLA-9.100 rotor, Beckman Coulter) pendant 20 min et le culot est resuspendu dans 25 mL de tampon de lyse (10 mM phosphate de sodium pH 7,5, 200 mM NaCl) avec une tablette d'anti-protéase par litre de culture (complete EDTA-free, Roche), du lysozyme (Roche) et de la DNase (Sigma-Aldrich). Les cellules sont lysées par trois sonications de 3 min et centrifugées à 16000 g à 4°C pour retirer les débris cellulaires. La fraction soluble est chargée sur une colonne HisTrap™ HP de 5 mL (GE Healthcare) équilibrée au préalable avec du tampon A (10 mM phosphate de sodium pH 7,5, 200 mM NaCl, 2 mM β -mercaptoéthanol). Après avoir chargé la fraction soluble, la colonne est lavée avec 50 ml de tampon A. Un second lavage est réalisé avec 10 ml de tampon A + 1 M NaCl afin d'éliminer l'ADN contaminant. Une première élution est ensuite effectuée avec 10% de tampon B (10 mM phosphate de sodium pH 7,5 ; 200 mM NaCl ; 2 mM β -mercaptoéthanol et 500 mM imidazole). Cette première élution permet de retirer une partie des contaminants

protéiques ayant une faible affinité voir une affinité aspécifique pour la colonne. L'élution se poursuit avec un gradient allant de 10% à 100% de tampon B. Les différentes fractions des étapes de purification sont analysées sur gel de polyacrylamide SDS 15% (ceci est valable pour l'ensemble des purifications). Les fractions contenant la protéine la plus pure possible sont ensuite rassemblées. La protéine est concentrée par centrifugation sur une unité de filtration amicon de 30 kDa (Milipore) et injectée sur une colonne d'exclusion stérique Superdex 200 10/300GL en utilisant un tampon composé de 10 mM phosphate de sodium pH 7,5 ; 150 mM NaCl ; 2 mM β -mercaptoéthanol. Les fractions éluées contenant la protéine pure sont rassemblées et concentrées jusqu'à 10 mg/mL.

2.2.3.2. Expression et purification de HpDnaB

Les cellules sont incubées à 37°C sous agitation jusqu'à une valeur de $DO_{600}=0,6$. La température est diminuée à 20°C et l'expression de la protéine est induite par l'ajout d'IPTG toute la nuit. Les cellules sont centrifugées à 6000 g pendant 20 min et le culot est resuspendu dans 25 mL de tampon de lyse (10 mM phosphate de sodium pH 7,5, 200 mM) avec une tablette d'anti-protéase par litre de culture (complete EDTA-free, Roche), du lysozyme (Roche) et de la DNase (Sigma-Aldrich). Les cellules sont lysées par trois sonications de 3 min et centrifugées de nouveau à 16000 g à 4°C (JA-20 rotor, Beckman Coulter) pour retirer les débris cellulaires. La fraction soluble est diluée pour atteindre une concentration de NaCl égale à 50 mM. Cette solution est ensuite chargée sur une colonne HiTrap Héparine HP de 5 mL (GE Healthcare) équilibrée au préalable avec du tampon A (10 mM phosphate de sodium pH 7,5, 2 mM β -mercaptoéthanol). Après avoir chargé la fraction soluble, la colonne est lavée avec 50 ml de tampon A. Une première élution est effectuée avec 10% de tampon B (10 mM phosphate de sodium pH 7,5 ; 2 mM β ME ; 1 M NaCl). Cette première élution permet de retirer une partie des contaminants protéiques ayant une faible affinité pour la colonne. L'élution se poursuit avec un gradient allant de 10% à 100% de tampon B. Les fractions contenant la protéine sont ensuite concentrées par centrifugation dans une unité de filtration amicon de 30 kDa (Milipore). La concentration en NaCl du tampon est diluée à 50 mM avec du tampon A et la protéine est injectée sur une colonne d'affinité échangeuse d'anions HiTrap Q HP. La colonne est lavée avec 50 mL de tampon A et la protéine est éluée avec un gradient allant de 10% à 100% de tampon B. Les fractions contenant la protéine pure sont rassemblées

et dialysées contre du tampon A + 150 mM NaCl. La protéine est ensuite concentrée par centrifugation jusqu'à 10 mg/mL.

2.2.3.3. Purification de HpDnaG^{HBD}

Les cellules ont été cultivées à 37°C jusqu'à une $DO_{600}=0,6$. L'expression de la protéine a été induite par l'ajout d'IPTG dans les cultures. Les cellules ont ensuite été placées à 20°C sous agitation toute une nuit. Les cellules ont été centrifugées à 6000 g pendant 20 min et resuspendues dans 25 mL de tampon de lyse (50 mM Tris pH 8, 200 mM NaCl, 5% glycérol) avec une tablette anti-protéase (Roche), du lysozyme (Roche) et de la DNase (Sigma-Aldrich). Les cellules ont été lysées par trois sonications de 3 min et centrifugées à 16000 g pendant 20 min (JA-20 rotor, Beckman Coulter). La fraction soluble est chargée sur une colonne HisTrap Hp 5 mL pré-équilibrée par du tampon A (50 mM Tris pH 8, 200 mM NaCl, 1 mM de dithiothréitol (DTT) et 5% glycérol). La colonne est ensuite lavée avec dix volumes colonne de tampon A. Un second lavage est effectué avec du tampon A +1 M NaCl afin d'éliminer l'excès d'ADN. Les premiers contaminants sont élués avec un lavage à 10% de tampon B (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 5% glycérol et 500 mM Imidazole) suivi d'un gradient d'éluion de 10 à 100% de tampon B. Les différentes étapes de la purification sont analysées sur gel polyacrylamide SDS 15%. Les fractions contenant la protéine sont rassemblées et l'étiquette histidine est clivée par l'ajout de Tobacco Etch Virus (TEV) protéase en présence de 0,5 mM EDTA, puis placées en dialyse contre le tampon A toute la nuit. La dialysat est chargé sur la colonne HisTrap 5 mL pour séparer la protéine de la TEV protéase.

2.2.3.4. Co-expression et purification du complexe HpDnaB•HpDnaG^{HBD}

Pour produire le complexe, les plasmides pACYC-HpDnaB (résistant au chloramphénicol) et pET151-HpDnaG^{HBD} (résistant à l'ampicilline) ont été introduits en même temps dans des cellules BL21* par choc thermique en suivant le même protocole que pour les plasmides seuls. Après inoculation de la préculture, les cellules ont été cultivées à 37°C jusqu'à une $DO_{600}=0,6$. L'expression du complexe est induite par l'ajout d'IPTG. La température est ensuite diminuée à 20°C et les cellules sont placées sous agitation toute la nuit. Celles-ci sont ensuite centrifugées 20 min à 6000 g et resuspendues dans 25 mL de tampon de lyse (50 mM Tris pH 8, 200 mM NaCl) avec une tablette anti-protéase (Roche), du lysozyme (Roche) et de la DNase (Sigma-Aldrich). Les cellules sont lysées par trois

sonications de 3 min et centrifugées 20 min à 16000 g. Le surnageant est ensuite chargé sur une colonne HisTrap Hp 5 mL (GE Healthcare) pré-équilibrée avec du tampon A (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 5% glycérol). La colonne est lavée avec 50 mL de tampon A puis avec 10 mL de tampon A + 1 M de NaCl (pour éliminer l'ADN). On effectue ensuite une première élution avec 10% de tampon B (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 5% glycérol et 500 mM Imidazole) pour retirer les contaminants aspécifiques, suivi d'un gradient d'élution de 10% à 100%.

2.2.3.5. Expression et purification de 6His-HpDnaB-Se

Une préculture de 20 mL de LB + chloramphénicol est inoculée avec une colonie ou un stock glycérol et placée à 37°C toute la nuit sous agitation. Ces 20 mL de culture en LB sont ensuite ajoutés à 1 litre de milieu M9 fraîchement préparé (voir 2.1.2) et placé à 37°C sous agitation jusqu'à une $DO_{600}=0,6$. On ajoute 50 mg de Thréonine, 50 mg de Lysine, 50 mg de Phenylalanine, 50 mg de Leucine, 25 mg d'Isoleucine, 25 mg de Valine (Sigma-Aldrich) et 4 mL de Séléno-méthionine liquide (Molecular Dimensions) dans la culture. Après 15 min d'incubation à 37°C, l'expression de la protéine est induite par l'ajout d'IPTG. Les cellules sont enfin placées à 20°C sous agitation toute la nuit. Les cellules sont centrifugées 20 min à 6000 g et le culot est resuspendu dans 25 mL de tampon de lyse (10 mM phosphate de sodium pH 7,5, 200 mM NaCl, 5 mM β ME). La culture est congelée à -80°C. Pour la purification, la culture est décongelée et une tablette anti-protéase, du lysozyme et de la DNase sont ajoutés. Les cellules sont lysées par trois sonications de 3 min et centrifugées 20 min à 16000 g. Le surnageant est chargé sur une colonne HisTrap Hp 5 mL et la colonne est lavée avec 50 mL de tampon A (10 mM phosphate de sodium pH 7,5, 200 mM NaCl, 5 mM β ME) jusqu'à ce que l'absorbance UV revienne à la ligne de base. On effectue un premier lavage avec 10% de tampon B (10 mM phosphate de sodium pH 7,5, 200 mM NaCl, 5 mM β ME et 500 mM Imidazole) suivi d'un gradient d'élution de 10% à 100%. Les fractions sont ensuite déposées sur un gel polyacrylamide SDS 15%. Les fractions contenant la protéine sont rassemblées puis concentrées par centrifugation pour ensuite être injectées sur une colonne d'exclusion stérique Superdex 200 10/300GL (10 mM phosphate de sodium pH 7,5, 150 mM NaCl, 5 mM β ME). La protéine éluée est concentrée par centrifugation sur une unité de filtration amicon de 30 kDa (Millipore) jusqu'à une concentration finale d'environ 10 mg/mL.

2.2.3.6. Purification de HpDnaA

Les cellules sont cultivées à 37°C jusqu'à une $DO_{600}=0,6$ puis l'expression de HpDnaA est induite par l'ajout d'IPTG. Les cellules sont ensuite placées à 20°C toute la nuit. Les cultures sont centrifugées 20 min à 6000 g et le culot bactérien est repris dans 30 mL de tampon de lyse (45 mM Hepes pH 7,5, 300 mM NaCl, 100 mM de Monopotassium L-glutamate monohydrate, 10% glycérol, 5% Sucrose, 1% Tween, 1 mM DTT, 5 mM $MgCl_2$). Les cellules sont congelées à -80°C et décongelées pour la purification. On ajoute une tablette anti-protéase par litre de culture, du lysozyme et de la DNase. Les cellules sont lysées par trois sonications de 3 min et centrifugées 15 min à 16000 g. Le surnageant est ensuite chargé sur une colonne HisTrap HP 5 mL (GE Healthcare) pré-équilibrée avec du tampon A (45 mM Hepes pH 7,5, 10 mM d'acétate de magnésium, 100 mM de Monopotassium L-glutamate monohydrate, 1 mM DTT, 10% sucrose). La colonne est lavée avec 50 mL de tampon A. Un lavage avec 10% de tampon B (45 mM Hepes pH 7,5, 10 mM d'acétate de magnésium, 100 mM de Monopotassium L-glutamate monohydrate, 1 mM DTT, 10% sucrose, 500 mM Imidazole) élue HpDnaA. Les fractions contenant la protéine sont rassemblées et chargées sur une colonne d'héparine pré-équilibrée avec du tampon C (45 mM Hepes pH 7, 10 mM d'acétate de magnésium, 1 mM DTT et 100 mM de monopotassium L-glutamate monohydrate). La colonne est lavée avec 50mL de tampon C. La protéine est éluée avec un gradient de tampon D (45 mM Hepes pH 7, 10 mM d'acétate de magnésium, 1 mM DTT et 100 mM de monopotassium L-glutamate monohydrate, 1 M NaCl). Les fractions contenant la protéine sont rassemblées, le NaCl est dilué jusqu'à 50 mM avec le tampon C. La solution est ensuite chargée sur une colonne échangeuse d'anions (MonoQ GE Healthcare). La protéine ne s'attache pas à la colonne et est éluée dans la fraction non retenue qui est ensuite dialysée contre du tampon C avec 150 mM NaCl. La protéine est concentrée jusqu'à 5 mg/mL et 20% de glycérol (concentration finale) est ajouté avant de congeler la protéine à -80°C.

2.2.3.7. Purification de HU

Le plasmide pET151 contenant le gène codant pour HU était déjà disponible à mon arrivée au laboratoire. Les bactéries BL21 portant le vecteur ont été cultivées à 37°C dans du LB + ampicilline sous agitation jusqu'à une $DO_{600} = 0,6$. L'expression de la protéine a ensuite été induite par l'ajout d'IPTG dans la culture, les bactéries ont ensuite été placées à 20°C sous agitation toute la nuit. Les bactéries sont centrifugées 20 min à 6000 g et le culot est repris

dans 25 mL de tampon de lyse (40 mM Tris pH 8, 250 mM NaCl, 5% glycérol). Le culot est ensuite congelé à -80°C. Pour la purification, une tablette anti-protéase (Roche), du lysozyme (Roche) et de la DNase (Sigma-Aldrich) sont rajoutés et les bactéries sont lysées par trois sonications de 3 min. Les cellules sont centrifugées pendant 20 min à 16000 g et le surnageant est injecté sur une colonne HisTrap Hp 5 mL pré-équilibrée avec du tampon A (45 mM Tris pH 8, 250 mM NaCl et 5% glycérol). La colonne est lavée avec 50 mL de tampon A. Un second lavage est effectué avec 10 mL de tampon A + 1 M NaCl pour éliminer l'ADN contaminant. Un second lavage est réalisé avec 10% de tampon B (40 mM Tris pH 8, 250 mM NaCl, 5% glycérol et 500 mM Imidazole). La protéine est éluée par un gradient de 10% à 100% de tampon B. Les fractions contenant la protéine pure sont rassemblées et l'étiquette histidine est coupée par la protéase TEV. Pour cela, on ajoute 1 mM DTT, 0,5 mM EDTA, la protéase TEV, et les fractions sont dialysées dans du tampon A à 4°C toute la nuit. Le dialysat est chargé une nouvelle fois sur la colonne HisTrap Hp 5 mL sur laquelle la protéine clivée n'est pas retenue. La protéine éluée est concentrée à 10 mg/mL et conservée à -80°C.

2.2.3.8. Purification de HpDnaB^{NTD}

Pour produire la protéine, les bactéries sont mises en culture dans du LB contenant de l'ampicilline à 37°C sous agitation jusqu'à une $DO_{600}=0,6$. L'expression de la protéine est induite par l'ajout d'IPTG et les bactéries sont placées à 20°C sous agitation toute la nuit. Les cellules sont centrifugées 20 min à 6000 g et le culot est repris dans 25 mL de tampon de lyse (50 mM Tris pH 8, 300 mM NaCl, 1 mM Imidazole). Les bactéries sont ensuite congelées à -80°C. Pour la purification, une tablette anti-protéase est ajoutée par litre de culture avec du lysozyme et de la DNase et les cellules sont lysées par trois sonications de 3 min. Elles sont ensuite centrifugées 20 min à 16000 g. Le surnageant est injecté sur une colonne HisTrap Hp 5 mL équilibrée avec du tampon A (50 mM Tris pH 8, 300 mM NaCl, 1 mM Imidazole). La colonne est lavée avec 50 mL de tampon A. L'excès d'ADN est élué avec un lavage de 10 mL de tampon A + 1 M NaCl. On effectue une première élution avec 10% de tampon B (50 mM Tris pH 8, 150 mM NaCl, 300 mM Imidazole) pour éliminer les contaminants liés de manière aspécifique. La protéine est éluée avec du tampon B et les fractions contenant la protéine sont rassemblées. La concentration en NaCl est diminuée à 50 mM avec du tampon A' (50 mM Tris pH 8, 5% glycérol et 1 mM DTT). La protéine est injectée sur une colonne échangeuse d'anions équilibrée en tampon A'. La colonne est lavée avec 50 mL de tampon A' puis la

protéine est éluée avec un gradient de tampon B (50 mM Tris pH 8, 5% glycérol et 1 mM DTT 1 M NaCl). Les fractions contenant la protéine sont rassemblées et concentrées par centrifugation jusqu'à 10 mg/mL.

2.3. Cristallographie des protéines

2.3.1. Cristallisation des protéines

La cristallographie aux rayons X est la technique la plus utilisée pour résoudre la structure tridimensionnelle des macromolécules biologiques. La résolution d'une structure de protéine nécessite trois étapes : 1) la surproduction et la purification de la protéine d'intérêt pour qu'elle soit la plus pure possible et en quantité suffisante. 2) l'obtention de cristaux de cette protéine (ou cristallogenèse) possédant une bonne qualité de diffraction. 3) La résolution de la structure de la protéine, c'est-à-dire résoudre le problème de phase inhérent à la cristallographie aux rayons X.

2.3.1.1. Paramètres influençant la cristallogenèse

La cristallisation d'une protéine est une transition de phase : elle nécessite de faire passer une protéine d'un état liquide vers un état solide organisé. Cette transition de phase est réalisée en faisant varier un ou plusieurs paramètres physico-chimiques. Ces paramètres sont variables d'une protéine à une autre et sont principalement :

- La nature et la concentration de l'agent précipitant :
 - Les polymères
 - Les sels
 - Les solvants organiques
- La nature et le pH du tampon
- La nature et la concentration d'un additif
- La température
- La pureté, l'homogénéité et la concentration de la protéine

2.3.1.2. Diagramme de phase

La variation de la concentration de l'agent précipitant en fonction de la concentration de la protéine à cristalliser va jouer un rôle critique dans la cristallogénèse qui peut être résumée par un diagramme de phase (Figure 27). Celui-ci met en évidence l'existence de plusieurs zones :

- La zone de solubilité (sous-saturation) dans laquelle la protéine reste soluble et dans laquelle aucun cristal ne pourra apparaître.

La zone de solubilité est délimitée par la courbe de solubilité qui représente la limite entre la phase soluble et la phase solide (cristal, précipité microcristallin ou amorphe).

- La zone de sursaturation qui est divisée en une zone de précipitation et une zone métastable.
 - La zone de précipitation : dans cette zone, la protéine n'est plus soluble. Elle se trouve alors dans un état de sursaturation qui lui permet d'initier la cristallisation : c'est le phénomène de nucléation. La nucléation revient à déstabiliser les interactions protéines-solvant au profit d'interactions protéines-protéines. Des interactions protéines-protéines non spécifiques amènent à la précipitation (phase solide amorphe) alors que des interactions spécifiques amènent à la formation de germes cristallins.
 - La zone métastable : dans cette zone la sursaturation plus faible permet la croissance cristalline. Le cristal croît et la concentration en protéine diminue jusqu'à atteindre la courbe de solubilité : condition d'équilibre où un cristal ajouté à la solution ne pourra ni croître ni se dissoudre.

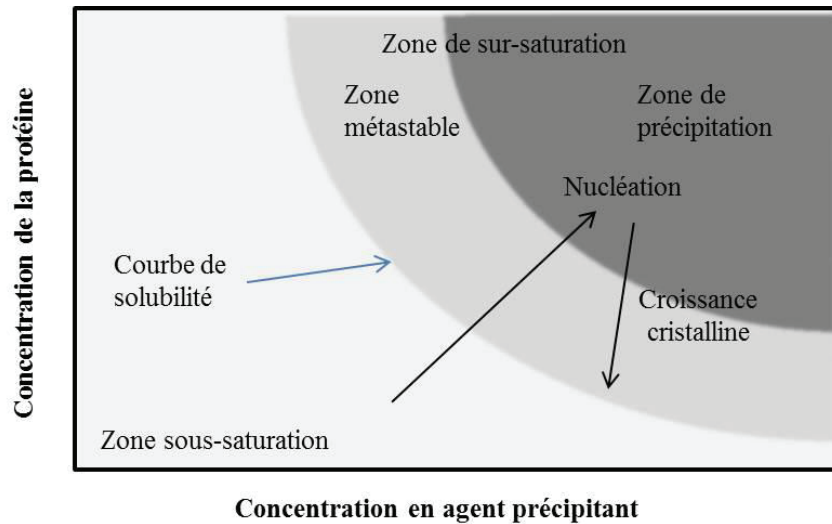


Figure 27 : Diagramme de phase pour la cristallisation d'une protéine

2.3.1.3. Méthodes de cristallogénèse utilisées

La méthode de cristallisation utilisée pendant cette thèse est celle de la diffusion de vapeur (Wlodawer et Hodgson, 1975 ; McPherson *et al.*, 1995) en gouttes suspendues ou en gouttes assises (Figure 28). Dans le cas de la goutte suspendue, le principe est de réaliser une goutte contenant la solution protéique et la solution de cristallisation (contenant l'agent cristallisant, les sels, etc...). On mélange généralement 1 μL de solution de cristallisation à 1 μL de solution protéique pour former la goutte sur une lamelle siliconée (pour éviter que la goutte ne s'étale). Cette lamelle est ensuite retournée au-dessus d'un réservoir contenant 500 μL de la solution de cristallisation. L'ensemble est scellé grâce à de la graisse. La solution de cristallisation dans la goutte étant diluée par la solution protéique, sa concentration est plus faible que dans le réservoir. La vapeur d'eau va alors diffuser (d'où le nom de diffusion de vapeur) jusqu'à ce que la concentration de l'agent précipitant soit la même entre la goutte et le réservoir. Le volume du réservoir étant très supérieur à celui de la goutte, le volume de cette dernière va diminuer et la concentration de la solution de cristallisation augmenter pour tendre vers celle du réservoir. La diminution du volume de la goutte provoque l'augmentation de la concentration de la protéine et de la solution de cristallisation pouvant amener à un état de sursaturation et donc de nucléation. Dans le cas de la goutte assise, le principe est le même à la différence que la goutte est déposée sur un socle dans le réservoir qui est ensuite fermé hermétiquement par un film collant. Les gouttes assises peuvent être effectuées manuellement

ou à l'aide d'un robot de cristallisation, qui permet de cribler plus rapidement un grand nombre de conditions et ce avec une faible quantité de protéine.

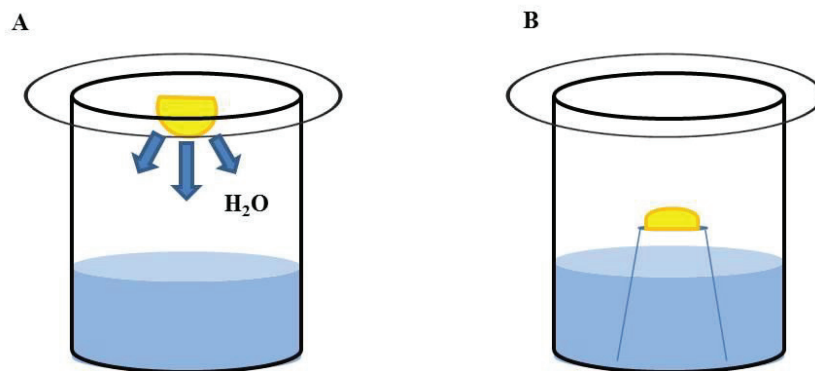


Figure 28 : Schéma du principe de la diffusion de vapeur. (A) Technique de la goutte suspendue. (B) Technique de la goutte assise.

2.3.1.4. Cryo-protection des cristaux

L'un des problèmes de la cristallographie des protéines est l'exposition aux rayons X lors de l'enregistrement des données de diffraction. Les rayons X génèrent des radicaux libres qui endommagent les molécules du cristal, entraînant une diminution progressive de la qualité de diffraction. Pour limiter le plus possible ce phénomène, la solution est de congeler les cristaux afin de limiter au maximum la diffusion des radicaux libres. Les cristaux sont prélevés individuellement à l'aide d'une boucle et refroidis rapidement à une température inférieure à 100 K par immersion dans de l'azote liquide. Pour éviter la formation de glace lors de la congélation, il est nécessaire d'utiliser un cryo-protectant. En effet, si on refroidit le cristal uniquement en présence de sa solution de cristallisation, l'eau contenue dans le cristal formera de la glace et détruira l'organisation cristalline. L'utilisation d'un cryo-protectant permet une vitrification de l'eau du cristal et évite la formation de cristaux de glace. La manière dont on utilise le cryo-protectant est empirique et varie d'un cristal à un autre. Le cryo-protectant peut être directement introduit dans la goutte mais cela peut créer un déséquilibre dans les concentrations des agents cristallisant pouvant entraîner une cassure et/ou une dissolution totale du cristal. Une autre méthode est de transférer au préalable le cristal dans une goutte contenant la solution de cristallisation avec du cryo-protectant. Pour éviter un écart brutal de concentration les cristaux sont transférés successivement dans plusieurs gouttes ayant une concentration croissante en cryo-protectant. La concentration finale en cryo-protectant est elle aussi variable d'une protéine à une autre. Dans notre cas, la

solution de cristallisation de 6His-HpDnaB a été modifiée par l'ajout de 10% (v/v) de sucrose comme cryo-protectant.

2.3.2. Diffraction des rayons X par les cristaux de protéines

La cristallographie est une technique physique qui peut être appliquée aux macromolécules biologiques afin de déterminer leur organisation tridimensionnelle. Pour résumer simplement, on peut comparer la cristallographie au fonctionnement d'un microscope optique: les rayons lumineux, traversant un objet observé, sont diffusés à travers celui-ci puis sont captés par des lentilles convergentes afin de restaurer l'image de l'objet. Dans notre cas, l'objet étudié est une protéine. Néanmoins, on ne peut pas utiliser un rayon de lumière dont la longueur d'onde se situe dans le visible pour observer une protéine. En effet, pouvoir observer la structure d'une protéine signifie pouvoir voir les atomes. De manière générale, pour pouvoir observer un objet, il faut utiliser une onde électromagnétique dont la longueur d'onde est de l'ordre de grandeur de cet objet. C'est pourquoi pour pouvoir observer des atomes, on doit utiliser les rayons X dont la longueur d'onde est de l'ordre de l'Angström (soit 10^{-10} m).

La cristallographie est cependant plus complexe que la microscopie optique classique puisqu'il n'existe pas à l'heure actuelle de lentilles capables de focaliser des rayons X. C'est pourquoi, en cristallographie, les lentilles sont remplacées par des programmes informatiques pour reconstituer l'image.

Mais la difficulté de cette technique ne s'arrête pas à l'utilisation de rayons X. Il est également impossible d'obtenir la structure d'une protéine en n'utilisant qu'une seule molécule. L'intensité des rayons diffusés serait alors beaucoup trop faible pour être utilisable. L'astuce consiste donc à utiliser un cristal de protéine à l'origine du phénomène de diffraction.

2.3.2.1. Définition et description du cristal

Le cristal est un état solide qui est ordonné. Les molécules qui le composent sont empilées de façon régulière et dans les trois dimensions de l'espace ; on parle d'arrangement périodique organisé en réseau. Le réseau est l'ensemble des mailles du cristal, c'est-à-dire l'unité reproduite par translation dans les trois directions. Ainsi, un seul cristal peut être décrit par différentes mailles. La maille peut contenir un arrangement d'éléments plus petit appelé

l'unité asymétrique. C'est l'unité asymétrique qui permet par différents opérateurs du système cristallin de reconstruire l'ensemble du cristal (Figure 29). A l'intérieur du cristal, la cohésion des macromolécules entre elles est assurée par des interactions de type Van der Waals, des ponts salins ou encore des liaisons hydrogènes. L'espace entre les molécules est occupé par le solvant (eau, tampon, ion...). Généralement, les canaux de solvants représentent 40 à 75% du volume cristallin. L'existence de ces canaux peut être un inconvénient car la présence de solvant rend le cristal fragile. Néanmoins, grâce à ces canaux, il est possible de faire diffuser à l'intérieur du cristal des molécules telles que des atomes lourds pour résoudre la structure, ou un substrat pouvant aller se fixer sur le site catalytique de l'enzyme cristallisée.

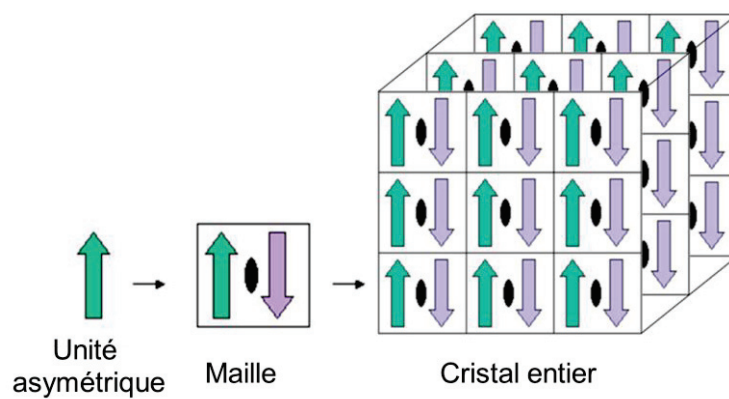


Figure 29 : Schéma représentant la composition d'une unité asymétrique, d'une maille et d'un réseau de cristal de protéine. La flèche représente une molécule de protéine.

2.3.2.2. Systèmes cristallins et groupes d'espace

En fonction de l'arrangement des molécules dans l'espace, on distinguera différents systèmes cristallins basés sur la symétrie de leur réseau. La maille est caractérisée par six valeurs : la norme des vecteurs $\vec{a}, \vec{b}, \vec{c}$ et de ses trois angles α, β, γ . Ces caractéristiques permettent de distinguer sept systèmes cristallins différents. De plus, l'arrangement de ses sept systèmes cristallins ne peut se faire que selon quatre types de réseaux : Primitif (P), Centré (I), Face centrée (F) et Base centrée (C). Les systèmes cristallins peuvent donc être décrits sous la forme de 14 réseaux de Bravais (Figure 30).

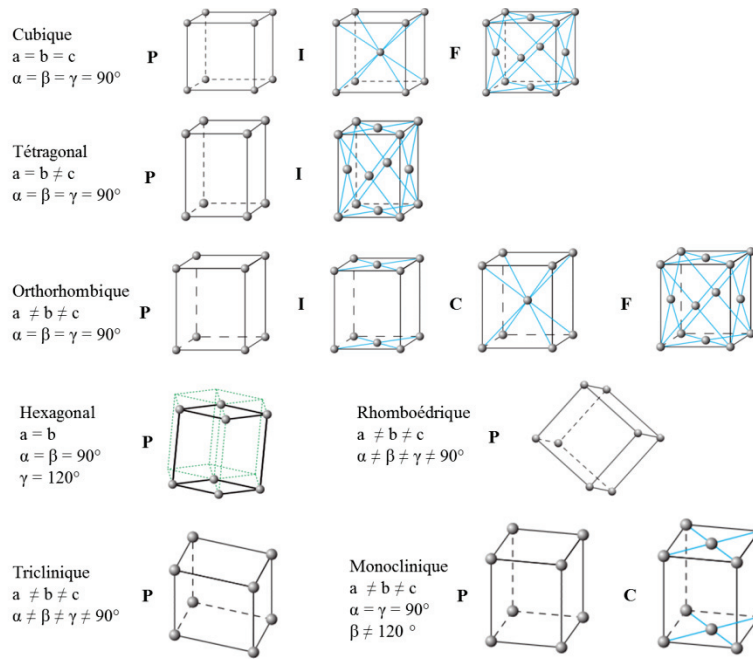


Figure 30 : Représentation des 14 réseaux de Bravais. Les réseaux sont représentés en fonction de leurs systèmes cristallins et du type de réseau : (P) pour Primitif, (I) pour Centré, (F) pour Face centrée et (C) pour Base centrée.

A partir des quatorze réseaux de Bravais et des différentes opérations de symétries possibles, il existe 230 manières d'empiler des objets dans l'espace : on parle de groupes d'espace. Le cristal appartiendra à un groupe ponctuel, dont les caractéristiques mathématiques permettent de décrire l'ensemble des opérations de symétrie qui le compose. Le groupe d'espace d'un cristal décrit non seulement les opérations de symétrie (groupe ponctuel) mais également les opérations de translation. Les propriétés des groupes d'espace sont rassemblées dans les tables internationales de cristallographie disponibles sur <http://it.iucr.org>. Les macromolécules biologiques possédant des propriétés de chiralités, certains groupes d'espace sont impossibles à générer. Les groupes d'espace comportant des symétries miroir ou des centres d'inversions n'existent pas. Il existe donc dans le cas des macromolécules biologiques seulement 36 groupes ponctuels (Tableau 3) et 65 groupes d'espace différents.

Famille cristalline	Réseaux de Bravais	Système cristallin (réticulaire)	Classification des groupes ponctuels
Cubique	<i>cP, cF, cI</i>	Cubique	23, m3, 432, 43m, m3m
Hexagonale	<i>hP</i>	Hexagonal	6, 622, 6mm, 6/m, 6/mmm, 6, 62m
Hexagonale	<i>hP</i>	Hexagonal	3, 32, 3m, 3, 3m
Hexagonale	<i>hR</i>	Rhomboédrique	3, 32, 3m, 3, 3m
Tétraгонаle (quadratique)	<i>tP, tI</i>	Tétraгонаle (quadratique)	4, 4, 422, 4mm, 42m, 4/m, 4/mmm
Orthorhombique	<i>oP, oS, oF, oI</i>	Orthorhombique	222, mm2, mmm
Monoclinique	<i>mP, mS</i>	Monoclinique	2, m, 2/m
Triclinique	<i>aP</i>	Triclinique	1, 1

Tableau 3 : Correspondance des différentes familles cristallines, réseaux de Bravais, systèmes cristallins et groupes ponctuels.

Lorsqu'un cristal est soumis à un faisceau de rayon X, la diffusion est restreinte à certaines directions privilégiées de l'espace ; on parle alors de phénomène de diffraction. En effet, la théorie ondulatoire montre que les rayons X peuvent être représentés par des ondes interagissant par interférences constructives et destructives. L'intensité diffractée va dépendre de la densité diffusante dans le cristal, dans notre cas les électrons à l'intérieur de la maille. Mathématiquement, la diffraction se décrit par les équations de Laue et de Bragg. Les équations de Laue relient le vecteur de diffusion \vec{S} ($S = s_0 - s$ avec s_0 la direction du vecteur incident et s la direction du vecteur diffracté) avec les vecteur définissant la maille cristalline ($\vec{a}, \vec{b}, \vec{c}$) (Figure 31).

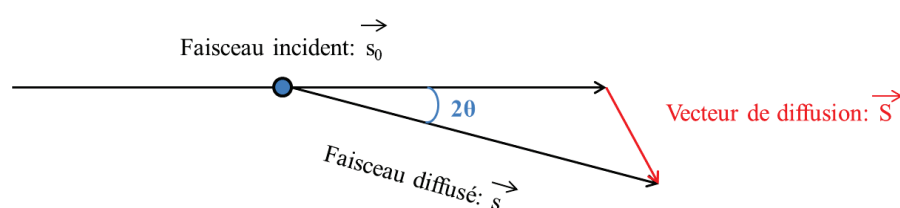


Figure 31 : Représentation schématique de la diffusion d'un rayonnement traversant la matière.

$$\vec{S} \cdot \vec{a} = h \quad \vec{S} \cdot \vec{b} = k \quad \vec{S} \cdot \vec{c} = l \quad \text{avec } h, k, l \text{ des entiers.}$$

Ces équations établissent l'ensemble des directions de diffraction qu'on appelle espace réciproque. Ce réseau est construit à partir du réseau direct (les vecteurs $\vec{a}, \vec{b}, \vec{c}$ de la maille cristalline).

$$\vec{a}^* = (\vec{b} \wedge \vec{c})/V \quad \vec{b}^* = (\vec{c} \wedge \vec{a})/V \quad \vec{c}^* = (\vec{a} \wedge \vec{b})/V$$

Le vecteur de diffusion \vec{S} peut alors s'écrire :

$$\vec{S} = h \cdot \vec{a}^* + k \cdot \vec{b}^* + l \cdot \vec{c}^*$$

Chaque réflexion correspond à un nœud du réseau réciproque et peut être indexée par un jeu de trois indices h, k, l appelés indices de Miller.

La loi de Bragg est une deuxième représentation mathématique de la diffraction. Cette loi définit le cristal comme étant un ensemble de plans réticulaires sur lesquels un faisceau de rayon X peut se réfléchir. La diffraction n'a lieu que si les ondes réfléchies sont en interférences constructives. La loi de Bragg représente la relation entre le faisceau incident de longueur d'onde « λ » qui irradie les plans réticulaires avec un angle « 2θ » (angle entre les faisceaux incident et diffracté). Il y a interférence constructive si $2d \sin\theta = n\lambda$ avec d la distance entre les plan réticulaires, n un entier et λ la longueur d'onde du faisceau incident. A chaque réflexion d'indices h, k, l dans l'espace réciproque correspond une résolution dans l'espace direct, mesurée par d_{hkl} , la distance inter réticulaire de la famille de plans donnant lieu à cette tache. La résolution de chaque intensité diffractée peut être estimée par la relation de Bragg. $2 \sin\theta/\lambda = 1/d_{hkl}$. Ainsi, la résolution dans l'espace réel sera déterminée par l'angle θ . Plus cet angle sera important meilleure sera la résolution.

2.3.2.3. Les facteurs de structure

Le faisceau diffracté étant une onde électromagnétique, celle-ci est représentée mathématiquement par une fonction périodique. Cependant, elle peut également s'écrire comme un nombre complexe. Pour un atome de position r , il est alors possible de représenter le facteur de diffusion atomique par un vecteur sur un plan complexe (ou diagramme d'Argand) dont la norme est l'amplitude de l'onde et l'angle entre l'axe des réels et le vecteur, la phase de l'onde (Figure 32).

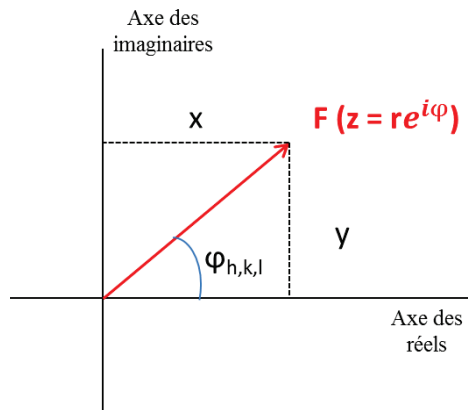


Figure 32 : Représentation du facteur de diffusion d'une onde électromagnétique sur un plan complexe.

Comme vu précédemment, chaque réflexion possède des indices h, k, l . En cristallographie, on utilise le facteur de diffusion atomique pour calculer le facteur de structure d'une réflexion, celui-ci peut alors s'écrire :

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

Une protéine n'est pas constituée d'un seul atome mais de N atomes (j). Pour une réflexion h, k, l , le facteur de structure sera donc défini par la somme des diffusions atomiques de N atomes (j) pondérée par leurs positions dans la maille.

$$F(hkl) = N \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

Avec f_j le facteur de diffusion de l'atome j de coordonnées (x_j, y_j, z_j) dans un cristal comprenant N mailles et n atomes par maille.

On peut reformuler le facteur de structure de chacune des ondes diffractées d'une deuxième façon. En effet, le cristal peut être vu comme le produit de convolution d'une fonction motif (la densité électronique contenu dans la maille) et d'une fonction réseau (la maille) qui répète le motif dans tout le cristal. On peut donc relier les facteurs de structure de chaque réflexion h, k, l à la somme des densités électroniques ρ de coordonnées x, y, z sur l'ensemble du volume de la maille.

$$F(hkl) = \int_V \rho(xyz) e^{i2\pi(hx+ky+lz)} dV$$

Avec ρ la densité électronique, V le volume de la maille et x, y, z les coordonnées fractionnaires de l'espace direct. Il est donc possible d'accéder à l'image de la densité électronique par transformée de Fourier inverse :

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

On peut donc reconstruire la densité électronique d'une protéine à partir de l'ensemble des réflexions enregistrées.

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)| e^{i\varphi_{hkl}} e^{-2i\pi(hx+ky+lz)}$$

Il suffit de mesurer l'amplitude des ondes diffractées (le module du vecteur $|F(hkl)|$) et de sa phase φ_{hkl} . Ces deux valeurs doivent être connues pour accéder à l'image de la densité électronique. Lors d'une expérience de diffraction aux rayons X par un cristal de protéine, il est possible d'accéder à l'amplitude des ondes car on mesure les intensités de ses ondes $I_{(hkl)}$ qui sont proportionnelles au carré de leurs amplitudes : $I_{(hkl)} \approx F_{(hkl)}^2$. Cependant, l'information de phase est perdue au cours de l'expérience. Il n'est donc pas possible de reconstruire directement la carte de densité électronique : « le problème de phase ». Plusieurs méthodes permettent de retrouver les phases : le remplacement moléculaire (à l'aide d'une protéine homologue dont on connaît déjà la structure) et des méthodes expérimentales utilisant le signal de différences isomorphes et/ou anormales.

2.3.2.4. Enregistrement des données de diffraction

Les collectes des données de diffraction ont été réalisées au synchrotron SOLEIL et à l'ESRF (European Synchrotron Radiation Facility). Les cristaux préalablement placés dans des boucles sont positionnés sur une tête goniométrique sous un flux de rayons X. Le cristal au niveau de la tête goniométrique est toujours soumis à un flux d'azote gazeux de 100 K pour éviter qu'il ne se réchauffe. La tête goniométrique permet d'effectuer des petits pas d'oscillation de l'ordre de $0,1^\circ$ à 1° lors de la collecte des données. Le temps d'exposition et l'intensité du faisceau de rayon X dépend du pouvoir de diffraction du cristal et de sa résistance aux dégâts d'irradiation. A chaque pas d'oscillation, les taches de diffraction sont enregistrées sur un détecteur bidimensionnel (on parle de cliché de diffraction). Il est nécessaire d'effectuer des pas d'oscillation afin de récolter un jeu de données complet sinon toutes les taches se superposeraient sur le détecteur. Le nombre de pas d'oscillation va dépendre de la symétrie du cristal.

2.3.3. Traitement et analyse des données de diffraction

Une fois les clichés de diffraction enregistrés, les images sont ensuite analysées :

- **Indexation** : Cette étape détermine les paramètres de mailles du cristal c'est-à-dire les vecteurs \vec{a} , \vec{b} , \vec{c} , leur module ainsi que les angles α , β , γ entre les vecteurs. Ceci permet de déterminer les indices de Miller pour chaque tache pour ensuite intégrer leurs intensités. A partir des paramètres de mailles, le logiciel propose le système cristallin le plus probable.
- **Intégration des intensités** : Cette étape mesure et corrige les intensités en fonction du groupe d'espace et en fonction de la dispersion spatiale des raies de diffraction. Cette dispersion est due à la mosaïcité du cristal, à la largeur spectrale et à la divergence du faisceau incident. Elle a pour conséquence l'étalement des taches sur un diamètre de moins de un à quelques millimètres. Les programmes établissent les profils moyens de ces taches et calculent l'intensité moyenne de chaque tâche par intégration numérique.

- **Mise à l'échelle** : Le cristal possède une symétrie. Cela signifie que les réflexions mesurées sont aussi reliées par symétrie ; on peut donc mesurer plusieurs fois une réflexion. En théorie, l'intensité de deux réflexions symétriques est identique. En pratique, cela peut ne pas être le cas à cause de facteurs tels que les dégâts d'irradiation du cristal par les rayons X ou encore une diminution de la puissance du faisceau incident. Le programme de mise à l'échelle détermine pour chaque image un facteur linéaire et un facteur d'agitation thermique en fonction de la résolution pour corriger ces effets. Une fois corrigées, les intensités équivalentes sont moyennées et les modules des facteurs de structure sont calculés. Cette mise à l'échelle permet d'obtenir l'ensemble des réflexions uniques nécessaires pour calculer l'image de la densité électronique par transformée de Fourier inverse.

Tous les jeux de données collectés ont été indexés et intégrés en utilisant le programme XDS (Kabsch, 1993). Ils ont été mis à l'échelle avec SCALA (Evans, 2006) de la suite de programmes CCP4 (Winn *et al.*, 2011) ou avec le programme XSCALE (Kabsch, 1993).

2.3.4. Paramètres des cristaux

2.3.4.1. Mosaïcité

Les perturbations thermiques ou mécaniques lors de la croissance ou de la manipulation des cristaux peuvent engendrer des défauts dans le réseau cristallin. Cette qualité est caractérisée par le facteur de mosaïcité exprimé en degré. La mosaïcité représente le désordre ou les imperfections à l'intérieur du cristal. Elle oscille généralement entre 0.2° et 1.5°. Plus la mosaïcité est élevée, plus elle entraîne des problèmes de superposition des réflexions ou des difficultés à mesurer des réflexions de faibles intensités. Elle a donc une grande importance dans le traitement et la qualité des données de diffraction.

2.3.4.2. Facteur B moyen

On a vu qu'un cristal de protéine était composé de solvant. A l'intérieur du cristal, les régions de la protéine qui sont exposées au solvant peuvent présenter une certaine variabilité dans leur position et/ou une certaine dynamique de conformation. Les interférences constructives sont donc moindres dans ces régions et cela provoque un affaiblissement de

l'intensité du signal diffracté. Il est alors difficile d'observer de la densité électronique dans ces régions. C'est le facteur B moyen qui reflète ce phénomène et il est exprimé en Å². Généralement, le facteur B moyen des cristaux de protéines varie entre 10 et 70 Å².

2.3.4.3. Estimation du nombre de molécules par unité asymétrique

L'unité asymétrique est le plus petit volume de la maille ne contenant aucune symétrie cristallographique. Elle peut contenir une ou plusieurs molécules. Si elle en contient plusieurs, alors ces molécules sont reliées entre elles par des symétries non cristallographiques. La méthode de Matthews permet d'estimer la quantité de solvant dans le cristal ainsi que le nombre de molécules de protéines contenues dans l'unité asymétrique à partir du poids moléculaire de la protéine cristallisée et des paramètres de l'arrangement cristallin (Matthews, 1968 ; Kantardjieff et Rupp, 2003).

$$V_m = \frac{V}{M * n * op}$$

V étant le volume de la maille, n le nombre de monomères dans l'unité asymétrique, op le nombre d'opérateurs de symétrie et M la masse en dalton d'un monomère. Pour les protéines, V_m est en général compris entre 1,7 et 3,53 Å³/dalton. La quantité de solvant peut être estimée par la relation :

$$S_{\%} = 1 - \frac{1.23}{V_m}$$

2.3.5. Caractéristiques, qualité et validation des données

Plusieurs critères doivent être pris en compte pour juger de la qualité des données de diffraction. Lorsque l'on traite un jeu de données, on obtient ces critères pour chaque palier de résolution.

- **Résolution** : elle détermine le degré de détail de la carte de densité électronique finale obtenue par transformée de Fourier inverse.

- **Rapport Signal/Bruit** : $I/\sigma(I)$ c'est le rapport de l'intensité des réflexions uniques par rapport à leur écart type correspondant au bruit de fond.
- **Complétude** : rapport entre le nombre de réflexions uniques mesurées sur le nombre total de réflexions mesurables. La complétude doit avoir la valeur la plus haute possible sinon cela signifie que la totalité des réflexions uniques n'ont pas été enregistrées. Il est possible pour obtenir une meilleur complétude de combiner des jeux de données isomorphes.
- **Redondance ou multiplicité** : c'est le nombre moyen de fois que les réflexions uniques ont été mesurées. Plus la valeur est grande, plus l'intensité moyenne d'une réflexion sera précise car on minimise les erreurs de mesures.
- **Facteur R_{sym}** : ce facteur représente les différences d'intensités entre les réflexions équivalentes par symétrie au sein d'un jeu de données. Plus le facteur est bas, meilleur est le jeu de données.

$$R_{sym} = \frac{\sum_{hkl} |I_{hkl} - \langle I \rangle|}{\sum_{hkl} I_{hkl}}$$

- **$CC_{1/2}$** : coefficient de corrélation calculé à partir du pourcentage de corrélation entre deux intensités aléatoires. Chaque intensité est issue d'une moitié du jeu de données (Karplus et Diederichs, 2012 ; Diederichs et Karplus, 2013).

2.3.6. Résolution du problème de phase

La détermination des phases est l'une des étapes essentielles de la détermination d'une structure cristallographique. En effet, la qualité de la carte de densité électronique dans laquelle un modèle pourra être construit dépendra en grande partie de la justesse des phases. Il existe plusieurs possibilités pour obtenir l'information de phases manquantes. La première est d'utiliser la structure d'une protéine similaire à la nôtre. On utilise donc les phases de cette structure combinées aux intensités de notre cristal pour obtenir la carte de densité électronique. Cette méthode est le remplacement moléculaire (Rossman et Blow, 1962). La seconde possibilité, dite « expérimentale », nécessite d'utiliser la présence d'un atome lourd

(tel qu'un métal) dans le cristal. Cet atome lourd peut être présent naturellement ou avoir été ajouté durant la production (c'est le cas pour les sélénio-méthionines) ou bien par diffusion dans le cristal avant de le congeler. On peut utiliser cet atome lourd comme source de phasage en utilisant soit les différences isomorphes (en comparaison avec les données obtenues avec le cristal de la protéine native), soit la diffusion anormale apportée par cet atome (une ou plusieurs longueurs d'onde) ou les deux lorsqu'on dispose de jeux de données adaptés. Le choix de l'utilisation de ces méthodes va dépendre des possibilités offertes par le cristal natif et les dérivés disponibles. Dans notre cas, nous avons utilisé la diffusion anormale de cristaux trempés dans le Bromure de Tantalate.

2.3.7. Diffusion anormale

En cristallographie des protéines, le réseau cristallin est centrosymétrique par définition. La loi de Friedel indique qu'en chaque point de la maille cristalline l'intensité d'une réflexion (h, k, l) est identique à celle en $(-h, -k, -l)$, soit $|F(hkl)| = |F(-h - k - l)|$. Par contre, lorsqu'un atome lourd est présent dans la structure et que la longueur d'onde du faisceau incident est proche du seuil d'absorption de ce dernier, ces électrons ne peuvent plus être considérés comme libres et vont entrer en résonance avec les photons. Ce phénomène est appelé diffusion anormale. Lorsque qu'il y a diffusion anormale, le facteur de diffusion atomique s'écrit $f(\lambda) = f^o + f'(\lambda) + if''(\lambda)$ avec f^o le facteur de diffusion en absence de diffusion anormale, $f'(\lambda)$ la partie anormale réelle (dispersive) et $f''(\lambda)$ la partie anormale imaginaire due à l'absorption des rayons X par les électrons de l'atome lourd qui entrent en résonance et réémettent le photon avec un déphasage de $\pi/2$. L'absorption par l'atome lourd rompt la loi de Friedel, dans ce cas $|F(hkl)| \neq |F(-h - k - l)|$ et $\varphi_{hkl} \neq \varphi_{-h-k-l}$ (Blundell et Johnson, 1976) (Figure 33).

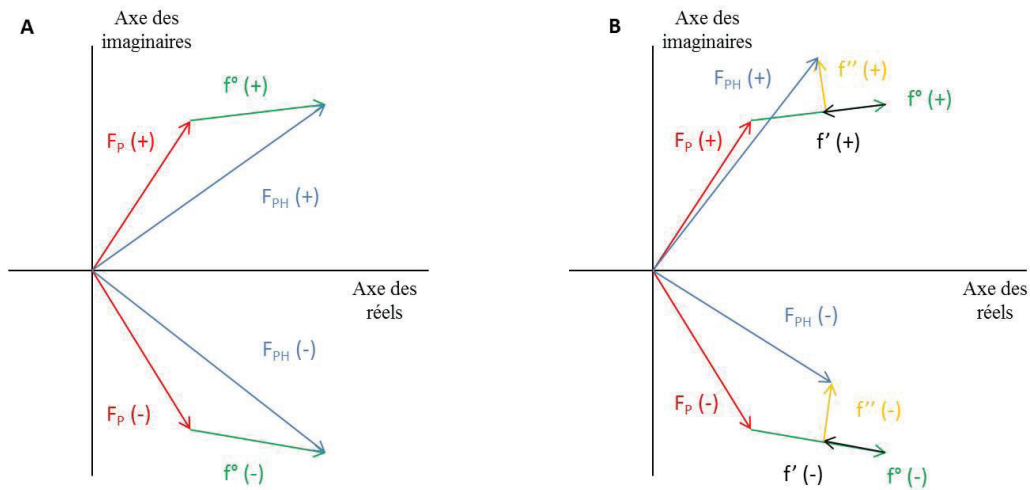


Figure 33 : Représentation vectorielle des facteurs de structure de la protéine (F_p) contenant un atome lourd ($f^o + f' + f''$). F_{PH} correspond au facteur de structure de la protéine et de l'atome lourd (A) à une longueur de d'onde en absence de diffusion anormale, et (B) à une longueur d'onde induisant une diffusion anormale.

2.3.8. Méthode SAD (Single wavelength anomalous dispersion)

Cette méthode consiste à utiliser la théorie de la diffusion anormale pour résoudre la structure de la protéine. Le principe consiste à déterminer la position de l'atome lourd dans la maille cristalline.

Les amplitudes de deux contributions réelle et imaginaire de la diffusion anormale sont constantes (à une longueur d'onde définie) pour un atome lourd donné. On utilise la spectroscopie EXAFS (Extended X-Ray Absorption Fine Structure) effectuant un scan de longueurs d'onde pour déterminer expérimentalement le seuil d'absorption de l'atome lourd. Avant de pouvoir obtenir les phases expérimentales, il est nécessaire de localiser la position des atomes lourds dans la maille. Trouver la position de quelques atomes est plus facile que de déterminer la position de tous les atomes de la maille cristalline. Pour cela, on utilise une transformée de Fourier des intensités diffractées : c'est la fonction de Patterson. Grâce à la spectroscopie EXAFS, on connaît l'amplitude des deux contributions réelle (f') et imaginaire (f'') de l'atome lourd. Les phases quant à elles ne dépendent uniquement que de la position dans la maille (déterminée par la fonction de Patterson). Il est donc maintenant possible de remonter au facteur de structure en absence de diffusion anormale : $f^o = f(\lambda) - f' - f''$ par la méthode vectorielle de Harker. Le SAD aboutit à une ambiguïté sur le choix des phases en absence de diffusion anormale car la méthode de Patterson ne permet pas de distinguer les positions énantiomériques des atomes lourds. Il y a donc deux cartes de densités

électroniques possibles correspondant au groupe d'espace du cristal ainsi que son énantiomère. Pour résoudre cette indétermination, un processus de modification de densité comme l'aplatissement de solvant est utilisé.

2.3.9. Résolution de la structure de la protéine

La détermination des phases par une méthode expérimentale permet d'éviter un biais possible introduit par un modèle de remplacement moléculaire. Il est possible que les premières cartes de densité électronique ne soient pas interprétables immédiatement. Il existe plusieurs méthodes visant à améliorer la qualité des phases expérimentales. Il est possible par exemple de réaliser un aplatissement de solvant ou de moyennner les densités électroniques reliées par symétrie non cristallographique (NCS). L'aplatissement de solvant permet de plus de lever l'indétermination sur la position des atomes lourds due à l'utilisation d'une seule longueur d'onde.

2.3.10. Construction du modèle et affinement

Une fois qu'une carte de densité électronique interprétable est obtenue, un modèle de la protéine peut-être construit. Il s'agit de placer les acides aminés de la protéine à l'intérieur de la carte de densité électronique. Ce modèle est approximatif au départ et il est amélioré par des cycles itératifs construction/modification du modèle et affinement. Les deux paramètres R et R_{free} permettent de suivre l'affinement. Ils comparent les facteurs de structure calculés du modèle avec les facteurs de structure expérimentaux. Ils s'expriment mathématiquement de la même façon :

$$R = R_{free} = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|}$$

La différence entre le facteur R et le facteur R_{free} est que ce dernier mesure l'accord entre un jeu d'amplitudes observées (F_{obs}) n'étant pas utilisées dans l'affinement (généralement 5% des réflexions) et celles obtenues avec le modèle (F_{calc}) (Brunger, 1992). L'écart entre les facteurs R et R_{free} est un moyen de déterminer si l'affinement du modèle est incorrect. Un écart faible (en général entre 5 % et 7%) entre ces deux facteurs signifie que le modèle est correctement affiné. Dans le cas d'un sur-affinement, le nombre d'observations n'est pas

suffisant par rapport au nombre de paramètres à affiner. Ainsi le facteur R aura tendance à diminuer alors que le facteur R_{free} augmentera (Kleywegt et Brunger, 1996) Il existe deux grandes méthodes d'affinement : l'affinement par minimisation d'un résiduel et le dynamisme moléculaire. Des méthodes d'affinement à basse résolution existent également comme par exemple la méthode d'affinement des corps rigides (rigid body) ou encore le « jelly body refinement ».

2.4. Etudes des protéines en solution

2.4.1. Diffusion des rayons X aux petits angles (SAXS)

La diffusion des rayons X aux petits angles (small angle X-ray scattering ou SAXS) est une technique d'étude structurale de molécules en solution à basse résolution. Cette technique est basée sur la diffusion des rayons X par les protéines et permet d'obtenir des informations sur la taille et la forme globale des macromolécules. En effet, si l'échantillon est monodisperse, c'est-à-dire si une seule espèce est présente, les données obtenues permettent d'estimer le rayon de giration (R_g), la distance maximale interatomique (D_{max}), ainsi qu'une estimation de la forme globale de la protéine grâce à la densité de diffusion de la solution ($p(r)$). Le SAXS permet également d'obtenir la masse moléculaire de la protéine étudiée par comparaison avec des protéines références comme la bovine serum albumine (BSA) ou le lysozyme. Il est également possible d'obtenir une enveloppe basse résolution de la protéine étudiée.

Lorsqu'une solution de macromolécules est placée sous un faisceau de rayons X de longueur d'onde λ , les atomes interagissent avec le rayonnement incident. Les atomes diffusent ensuite le rayonnement dans toutes les directions. Ces ondes réémises autour du faisceau direct sont ensuite enregistrées sur un détecteur. Dans le cas d'une diffusion élastique, c'est-à-dire si les photons incidents n'échangent pas d'énergie avec les atomes, le module du vecteur incident ($q_0 = 2\pi/\lambda$) est égal au module du vecteur diffracté (q_1) : $q_1 = q_0$. Le vecteur de diffusion \vec{q} est alors défini par $\vec{q} = \vec{q}_1 - \vec{q}_0$ avec $q = 4\pi \sin\theta/\lambda$ où 2θ est l'angle de diffusion et λ la longueur d'onde. L'amplitude de l'onde diffusée est la transformée de Fourier de la distribution des atomes (Feigin et Svergun, 1987):

$$A(\vec{q}) = \sum_{i=1}^N f_i(\vec{q}) \exp(i\vec{q} \cdot \vec{r}_i)$$

Néanmoins, le détecteur mesure des intensités et non pas des amplitudes. L'intensité mesurée est :

$$I(\vec{q}) \approx |A(\vec{q})|^2 = \sum_{i=1}^N \sum_{j=1}^N f_i(\vec{q}) f_j(\vec{q}) \exp(i\vec{q} \cdot (\vec{r}_i - \vec{r}_j))$$

Les molécules sont orientées aléatoirement dans la solution. L'intensité mesurée est donc la moyenne de toutes les intensités diffusées pour une molécule dans toutes les orientations moléculaires possibles :

$$I(\vec{q}) = \langle I(\vec{q}) \rangle$$

Pour obtenir l'intensité diffusée par les particules, il convient de soustraire l'intensité diffusée par le solvant à l'intensité diffusée par la solution (particules + solvant). Dans le cas d'un système monodisperse idéal, l'intensité mesurée peut s'écrire alors (Svergun et Koch, 2003 ; Jacques et Trehwella, 2010) :

$$I(\vec{q}) = \left| \int_{vp} (p(r) - p(s) \exp(i\vec{q} \cdot \vec{r})) d\vec{r} \right|^2$$

Avec $p(r)$ la densité de diffusion de la solution, $p(s)$ la densité de diffusion du solvant et vp le volume de la particule. Pour une solution monodisperse, on peut extraire le rayon de giration (R_g) de la protéine par la loi de Guinier (Guinier, 1938 ; Svergun et Koch, 2013):

$$I(\vec{q}) = I(0) e^{-\frac{q^2 R_g^2}{3}}$$

Avec $I(0)$, l'extrapolation de l'intensité à l'origine et R_g le rayon de giration. Le R_g est la distance quadratique moyenne de tous les points de la particule au centre de gravité pondérée par la masse diffusante en chaque point. Les objets sphériques auront donc des rayons de giration plus petits que des objets allongés de même masse. La loi de Guinier peut s'écrire également sous forme logarithmique :

$$\ln I(\vec{q}) = \ln I(0) - \frac{q^2 R_g^2}{3}$$

En pratique, cette équation permet de déterminer le $I(0)$ et le R_g par le graphique de Guinier ($\ln I(Q)$ sur q^2). Dans la zone de Guinier ($q \cdot R_g < 1,3$), l'intersection de la droite obtenue par régression linéaire avec l'axe des ordonnées donne le $I(0)$ et la pente de la droite ($-1/3 R_g^2$)

permet d'obtenir le R_g . Cette valeur est une première indication concernant la forme de la protéine.

La fonction de distribution des distances est une deuxième indication de la forme d'une protéine. La fonction $p(r)$ (ou facteur de forme) correspond à la distribution de toutes les distances interatomiques de la protéine. Elle est définie par :

$$p(r) = \frac{r^2}{2\pi^2} \int_0^\infty q^2 I(q) \frac{\sin qr}{qr} dr$$

La fonction $p(r)$ est la représentation des données de diffusion dans l'espace réel. Elle permet de représenter graphiquement les paramètres globaux de la protéine. Une simple inspection de la courbe donne une indication sur la forme de la protéine (Figure 34). Le diamètre maximal de la protéine (D_{max}) peut être déterminé grâce à la $p(r)$ qui est égale à zéro pour une valeur de q égale au D_{max} (Svergun et Koch, 2003).

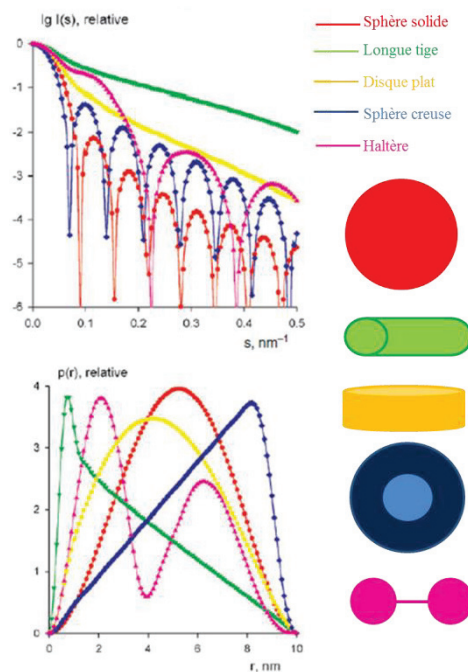


Figure 34 : Intensités diffusées et fonctions de distribution des distances de différentes formes géométriques (adapté de Svergun et Koch, 2003).

Le calcul de la $p(r)$ à partir des intensités $I(\vec{q})$ est possible grâce à des programmes utilisant les transformées de Fourier comme par exemple le programme GNOM (Svergun, 1992) utilisé dans cette thèse pour analyser les données SAXS. Les données de SAXS de 6His-

HpDnaB et HpDnaB seule ou avec HpDnaG^{HB}D, ATP, AMP-PNP et un oligonucléotide dT20 ont été enregistrées au synchrotron SOLEIL sur la ligne SWING. Des échantillons de 50 μ L ont été injectés dans du tampon 50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT et 5% glycérol sur une colonne d'exclusion stérique SHODEX KW405-4F branchée sur une HPLC Agilent à 15°C (David et Perez, 2009). Les échantillons après élution passent par un microcapillaire à 200 μ L/min pour la mesure SAXS. Les données ont été traitées en utilisant le programme de la ligne FOXTROT (<http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING>) et analysées en utilisant le logiciel PRIMUS (Konarev *et al.*, 2003). Les courbes théoriques issues des modèles PDB ont été générées et superposées aux données expérimentales par le serveur FoXS (<http://www.modbase.compbio.ucsf.edu/foxs/>) (Schneidman-Duhovny *et al.*, 2010, 2013).

2.4.2. Chromatographie d'exclusion de taille couplée à la diffusion de lumière multi-angles et réfractométrie (SEC-MALS)

Cette technique permet de déterminer le poids moléculaire d'une protéine en solution et par conséquent son état oligomérique. Quand un rayonnement électromagnétique rencontre de la matière, la plus grande partie de la radiation continue dans la direction originelle mais une petite fraction de cette radiation est diffusée. Le MALS (Multi-angle Light Scattering) repose sur deux principes de la diffusion statique de la lumière :

- La quantité de lumière diffusée est directement proportionnelle au produit de la masse molaire et de la concentration
- La variation d'intensité de lumière diffusée en fonction de l'angle de mesure est proportionnelle à la taille des molécules.

En mesurant la quantité de lumière diffusée ainsi que l'indice de réfraction via le dn/dc (l'incrément de l'indice de réfraction en fonction de la concentration), on peut calculer la masse d'une protéine en solution (Wyatt, 1993).

Les échantillons sont injectés sur une colonne d'exclusion stérique Superdex (GE Healthcare) S200 10/300GL ou SHODEX KW405 équilibrée avec du tampon contenant 50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT et 5% glycérol. Pour les expériences réalisées sur le complexe HpDnaB•HpDnaG^{HB}D, deux échantillons ont été analysés : (1) un mélange équimolaire des deux protéines purifiées séparément et incubées ensemble pendant 10 min et (2) un

échantillon issu de la co-expression et de la purification du complexe. L'analyse par MALS a été réalisée avec un détecteur TREOS-miniDAWN (Wyatt Technology, Santa Barbara, CA, USA) utilisant un laser émettant à 690 nm et en mesurant l'indice de réfraction grâce à un système Optilab T-REX (Wyatt Technology). Les masses molaires moyennes ont ensuite été calculées à l'aide du logiciel ASTRA (Wyatt Technology).

2.4.3. Titration calorimétrique isotherme (ITC)

La titration calorimétrique isotherme (ITC) est une méthode durant laquelle une molécule (M) est titrée par un ligand (L). La méthode permet de mesurer la différence d'enthalpie (ΔH) d'interaction. Les valeurs de la constante d'association K_a et de la stœchiométrie du complexe sont également déduites. Le ligand L est ajouté progressivement à une solution contenant le partenaire protéique (M) présent dans la cellule du microcalorimètre. Après chaque addition de L, la chaleur (consommée ou libérée) est mesurée par rapport à la cellule de référence contenant le tampon seul. La chaleur échangée est calculée à partir de la puissance électrique ($J s^{-1}$) nécessaire pour maintenir la cellule à la température de départ. L'addition de (L) est automatisée via une seringue que l'on place dans la cellule de mesure pilotée par un ordinateur. Les pics obtenus lors de l'expérience correspondent à la chaleur libérée ou consommée lors de l'addition de (L). L'intégration de ces pics en fonction du temps donne la chaleur apparente échangée et donne accès aux paramètres thermodynamiques du système (Wiseman *et al.*, 1989).

Les protéines 6His-HpDnaB ou HpDnaB sont placées dans la cellule de mesure à des concentrations de 110 μM et 50 μM , respectivement. Elles sont ensuite titrées par 16 injections de 2,5 μL de HpDnaG^{HBD} à 305 μM (pour 6His-HpDnaB) et 500 μM (pour HpDnaB). Les mesures sont réalisées à 25°C par un ITC 200 Microcal (GE Healthcare). Avant la mesure, les deux protéines sont dialysées toute la nuit contre un tampon contenant 50 mM Tris pH 8, 200 mM NaCl et 5% glycérol.

2.5. Etude de l'interaction ADN-Protéine

2.5.1. Expériences d'ouverture de la fourche de réplication

Le principe de cette expérience est d'utiliser l'origine de réplication de *H. pylori*, insérée dans un plasmide. Le plasmide est incubé avec HpDnaA ce qui provoque l'ouverture de l'origine en deux ADN simples brins. Afin de vérifier s'il y a eu ouverture, on utilise la nucléase P1 qui hydrolyse uniquement l'ADNsb. Le plasmide est ensuite coupé avec une deuxième enzyme de restriction. Si il y a eu ouverture de l'origine par HpDnaA, la nucléase P1 aura donc coupé le plasmide une première fois et l'enzyme de restriction une deuxième fois. On observe alors sur gel d'agarose deux bandes d'ADN.

L'essai d'ouverture à la nucléase P1 a été réalisé comme décrit dans Chodavarapu *et al.*, 2008 avec plusieurs modifications. Le volume de réaction est de 25 μ L dans du tampon 25 mM Hepes pH 7,6, 1 mM CaCl₂, 0,2 mM EDTA, 12% glycérol, 5 mM ATP et 0,1 mg/mL de BSA, 200 ng de plasmide por1-ori2 et 200 ng de HpDnaA. Les échantillons sont incubés 15 min à 30°C. La nucléase P1 est rajoutée (0,3 unités dans 25 mM Hepes pH 7,6, 1 mM EDTA, 15% glycérol, 4 mM DTT) et les échantillons sont incubés 5 min à 30°C. La réaction est stoppée par l'addition de 85 μ L d'eau et de 300 μ L de tampon QG (Qiagen). L'ADN est immédiatement purifié sur des colonnes de purification d'ADN QIAquick suivant le protocole du kit de purification (Qiagen) puis élué avec 50 μ L de tampon EB (10 mM Tris pH 8,5) du kit dilué 4 fois. Les échantillons sont séchés par centrifugation sous vide pour atteindre un volume de 12 à 10 μ L. L'ADN est finalement digéré par EcoR1 (Thermo scientific) et chargé sur un gel d'agarose 1%.

2.5.2. Expériences de chargement de l'hélicase HpDnaB sur l'ADN

Classiquement, un plasmide non digéré migre sur gel d'agarose de trois façons différentes. La forme dite superenroulée (forme F1) migre plus rapidement à travers les mailles du gel dû à sa conformation superenroulée. Au-dessus on retrouve la forme du plasmide correspondant à la taille linéaire (forme F3, même taille que le plasmide coupé une fois). Enfin la bande la plus haute correspond à la forme circulaire et/ou nicked (forme F2) c'est-à-dire que le plasmide est coupé une fois sur un simple brin. Cette expérience consiste à mettre en évidence une quatrième forme, une forme qui migre encore plus loin que la forme superenroulée. Cette bande correspond à une forme circulaire fermée et très déroulée (forme

F1*) (Figure 35). Dans notre cas, cela signifie que l'hélicase replicative s'est chargée sur le plasmide comportant l'origine de réplication de *H. pylori*.

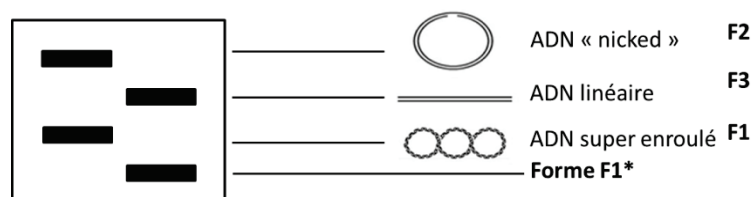


Figure 35 : Représentation schématique des différentes formes d'un plasmide lors d'une migration sur gel d'agarose 1% après chargement de l'hélicase DnaB.

Les essais de chargement sont réalisés dans du tampon 40 mM Hepes pH 8, 25 mM Tris pH 7,4, 11 mM acétate de magnésium, 4 mM DTT, 4% sucrose, 2 mM ATP, 80 µg/mL BSA. Le volume total de la réaction est de 25 µL. Les protéines utilisées pour cet essai sont HpDnaA (3,86 pmoles), HobA (0,2 pmoles), la gyrase (1,3 pmoles), HU (0,1 pmoles), HpDnaB (17 pmoles), HpDnaG^{HBD} (51 pmoles). Les échantillons sont incubés 30 min à 32°C. La réaction est stoppée en rajoutant 5 µL d'un mélange 10 mM EDTA et 2% SDS. Les échantillons sont chauffés à 65°C pendant 2 min et sont chargés sur un gel 1% agarose.

2.5.3. Mesure d'activité ATPase

Les mesures d'activité ATPase effectuées pendant la thèse reposent sur la réaction de la voie métabolique de la fermentation lactique. On utilise donc une molécule, le phosphoénolpyruvate (PEP), avec deux enzymes, la pyruvate kinase et la lactate déshydrogénase (PK/LDH). La pyruvate kinase convertit une molécule d'ADP en ATP en transformant une molécule de PEP en pyruvate. La lactate déshydrogénase utilise ensuite une molécule de nicotinamide adénine dinucléotide oxydé (NADH) pour convertir une molécule de pyruvate en une molécule de lactate ou acide lactique (Figure 36). Etant donné que la consommation de NADH et la formation d'ATP sont stœchiométriques au cours de la réaction, on peut mesurer l'activité ATPase d'une protéine en mesurant par spectrophotométrie la diminution d'absorbance due à la conversion de NADH en NAD au cours de la réaction.

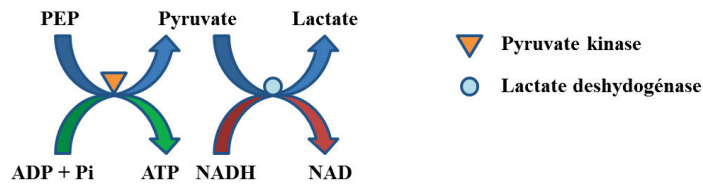


Figure 36 : Représentation schématique des réactions mises en jeu lors de la mesure de l'activité ATPase par le couplage des enzymes pyruvate kinase et lactate deshydrogénase.

Les tests ATPase ont été réalisés à 37°C dans du tampon 50 mM Tris pH 8, 100 mM NaCl, 5 mM MgCl₂. La réaction est complétée avec 0,5 mM de NADH, 2 mM PEP, 2,5 unités de pyruvate kinase, 1,6 unités de lactate deshydrogénase ainsi que 5 mM d'ATP. Pour les tests avec l'ADN, 500 μM d'oligonucléotide simple brin composé de vingt thymidines a été rajouté. La réaction est incubée 5 min à 37°C avant d'ajouter les protéines respectives (HpDnaB et HpDnaG^{HBD} à 500 μM chacune). Après le rajout des protéines, la décroissance d'absorbance du NADH a été mesurée dans un spectrophotomètre (Jasco) à une longueur d'onde de 340 nm pendant 120 sec. La conversion de l'ATP en ADP couplée à la réduction du NADH en NAD est stœchiométrique et calculée en utilisant l'équation suivante :

$$\text{Activité ATPase } (\mu\text{M}/\text{sec}) = \frac{(\Delta\text{abs } 340 \text{ nm}) \div (\Delta\text{temps})}{(\text{longueur de la cuve cm}) \times (6220 \text{ M}\cdot\text{cm}^{-1})}$$

Avec 6220 M·cm⁻¹, le coefficient d'extinction du NADH à 340 nm.

2.6. Microscopie électronique à coloration négative

La microscopie électronique à transmission (ME) est une technique qui utilise un faisceau d'électrons à haute énergie pour obtenir une image d'objets de faible épaisseur. La résolution correspond à la plus petite distance en dessous de laquelle deux points ne seront plus distingués. Cette distance est proportionnelle à la longueur d'onde utilisée pour observer l'objet étudié. L'hypothèse de De Broglie, formulée en 1924, prédit que les particules doivent aussi se comporter comme des ondes. La longueur d'onde d'un électron étant plus petite que celle d'un photon, les microscopes électroniques ont un plus grand pouvoir de résolution que les microscopes optiques. En ME, le faisceau d'électrons généré traverse l'échantillon et produit deux types d'interactions avec ce dernier : des interactions inélastiques et des interactions élastiques. Lors d'une diffusion élastique, les électrons incidents interagissent

avec le noyau des atomes et sont ainsi déviés sans perte d'énergie. Lors d'une diffusion inélastique, les électrons incidents percutent un électron de la couche externe d'un atome et sont légèrement déviés. Cette interaction induit une perte d'énergie, la formation de radicaux libres et peut alors endommager l'échantillon. Seuls, les électrons traversant les échantillons sont utilisés pour l'analyse en ME (Colliex, 1996).

En ME la formation d'une image haute résolution provient essentiellement des électrons diffusés élastiquement. Les électrons diffusés inélastiquement induisent du bruit de fond et peuvent être éliminés par l'utilisation de filtres. Comme dit précédemment, un faisceau d'électrons peut être considéré comme une onde. En traversant l'échantillon, les électrons diffusés élastiquement ont un décalage de phase par rapport aux électrons incidents (on parle de contraste de phase). Les électrons diffusés inélastiquement ayant perdu de l'énergie, l'amplitude de l'onde est diminuée (on parle de contraste d'amplitude). En ME, il est possible de modifier le contraste grâce à un diaphragme de contraste situé après l'échantillon et permet de sélectionner les électrons diffusés. Afin de visualiser l'image de l'objet étudié, les électrons sont captés par une lentille électromagnétique qui comme une lentille optique possède des aberrations: l'astigmatisme, l'aberration sphérique et l'aberration chromatique. Le faisceau d'électrons est ensuite projeté sur un écran phosphorescent grâce à des lentilles magnétiques transformant l'image électronique en image optique. Les images peuvent également être collectées sur un film argentique ou sur une caméra digitale. Elles sont par la suite traitées afin de corriger les aberrations induites par les lentilles ainsi que la fonction de transfert de contraste grâce à des programmes informatiques (Reimer, 1997).

2.6.1. Microscopie électronique à coloration négative

Le matériel biologique est essentiellement composé d'atomes légers (carbone, hydrogène, oxygène,...), et possède par conséquent un faible pouvoir diffuseur. En ME à coloration négative, on utilise un sel d'atome lourd comme l'acétate d'uranyl, qui a un fort pouvoir diffuseur et permet d'augmenter le contraste des échantillons biologiques. Soumis au faisceau d'électrons, le colorant d'atome lourd dévie donc davantage les électrons, L'échantillon apparaîtra plus clair comparé au bruit de fond sur l'image finale d'où le nom de coloration négative.

2.6.2. Préparation de l'échantillon

La protéine HpDnaB est diluée au préalable à 0,1 mg/mL, puis quelques microlitres sont placés entre une couche de carbone préalablement hydrophylisé et une couche de mika. L'ensemble est ensuite placé dans une goutte de solvant. La couche de mika tombe au fond de la goutte et une plaque de cuivre est adsorbée sur l'échantillon. L'excès de solvant est absorbé par un papier filtre et l'ensemble est placé dans une goutte d'acétate d'uranyl pendant quelques secondes. Le colorant en excès est retiré grâce à un papier filtre et la plaque est enfin séchée à l'air libre.

3. Résultats

3.1. Préparation et purification des protéines et complexes étudiés

Le protocole de purification de 6His-HpDnaB (Stelter *et al.*, 2012) consiste en une étape de purification sur colonne de nickel, puis sur colonne échangeuse d'anions et enfin sur colonne d'exclusion stérique. Les tampons utilisés dans le protocole contenaient du Tris 50 mM pH 7,5. Avec la protéine issue de ce protocole de purification, nous n'avons obtenu que des cristaux ne diffractant pas ou à très basse résolution ($< 9\text{\AA}$) (voir partie 3.2.1). Nous avons donc modifié le protocole et purifié la protéine dans un tampon à base de phosphate de sodium. Les essais de cristallisation ont été réitérés et ont permis l'apparition d'une nouvelle forme cristalline. La purification de 6His-HpDnaB utilisée pour la production de ces cristaux est donc décrite ci-dessous.

3.1.1. Purification de 6His-HpDnaB

Après lyse bactérienne par sonication et centrifugation, le surnageant est injecté sur une colonne de nickel. La protéine soluble s'accroche à la colonne de nickel et une partie des contaminants cellulaires sont éliminés grâce au lavage avec du tampon A (10 mM phosphate de sodium pH 7,5, 200 mM NaCl, 2 mM β ME). Le lavage avec 10% (v/v) de tampon B (10 mM phosphate de sodium pH 7,5 ; 200 mM NaCl ; 2 mM β ME et 500 mM imidazole) ne retire quant à lui que peu de contaminants. Lors de l'élution, un pic majoritaire d'absorbance à 280 nm contenant 6His-HpDnaB est observé (Figure 37).

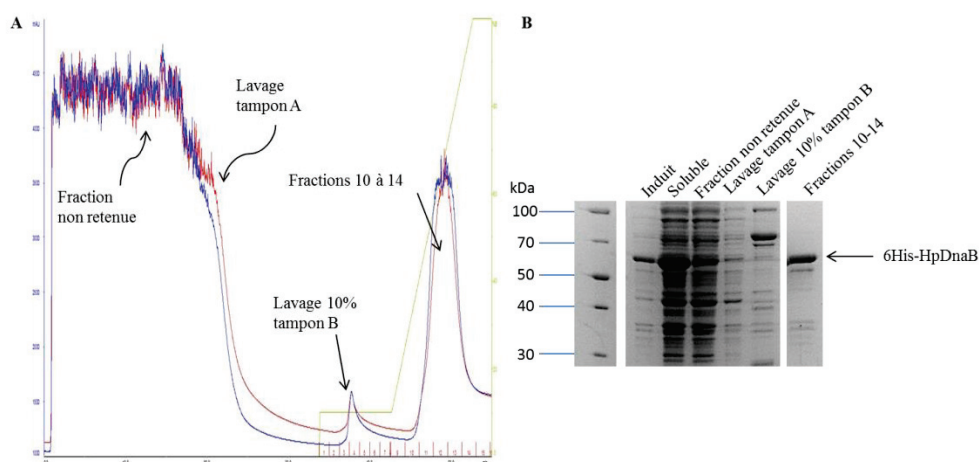


Figure 37 : Purification de 6His-HpDnaB sur colonne d'affinité au nickel. (A) Chromatogramme d'absorbance à 280 nm (bleu), à 254 nm (rouge) et du gradient d'élution (vert). (B) Analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (kDa) est indiqué sur la première piste du gel.

Les fractions contenant 6His-HpDnaB sont regroupées et injectées sur une colonne d'exclusion stérique (Superdex S200) équilibrée avec le tampon final de la protéine : 10 mM phosphate de sodium, 200 mM NaCl et 2 mM β ME. On observe un pic d'absorbance à 280 nm. Le gel SDS indique que les fractions de ce pic contiennent HpDnaB et quelques contaminants. Seules les fractions correspondant au maximum d'absorption et contenant le moins de contaminants sont regroupées et concentrées à 10 mg/mL. La protéine peut être conservée à -20°C pour une utilisation ultérieure (Figure 38).

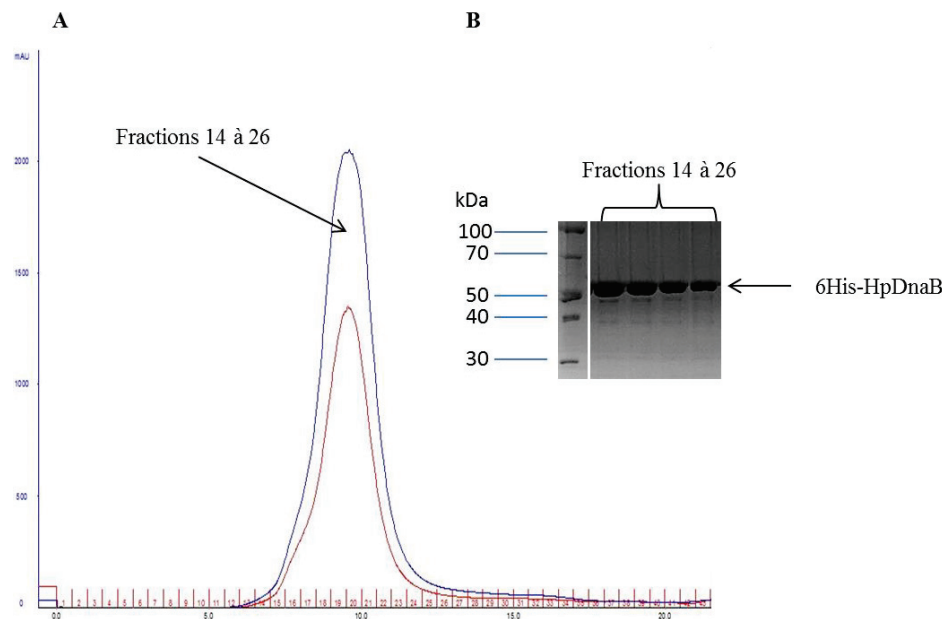


Figure 38 : Purification de 6His-HpDnaB sur colonne d'exclusion stérique. (A) Chromatogramme d'absorbance à 280 nm (bleu) et à 254 nm (rouge). (B) Analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (kDa) est indiqué sur la première piste du gel

3.1.2. Purification de HpDnaB

Après lyse bactérienne et centrifugation, la fraction soluble est injectée sur une colonne d'héparine. L'héparine permettant de mimer la fixation à l'ADN, les protéines ayant une affinité pour les acides nucléiques peuvent se fixer. Lors de l'élution, un premier pic à 280 nm apparaît suivi de plusieurs autres. L'analyse des fractions sur gel SDS montre que la protéine s'est fixée à la colonne et une partie des contaminants cellulaires ont été éliminés dans la fraction non retenue par la colonne. Les fractions du premier pic apparu lors de l'élution (entre 15% et 30% (v/v) de tampon B : 10 mM phosphate de sodium pH 7,5 ; 2 mM β ME ; 1 M NaCl) correspondent à HpDnaB et les pics suivants correspondent à des contaminants. Les fractions contenant HpDnaB sont regroupées et injectées sur une colonne échangeuse d'anions (Figure 39).

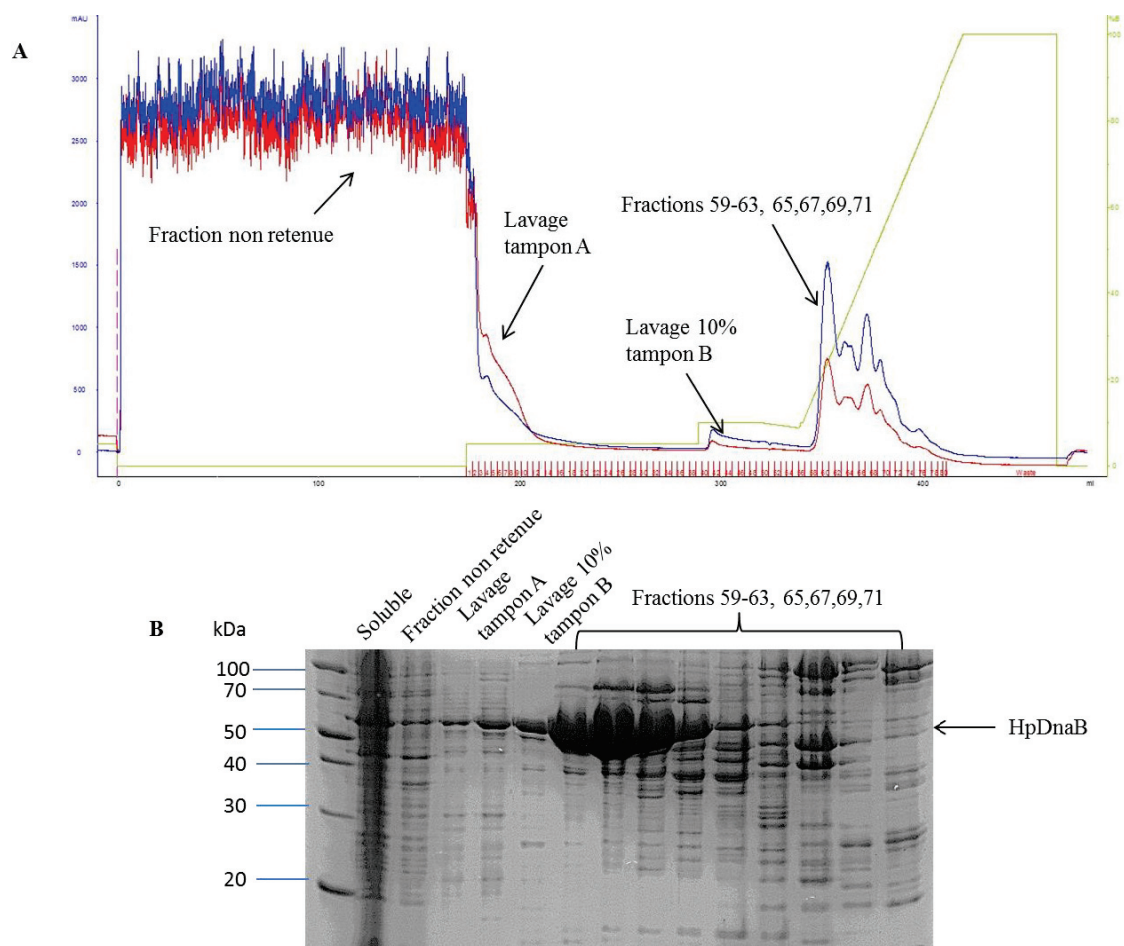


Figure 39 : Purification de HpDnaB sur colonne d'héparine. (A) Chromatogramme d'absorbance à 280 nm (bleu), à 254 nm (rouge) et du gradient d'éluion (vert). (B) Analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire est indiqué sur la première piste du gel.

Après avoir injecté HpDnaB sur la colonne échangeuse d'anions, la colonne est lavée avec du tampon A (10 mM phosphate de sodium pH 7,5, 2 mM β -mercaptoéthanol). Le point isoélectrique de HpDnaB étant de 5,20, la protéine se fixe à la colonne puisqu'à un pH 7,5, HpDnaB est chargée négativement. Un deuxième lavage à 10% de tampon B est réalisé, suivi du gradient d'éluion. Les premiers pics à 280 nm qui apparaissent au début de l'éluion correspondent à des contaminants et le pic majoritaire correspond à HpDnaB (éluion entre 45% et 55% (v/v) de tampon B). La colonne échangeuse d'anions a ainsi permis d'améliorer significativement la pureté de notre protéine (Figure 40).

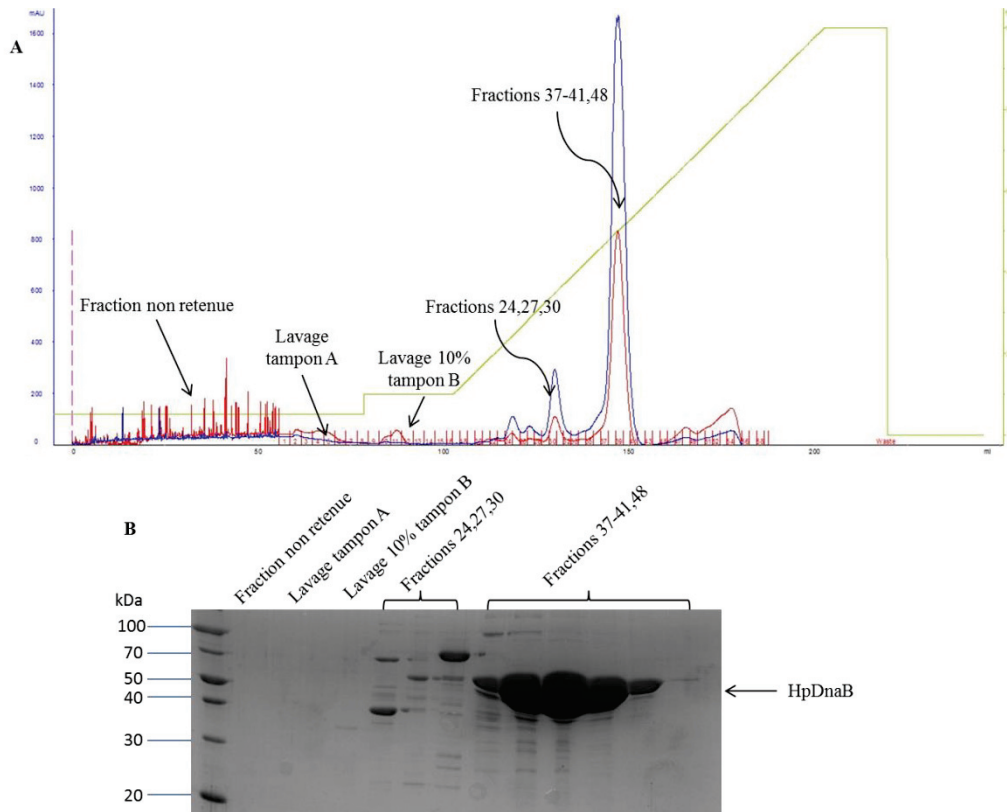


Figure 40 : Purification de HpDnaB sur colonne échangeuse d’anions. (A) Chromatogramme d’absorbance à 280 nm (bleu), à 254 nm (rouge) et du gradient d’élution (vert). (B) Analyse des fractions sur gel SDS coloré au bleu de Coomassie avec le marqueur de poids moléculaire (en kDa) indiqué sur la première piste du gel.

3.1.3. Purification de HpDnaG^{HBD}

La fraction soluble obtenue après lyse bactérienne et centrifugation est injectée sur une colonne de nickel. Une partie des contaminants cellulaires sont éliminés par le lavage avec le tampon A (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT et 5% glycérol). Certains contaminants accrochés à la colonne sont ensuite éliminés avec le lavage à 10% (v/v) de tampon B (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 5% glycérol et 500 mM Imidazole) et HpDnaG^{HBD} est éluée avec le gradient de tampon B (élution à environ 50% (v/v) de tampon B) (Figure 41). Les fractions contenant HpDnaG^{HBD} les plus pures sont rassemblées et l’étiquette histidine est coupée par la TEV.

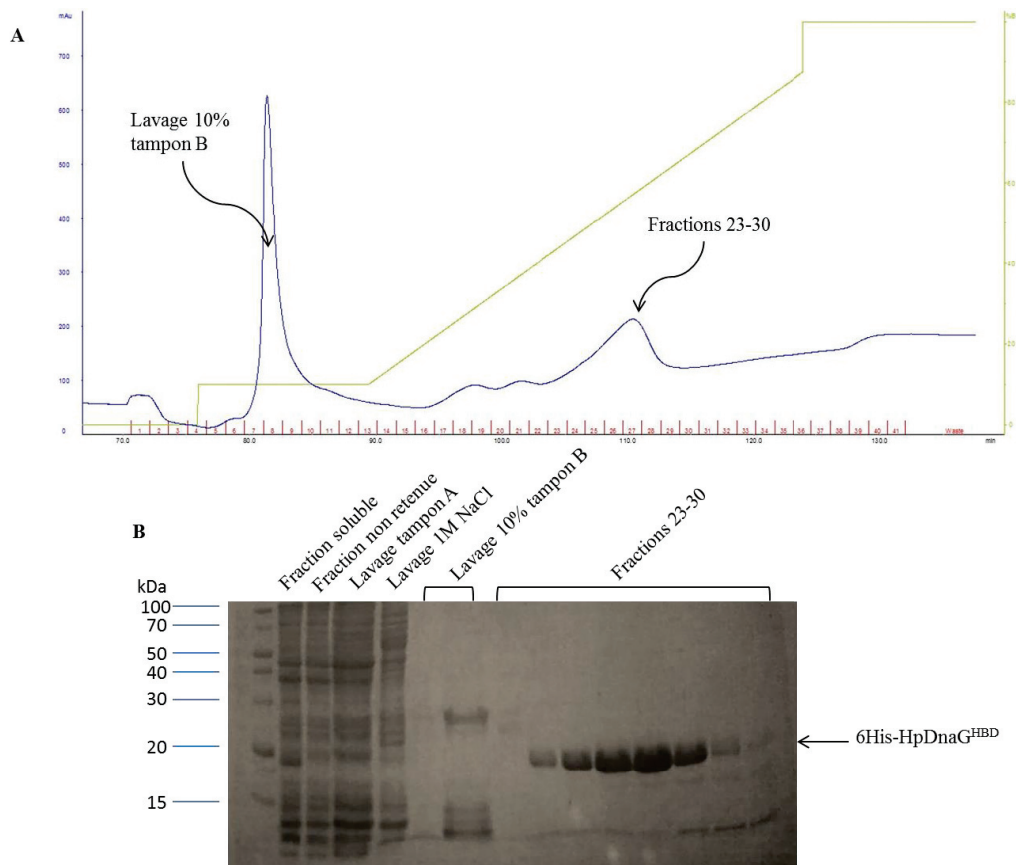


Figure 41 : Purification de HpDnaG^{HBD} sur colonne d'affinité au nickel. (A) Chromatogramme d'absorbance à 280 nm (bleu) et du gradient d'éluion (vert). (B) L'analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

Après coupure de l'étiquette histidine par la TEV protéase, la protéine est séparée de l'étiquette et de la TEV en utilisant une nouvelle fois la colonne d'affinité au nickel. Bien que la protéine ne possède plus d'étiquette histidine, celle-ci s'accroche à la colonne mais elle peut être éluée avec 5% (v/v) de tampon B. La protéine est séparée de la TEV puisque celle-ci se fixe à la colonne et est éluée entre 40% et 70% (v/v) de tampon B (Figure 42). La protéine est rapidement concentrée à 10 mg/mL et congelée à -20°C.

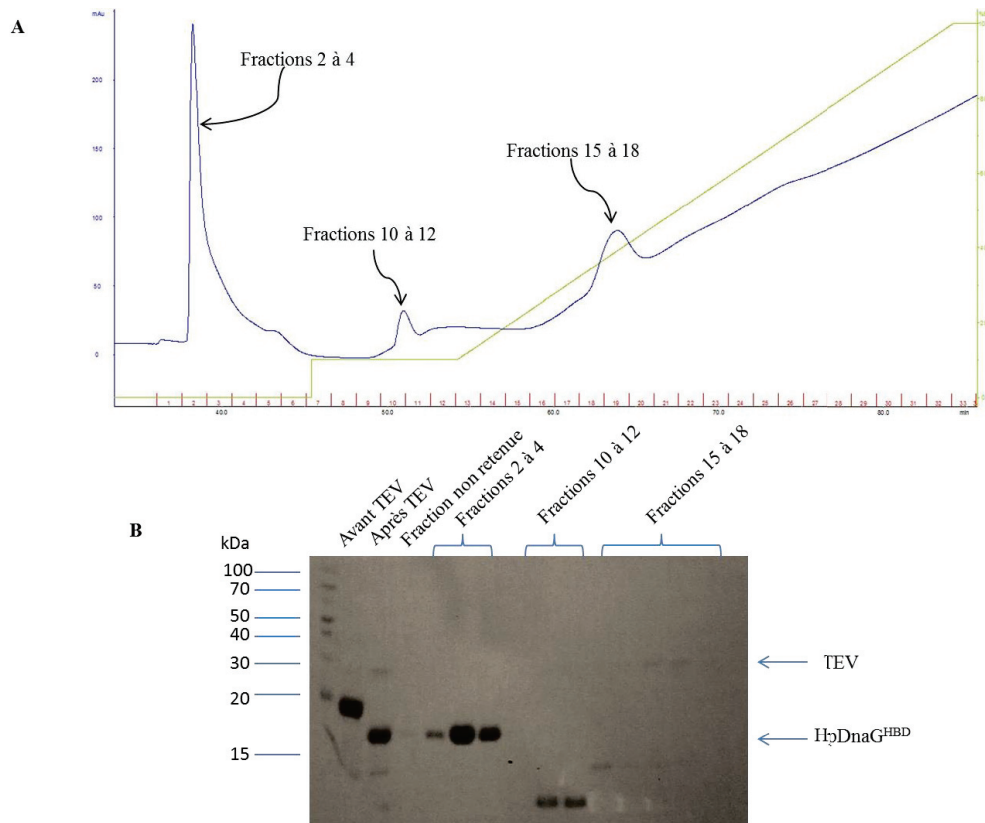


Figure 42 : Purification de HpDnaG^{HBD} sur colonne d'affinité au nickel après coupure à la TEV protéase. (A) Chromatogramme d'absorbance à 280 nm (en bleu) et du gradient d'éluion (vert). (B) Analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.1.4. Purification du complexe HpDnaB•HpDnaG^{HBD}

Pour purifier le complexe, HpDnaB et HpDnaG^{HBD} ont été exprimées simultanément dans la bactérie. Seule HpDnaG^{HBD} possède une étiquette histidine ; le complexe est donc purifié par co-éluion des deux protéines sur une colonne d'affinité au nickel. Le complexe se fixe à la colonne et une partie des contaminants cellulaires sont éliminés dans la fraction non retenue. Il n'y a pas de lavage avec 1 M de NaCl car ce lavage déstabilise le complexe et sépare les deux protéines. Un premier lavage avec 10% (v/v) de tampon B (50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT, 5% glycérol et 500 mM Imidazole) permet d'éliminer des contaminants accrochés à la colonne. Le complexe est ensuite élué avec un gradient de tampon B. L'analyse des fractions sur gel SDS montre qu'il y a un excès de HpDnaB produit par rapport à HpDnaG^{HBD} et que cet excès est éliminé par le lavage avec 10% (v/v) de tampon B. HpDnaB ne possédant pas d'étiquette histidine, seule la protéine en complexe avec HpDnaG^{HBD} est éluée lors du gradient de tampon B. Le complexe élué est pur et sa stoechiométrie sur gel semble être d'une molécule de HpDnaB pour une molécule de HpDnaG^{HBD} (Figure 43).

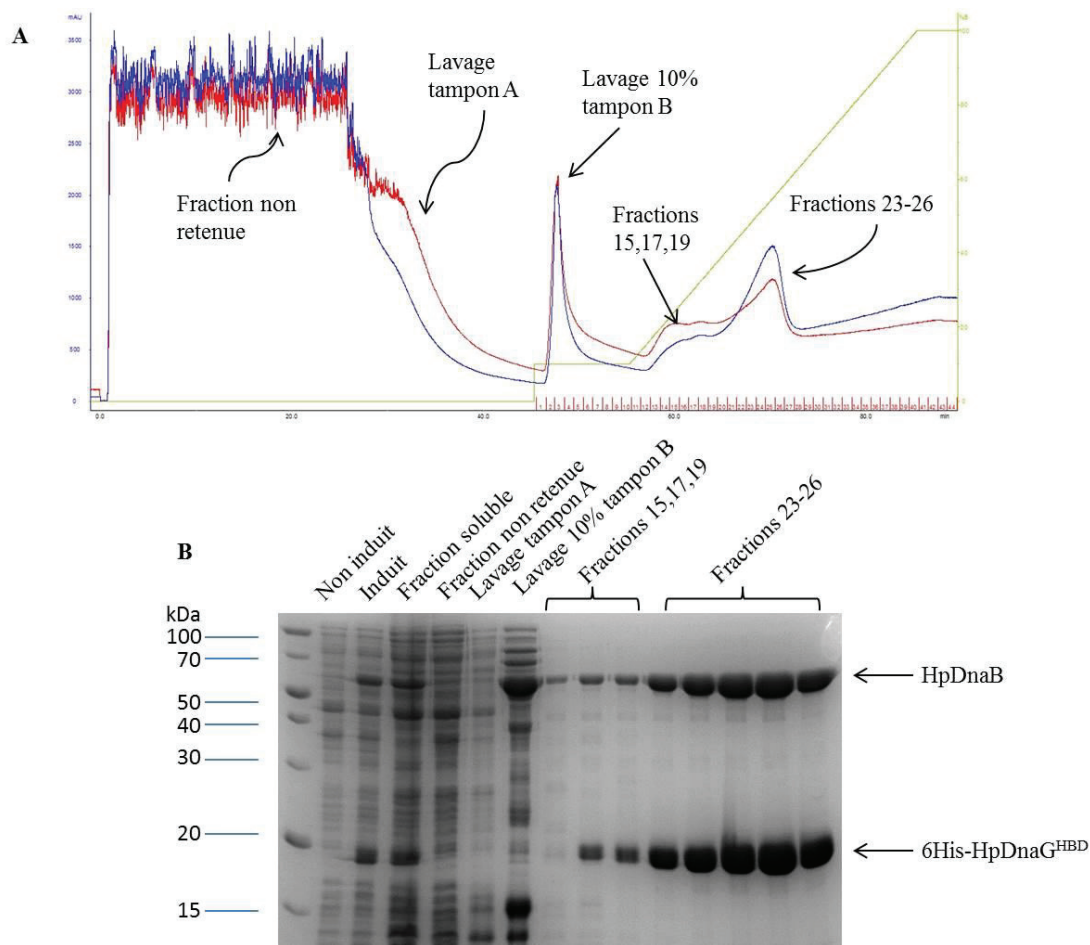


Figure 43 : Purification du complexe HpDnaB•HpDnaG^{HBD} sur colonne d'affinité au nickel. (A) Chromatogramme d'absorbance à 280 nm (bleu), à 254 nm (rouge) et du gradient d'éluion (vert). (B) L'analyse des fractions sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.1.5. Purification de HpDnaA

Lors de mon arrivée au laboratoire, le clonage et le protocole de purification de HpDnaA avaient déjà été mis au point. Le vecteur d'expression de HpDnaA permet d'exprimer une protéine sans étiquette mais la protéine est capable de se fixer avec une faible affinité sur une résine de nickel. Après la lyse bactérienne, la fraction soluble est donc injectée sur une colonne d'affinité au nickel. Cela permet d'éliminer une partie des contaminants cellulaires dans la fraction non retenue par la colonne et la protéine est éluée très tôt avec un lavage à 10% (v/v) de tampon B (45 mM Hepes pH 7,5, 10 mM d'acétate de magnésium, 100 mM de Monopotassium L-glutamate monohydrate, 1 mM DTT, 10% sucrose, 500 mM Imidazole). Les fractions sur gel SDS montrent qu'il reste encore des contaminants présents avec HpDnaA qui reste cependant majoritaire (Figure 44).

Pour améliorer la pureté, les fractions contenant HpDnaA sont assemblées et injectées sur une colonne d'héparine. Les fractions sur gel SDS montrent que la protéine s'est bien attachée à la colonne. Les lavages en tampon C (45 mM Hepes pH 7, 10 mM d'acétate de magnésium, 1 mM DTT et 100 mM de monopotassium L-glutamate monohydrate) et avec 10% (v/v) de tampon D (45 mM Hepes pH 7, 10 mM d'acétate de magnésium, 1 mM DTT et 100 mM de monopotassium L-glutamate monohydrate, 1 M NaCl) retirent quelques contaminants et la protéine est éluée avec le gradient de tampon D. Le gradient d'éluion permet d'améliorer la pureté de la protéine car les contaminants sont élués plus tôt que HpDnaA pendant le gradient d'éluion (Figure 44).

Les fractions contenant HpDnaA sont une nouvelle fois assemblées, diluées jusqu'à atteindre une concentration de 50 mM de NaCl et injectées sur une colonne échangeuse d'anions. La majorité de la protéine ne se fixe pas sur la colonne et se retrouve dans la fraction non retenue. La colonne échangeuse d'anions n'améliore pas la pureté de la protéine mais permet d'éliminer beaucoup d'ADN (Figure 44).

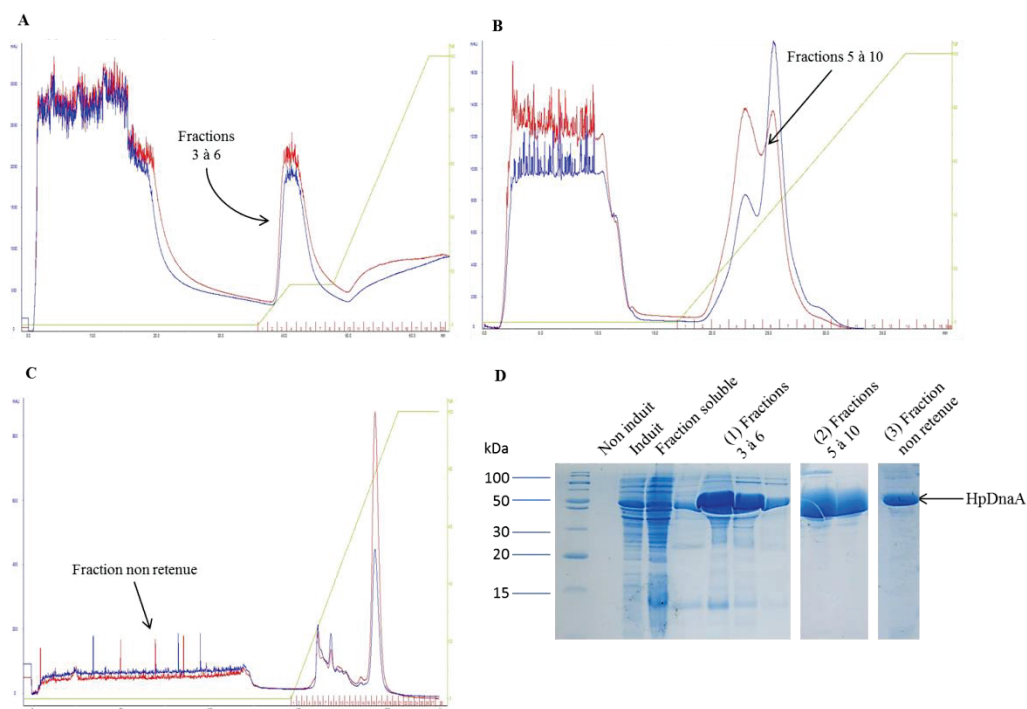


Figure 44 : Purification de HpDnaA. (A) Chromatogrammes de la purification sur colonne d'affinité au nickel. (B) Chromatogrammes de la purification sur colonne héparine. (C) Chromatogrammes de la purification sur colonne échangeuse d'anions. Les courbes d'absorbance à 280 nm sont représentées en bleu, celles à 254 nm en rouge et les courbes vertes correspondent aux gradients d'éluion. (D) L'analyse des fractions sur gels SDS des trois purifications colorés au bleu de Coomassie avec en (1) les fractions contenant HpDnaA de la colonne d'affinité au nickel, (2) de la colonne héparine et (3) de la colonne échangeuse d'anions. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.1.6. Purification de la protéine HU

Après lyse bactérienne, la fraction soluble est injectée sur une colonne d'affinité au nickel. Le lavage avec 1M NaCl permet d'éliminer une quantité importante d'ADN ainsi que des contaminants. D'autres contaminants sont éliminés grâce au lavage avec 10% (v/v) de tampon B (40 mM Tris pH 8, 250 mM NaCl, 5% glycérol et 500 mM Imidazole) et lors du gradient d'élution. La protéine HU est éluée avec des contaminants lors du gradient d'élution comme le montre le gel SDS. Les fractions contenant HU sont regroupées et l'étiquette histidine est clivée par action de la TEV protéase.

La protéine est de nouveau injectée sur la colonne d'affinité au nickel pour séparer la TEV protéase et HU. Lors de l'élution un pic important à 280 nm apparaît. Ce pic correspond à la protéine HU dont l'étiquette histidine n'a pas été clivée. Celle-ci s'est donc fixée sur la colonne d'affinité au nickel comme le montre les fractions du pic sur gel. Néanmoins, on retrouve une petite quantité de HU pure sans étiquette histidine dans les fractions non retenues. Les fractions contenant HU avec son étiquette histidine ont été récupérées et incubées de nouveau avec la TEV protéase. Néanmoins, seule une petite quantité de la protéine a pu être séparée de l'étiquette histidine. La TEV protéase ne semble donc pas pouvoir cliver la totalité de la protéine purifiée. HU formant des dimères, il est probable que les étiquettes des deux sous-unités soient peu accessibles dans ce contexte (Figure 45). Cependant nous avons tout de même obtenu une quantité de HU sans étiquette histidine pure. La protéine a été concentrée et stockée à -20°C pour une utilisation ultérieure lors des essais d'ouverture de l'origine de réplication et des essais de chargement de HpDnaB (voir partie 3.5.1 et 3.5.2).

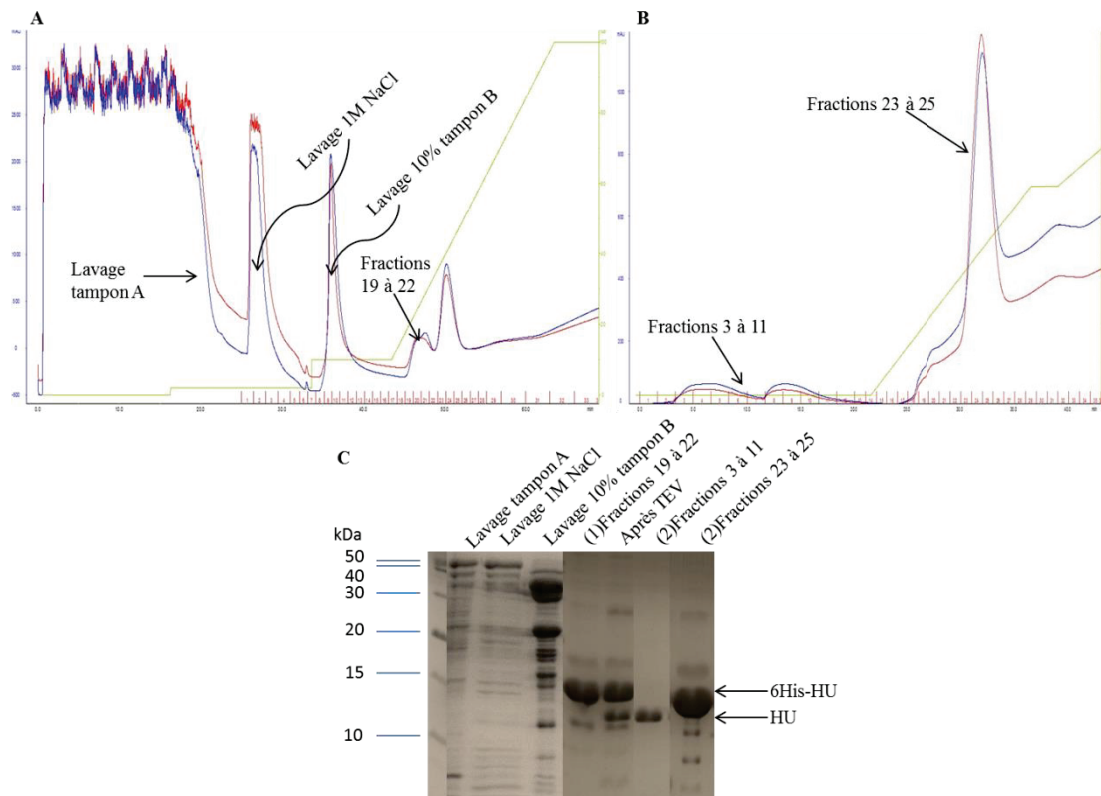


Figure 45 : Purification de la protéine HU. (A) et (B) Chromatogrammes d'absorption à 280 nm (bleu), à 254 nm (rouge) et des gradients d'éluion (vert). (C) Analyse de la purification de HU sur gels SDS colorés au bleu de Coomassie. Les fractions (1) et (2) contiennent HU avant et après coupure à la protéase TEV. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.1.7. Purification de HpDnaB^{NTD}

Le domaine NTD de HpDnaB (HpDnaB^{NTD}) a été purifié selon le protocole de purification de Kashav *et al.*, 2010. Celle-ci a été réalisée sur une colonne d'affinité et échangeuse d'anions. Les bactéries ont été lysées et le surnageant a été injecté sur une résine de nickel. La résine a ensuite été lavée avec du tampon A (50 mM Tris pH 8, 300 mM NaCl, 1 mM Imidazole), 10% (v/v) de tampon B (50 mM Tris pH 8, 150 mM NaCl, 300 mM Imidazole) et la protéine a été éluée avec 100% (v/v) de tampon B. L'analyse des fractions sur gel SDS indique qu'une quantité non négligeable de la protéine a été éluée avec le lavage à 10% de tampon B. La fraction correspondant à ce lavage a donc été conservée avec les fractions de l'éluion. Celles-ci ont été rassemblées puis la concentration imidazole ainsi que celle du NaCl ont été réduites par dialyse. La protéine a été injectée sur une résine échangeuse d'anions qui a été lavée avec du tampon A et la protéine a été éluée. L'analyse des fractions sur gel SDS montre que la pureté de la protéine a légèrement augmenté. Pour finir la purification, l'étiquette histidine a été clivée avec la TEV protéase sous dialyse à 4°C. La

protéine a été ensuite déposée sur la résine de nickel pour la séparer de la TEV protéase. L'analyse des fractions montre que l'étiquette histidine a bien été clivée et la protéine se retrouve dans la fraction non retenue par la résine. Ces fractions sont donc récupérées et la protéine a été concentrée jusqu'à 10 mg/mL et conservée à -20°C (Figure 46).

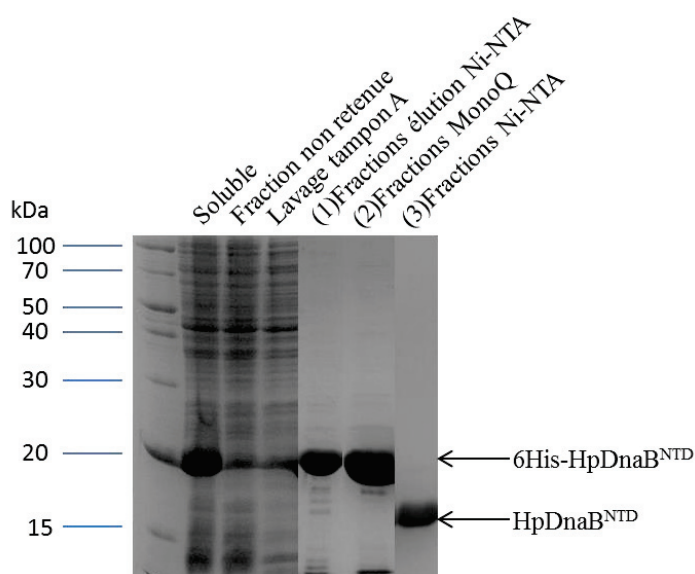


Figure 46 : Analyse des fractions sur gel SDS de la purification de HpDnaB^{NTD} coloré au bleu de Coomassie. Avec en (1) HpDnaB^{NTD} après la première colonne d'affinité (Ni-NTA), en (2) HpDnaB^{NTD} après la colonne échangeuse d'anions (MonoQ) et en (3) HpDnaB^{NTD} après la deuxième colonne d'affinité (Ni-NTA) et dont le tag histidine a été clivé par la protéase TEV. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.2. Structure cristallographique de HpDnaB

3.2.1. Cristallisation de 6His-HpDnaB

Les premiers tests de cristallisation ont été effectués avec la protéine purifiée en tampon Tris. Un criblage large a d'abord été réalisé en utilisant trois jeux de solutions commerciales : JCSG +, PEGion/Natrix et Cristal screen ½. Des microcristaux sont apparus dans la condition B5 de kit Cristal screen ½ contenant 30% de poly-éthylène glycol (PEG) 4000, 0,2 M LiSO₄ 0,1 M Tris pH 8,8. A partir de ces conditions initiales nous avons optimisé les cristaux. Ces expériences ont été réalisées dans des boîtes 24 puits en gouttes suspendues ou assises en mélangeant 1 µl de solution de cristallisation et 1 µl de protéine à 10mg.mL⁻¹.

Au fil des expériences, deux formes cristallines sont apparues. La première a été obtenue à partir de solutions contenant 24% à 27% de PEG 5KMME, 0,2M LiSO₄ et 0,1M Tris pH 8,0 à 8,8. Ces cristaux avaient une forme hexagonale en plaques (Figure 47). Afin d'obtenir des

cristaux en trois dimensions, 96 additifs ont été testés manuellement. Des cristaux plus épais sont ainsi apparus dans la condition de cristallisation contenant 26% de PEG 5KMME, 0,2 M LiSO₄, 0,1 M Tris pH 8,8 et 0,1 M spermine tetrahydrochloride (Figure 47). Cependant ces cristaux ne diffractaient que très peu (résolution < 9 Å) et de manière anisotrope.

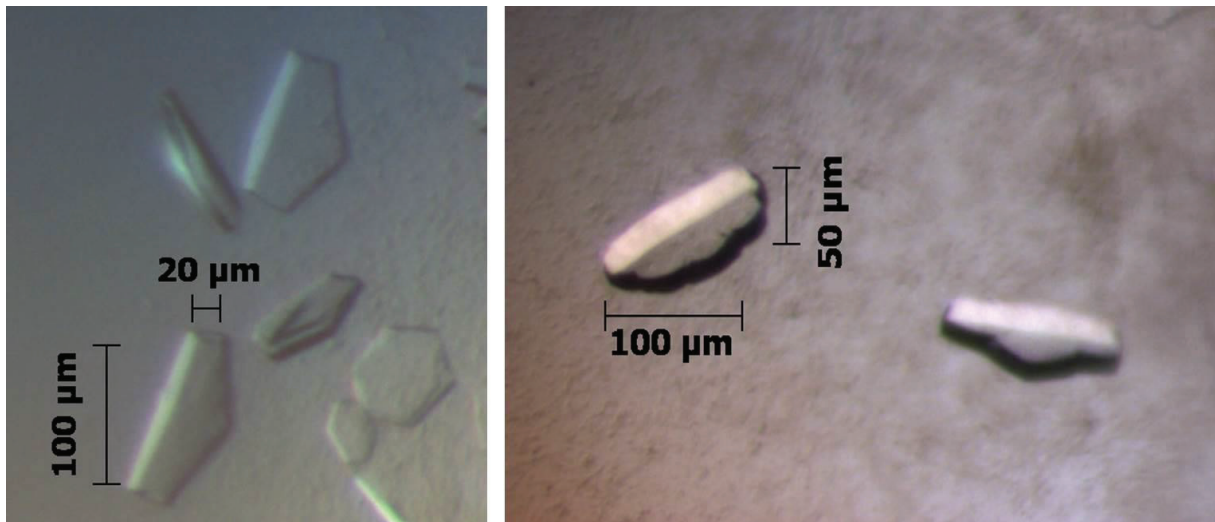


Figure 47 : A gauche, photographies des premiers cristaux de 6His-HpDnaB obtenus. A droite, cristaux de 6His-HpDnaB obtenus en ajoutant de la spermine tetrahydrochloride comme additif aux conditions de cristallisation.

La seconde forme cristalline est apparue grâce à la protéine purifiée en phosphate de sodium; après plusieurs mois (2-3) et en utilisant la même condition de cristallisation à l'exception du précipitant qui était cette fois le PEG 3350. Ces cristaux adoptaient une forme hexagonale et croissaient dans les trois dimensions. L'ajout d'ATP, d'ADP, ou d'AMP-PNP (1 mM) dans la goutte permet d'accélérer la formation des cristaux mais ne change pas la forme cristalline. Les conditions dans lesquelles sont apparus ces cristaux sont composées de 28-33% de PEG 3350, 0,2 M LiSO₄ et 0,1 M Tris pH 8,4 à 8,6 (Figure 48).

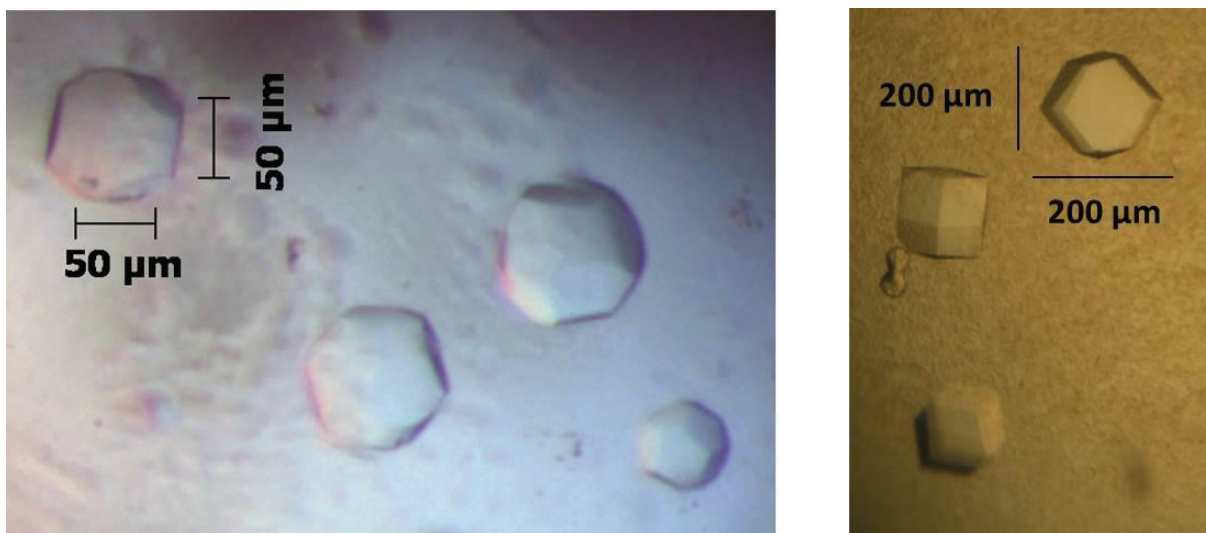


Figure 48 : Photographies des cristaux de 6His-HpDnaB en trois dimensions de forme hexagonale (gauche) et cocrystallisés avec de l'ADP (droite).

3.2.2. L'enregistrement et le traitement des données de diffraction

La méthode de cryoprotection utilisée pour les cristaux de 6His-HpDnaB consiste à incuber rapidement les cristaux dans une goutte de solution de cristallisation contenant 10% (v/v) d'une solution de sucrose saturée. Les cristaux étant fragiles et difficiles à pêcher à cause du lithium sulfate, l'autre technique utilisée était de remplacer progressivement la goutte de cristallisation par la solution de cryoprotection afin d'éviter le plus possible la manipulation des cristaux avant de les congeler dans de l'azote liquide. Le meilleur jeu de données que nous ayons collecté pour les cristaux natifs de 6His-HpDnaB a été indexé et intégré avec le programme imosflm (Leslie, 2006 ; Battye *et al.*, 2011). Le groupe d'espace I23 (ou I2₁3) a été déterminé comme étant le plus probable par le programme. La mise à l'échelle a été faite avec le programme scala de la suite CCP4 (Winn *et al.*, 2011). Le jeu de données avait des statistiques acceptables jusqu'à une résolution de 7,5 Å (Tableau 4).

Longueur d'onde (Å)	0.96112
Groupe d'espace	I2 ₁ 3
Paramètres de maille (Å, °)	a = b = c = 282.65, β = 90°
Limite de résolution (Å)	48,8-7,5 (7,91-7,50)
Nombre total de réflexions mesurées	171283 (25533)
Nombre de réflexions uniques mesurées	4955 (725)
Complétude (%)	99,8 (100)
Multiplicité	34,1 (35,2)
R _{meas}	0,13 (>1)
R _{pim}	0,03 (0,27)
R _{merge}	0,11 (>1)
CC _{1/2}	99,9 (78,9)
I/σ(I)	27,6 (2,8)
Facteur B moyen (Å ²)	334

Tableau 4 : Tableau des statistiques du jeu de données des cristaux natifs de 6His-HpDnaB. Les valeurs entre parenthèses correspondent aux statistiques de la dernière coquille de résolution.

3.2.3. Tentatives de phasage par remplacement moléculaire

La détermination du coefficient de Matthews nous a permis d'estimer entre quatre et huit le nombre de monomères dans l'unité asymétrique (Tableau 5). La protéine pouvant s'oligomériser en hexamère, nous avons donc considéré que l'hypothèse la plus probable était que six monomères composaient l'unité asymétrique. Afin de résoudre la structure par remplacement moléculaire, un modèle de HpDnaB en hexamère a été construit à partir de la structure apo de BstDnaB en complexe avec BstDnaG^{HBD}. Une carte de densité électronique a pu être obtenue mais le facteur R_{free} ne diminuait pas après affinement, au contraire il augmentait. De plus, aucune densité supplémentaire n'apparaissait et celle correspondant au modèle se dégradait. D'autres tentatives de remplacement moléculaire ont été réalisées avec plusieurs autres modèles : monomère, dimère en faisant varier le nombre de molécules à chercher mais aucune solution n'a pu être obtenue.

Nombre de molécules/unité asymétrique	Coefficient de Matthews	Pourcentage de solvant	Probabilité
1	16,17	92,40	0,00
2	8,08	84,79	0,00
3	5,39	77,19	0,00
4	4,04	69,59	0,02
5	3,23	61,99	0,09
6	2,69	54,38	0,26
7	2,31	46,78	0,40
8	2,02	39,18	0,21
9	1,80	31,57	0,02
10	1,62	23,97	0,00
11	1,47	16,37	0,00
12	1,35	8,77	0,00
13	1,24	1,16	0,00

Tableau 5 : Tableau de probabilité du nombre de molécules par unité asymétrique.

3.2.4. Phasage de la structure 6His-HpDnaB par single anomalous dispersion

Afin de résoudre le problème de phase, les cristaux de 6His-HpDnaB ont été trempés avec du bromure de tantalate ($Ta_6Br_{12}^{2+}$) pour que l'atome lourd diffuse à l'intérieur du cristal. Après deux heures de trempage, le bromure de tantalate avait pénétré les cristaux qui avaient pris une couleur verte. Les cristaux ont été cryoprotégés en utilisant la solution de cristallisation complétée avec 10% (v/v) de sucrose saturé puis congelés dans l'azote liquide (Figure 49).

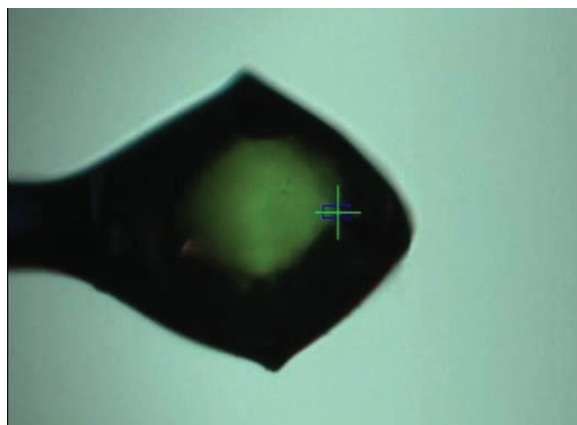


Figure 49 : Cristal de 6His-HpDnaB trempé avec du $Ta_6Br_{12}^{2+}$. La photo a été prise sur la ligne de lumière Proxima 2a du synchrotron SOLEIL.

La première étape était de vérifier la présence de l'atome lourd dans les cristaux. Pour cela, une expérience de fluorescence a été réalisée sur la ligne de lumière Proxima 2a au synchrotron SOLEIL. Le spectre indique un pic d'absorption à l'énergie ($E = 9982$ eV), proche du seuil théorique du tantalate (9881 eV). A partir du spectre, nous avons ainsi pu déterminer l'énergie correspondant au pic et au point d'inflexion du tantalate (Figure 50).

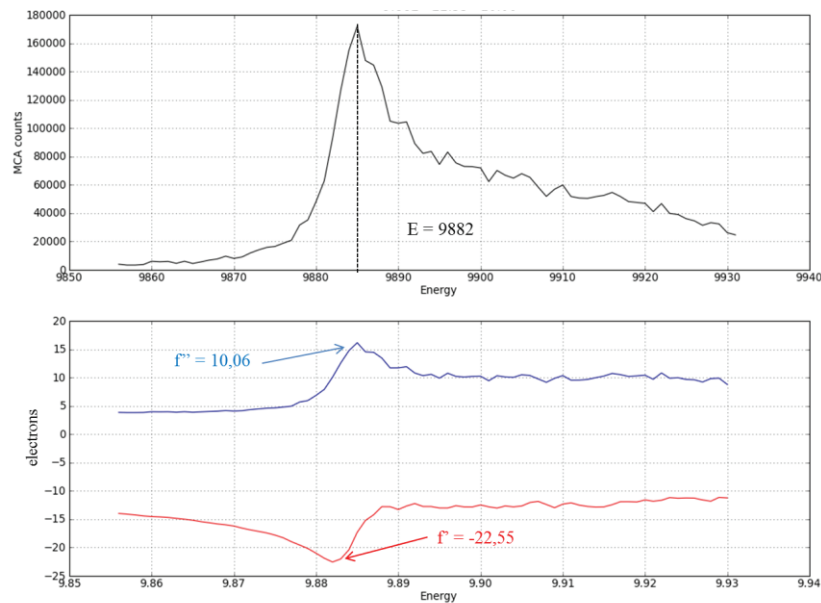


Figure 50 : Spectre de fluorescence réalisé sur la ligne de lumière Proxima 2a (en haut). Estimation des valeurs de f'' (bleu) et f' (rouge) (en bas)

Nous avons ensuite enregistré deux jeux de données : Tant 1 et Tant 2. Nous avons utilisé la méthode de collection inversée (« inversed beam »). Cette technique a pour objectif d'optimiser l'enregistrement du signal anomal. En collectant alternativement 10 images sur un angle puis à 180° , elle permet d'éviter des dégâts d'irradiation pouvant subvenir lors de la collection qui dégradent la qualité du jeu de données. On obtient ainsi des jeux de données isomorphes qu'il est possible de combiner, permettant ainsi d'augmenter la quantité de signal anomal (Tableau 6). Les jeux Tant 1 et Tant 2 ainsi obtenus ont été collectés sur le même cristal mais en exposant deux endroits du cristal et en augmentant le temps d'exposition de 30% pour Tant 2.

L'indexation et l'intégration des jeux de données ont été réalisées avec le programme XDS (Kabsch, 1993) qui a identifié le groupe d'espace le plus probable comme étant $I2_13$. Le logiciel POINTLESS (Evans, 2011) a également évalué que ce même groupe d'espace était le plus probable. Les paramètres de maille sont presque identiques à ceux des cristaux natifs : $a = b = c = 284,03$; $\alpha = \beta = \gamma = 90^\circ$. La résolution des jeux de données a été estimée grâce au facteur R mais surtout grâce au facteur $CC_{1/2}$. Le facteur $CC_{1/2}$ est un coefficient calculé à partir du pourcentage de corrélation entre deux intensités (I_1 et I_2) sélectionnées de façon aléatoire à l'intérieur d'un même jeu de données : $CC_{1/2} = \text{Corrélation}(I_1, I_2)$ (Diederichs & Karplus, 2013 ; Karplus & Diederichs, 2012). En ce qui concerne Tant1, nous avons limité la résolution à 7,5 Å pour obtenir des phases de meilleure qualité avec un rapport signal sur bruit ($I/\sigma(I)$) supérieur à 2,5. Pour Tant 2, nous avons conservé le maximum de réflexions possible, tout en conservant un rapport $I/\sigma(I)$ supérieur à 1. Le tableau de statistiques finales des deux jeux de données est présenté ci-dessous (Tableau 6).

Jeu de données	Tant 1 ($Ta_6Br_{12}^{2+}$) SOLEIL-Proxima2a	Tant 2 ($Ta_6Br_{12}^{2+}$) SOLEIL-Proxima2a
Longueur d'onde (Å)	1,2546	1,2546
Groupe d'espace	$I2_13$	$I2_13$
Paramètres de la maille (Å, °)	$a = b = c = 284,03$; $\beta = 90$	$a = b = c = 283,47$; $\beta = 90$
Limite de résolution (Å)	48,8-7,5 (7,91-7,50)	47,2 - 6,7 (6,93 - 6,70)
Nombre total de réflexions	171283 (25533)	308609 (45815)
Nombre de réflexions uniques	5016 (725)	6954 (1005)
Complétude (%)	99,6 (100)	99,7 (100)
Complétude anormale (%)	99,7(100)	99,8 (100)
Redondance	34,1 (35,2)	44,3 (45,6)
Redondance anormale	18,1(18,2)	
Rmeas	0,13 (>1)	0,18 (>1)
Rpim	0,03 (0,27)	0,03 (0,99)
Rmerge	0,11 (>1)	0,17 (>1)
$CC_{1/2}$	99,9 (78,9)	99,9 (41,6)
$I/\sigma(I)$	27,6 (2,8)	18,3 (1,2)
Facteur B moyen (Å ²)	334	502

Tableau 6 : Statistiques des données de diffraction des deux jeux Tant 1 et Tant 2. Les valeurs entre parenthèses correspondent aux statistiques de la dernière coquille de résolution.

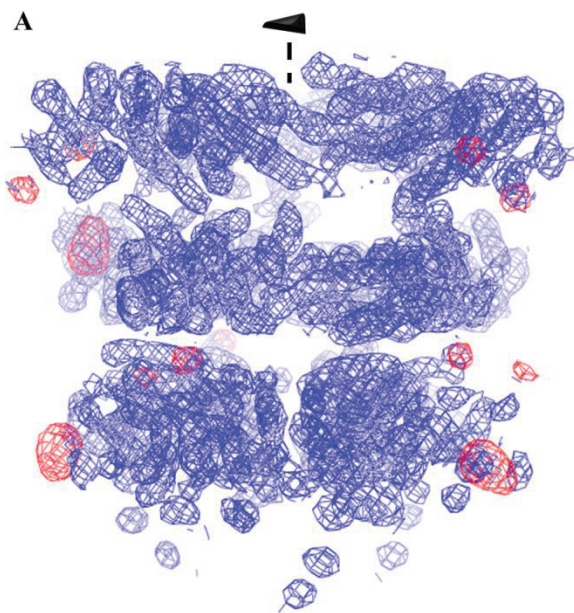
3.2.4.1. Localisation des sites, phasage et amélioration de la carte de densité électronique

Le principe d'un phasage par SAD est de localiser la position des atomes lourds en utilisant le signal anomal afin de déterminer les phases. Nous avons utilisé le jeu de données Tant 1 et déterminé la position des groupes d'atomes de bromure de tantalate grâce au programme Autosolve du logiciel PHENIX (Adams *et al.*, 2010). Les valeurs de f' et f'' ont été utilisées dans le logiciel HYSS (McCoy *et al.*, 2004) pour déterminer la position des atomes lourds. Le logiciel a trouvé la position de trois groupes d'atomes. Leurs coordonnées ont ensuite été utilisées dans Phaser (McCoy *et al.*, 2007) intégré dans le programme Resolve de PHENIX (Adams *et al.*, 2010 ; Terwilliger *et al.*, 2000) pour l'affinement des coordonnées et de l'occupation. La méthode SAD conduit à une ambiguïté sur la position des atomes lourds car elle ne permet pas de faire la distinction entre les deux « mains » possible du cristal. Deux solutions sont possibles, une position dans le groupe d'espace dans lequel nous avons intégré les données et une position dans un groupe différent mais indistinguables au niveau des réflexions présentes (dans notre cas $I23$ ou $I2_13$). Le logiciel effectue une modification de densité dans les deux groupes d'espace. On obtient donc deux cartes de densité électronique expérimentales correspondant à la tentative de phasage dans les deux groupes d'espaces après modification de densité. L'examen visuel des cartes indiquait clairement que le bon groupe d'espace était $I2_13$ (Tableau 7). Après modification de densité, la carte électronique obtenue est de très bonne qualité et permet de voir la densité électronique de la protéine (Figure 51).

<i>Statistiques de phasage</i>	
Nombre de sites de $Ta_6Br_{12}^{2+}$ trouvés	3
Figure de mérite (FOM) avant modification de densité	0,36
Figure de mérite (FOM) après modification de densité	0,78

Tableau 7 : Tableau de statistiques après phasage par SAD en utilisant le groupe d'espace $I2_13$.

La carte de densité électronique obtenue à 7,5 Å de résolution révèle que l'unité asymétrique est composée de quatre monomères. On voit très clairement que HpDnaB forme des dodécamères générés par l'axe cristallographique de symétrie trois (Figure 51). L'assemblage est constitué de deux hexamères qui se font face par l'anneau NTD dans le cristal. Dans l'unité asymétrique les quatre monomères sont répartis en deux dimères (un pour chaque hexamère) qui se font face. Avec quatre monomères dans l'unité asymétrique, le calcul du pourcentage de solvant grâce au coefficient de Matthews, estime à 70% le taux de solvant dans le cristal. Un taux aussi élevé est beaucoup pour un cristal d'une macromolécule biologique. En général, les cristaux de protéines, contiennent environ 50% de solvant. Nous avons pu ensuite étendre les phases à 6,7 Å de résolution grâce au jeu de donnée Tant 2 et débiter l'affinement.



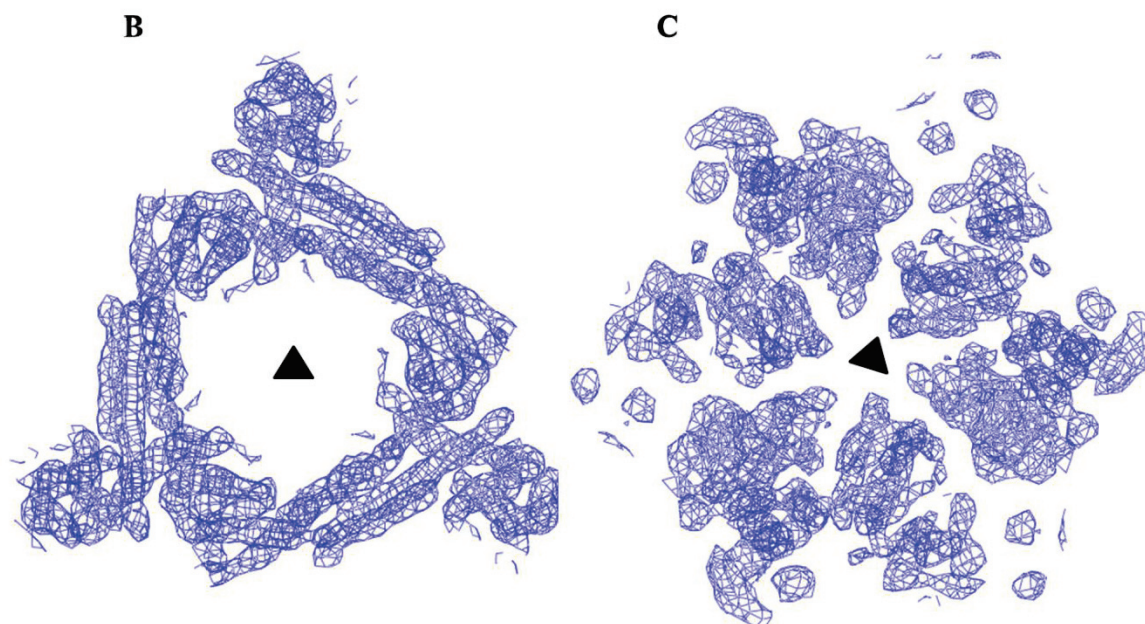


Figure 51 : Carte de densité électronique obtenue après phasage par Autosolve. (A) Densité électronique du dodécamère. Vue selon l'axe de symétrie trois, (B) densité électronique correspondant à l'anneau NTD et (C) densité électronique correspondant à l'anneau CTD.

3.2.5. Construction du modèle

Pour construire le modèle, les structures haute résolution des domaines NTD (code PDB 3GXV; Kashav *et al.*, 2009) et CTD (code PDB 4A1F; Stelter *et al.*, 2012) de HpDnaB ont été utilisées. Il a été possible de placer quatre domaines NTD (Figure 52) manuellement. Ces quatre domaines ont été ensuite utilisés pour améliorer les phases en calculant des symétries non cristallographiques. Ces symétries non cristallographiques ont été utilisées pour faire une moyennation de la densité électronique. La carte ainsi améliorée nous a permis de placer deux domaines CTD dans l'unité asymétrique. En revanche la densité électronique au niveau de l'anneau CTD du deuxième hexamère est absente. Ceci est assez surprenant compte tenu que les NTD étaient présents.

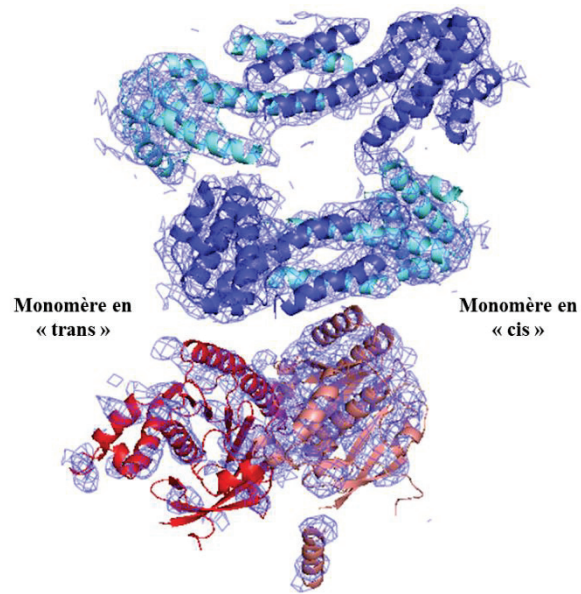


Figure 52 : Densité électronique correspondant à l'unité asymétrique avec les structures des domaines NTD (colorés en bleu et cyan) et des domaines CTD (en rouge et rose) placés dans la densité.

Pour vérifier que le cristal contenait la protéine entière, nous avons déposé les cristaux sur gel SDS (Figure 53). Le gel indique qu'il n'y a aucune dégradation de la protéine dans le cristal. L'examen des contacts cristallins indique que la région du deuxième CTD est exposée à un canal de solvant très large. De plus aucun contact cristallin n'est présent dans cette zone. Ceci suggère donc que cette région est très dynamique et/ou que l'anneau CTD adopte différentes conformations dans le cristal. C'est sans doute à cause de ce dynamisme que cette région n'est pas visible.

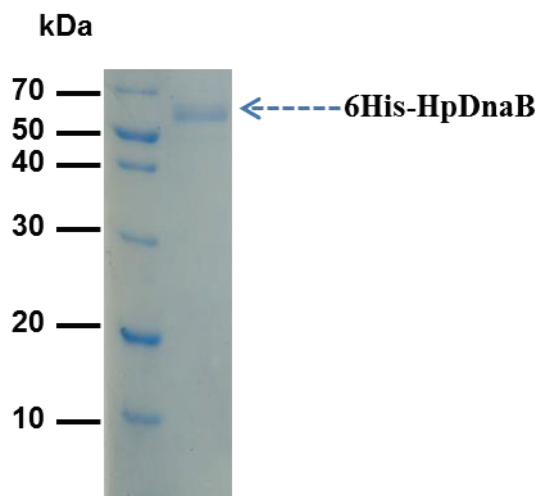


Figure 53 : Gel SDS des cristaux de 6His-HpDnaB coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

A l'issue des étapes de phasage et de la construction du premier modèle, la structure a été affinée. Le processus d'affinement a consisté en des phases de reconstructions manuelles, principalement par déplacement d'hélices entières par corps rigide à l'aide du logiciel COOT (Emsley et Cowtan, 2004) et des cycles d'affinement numériques. L'affinement a été réalisé en utilisant le programme Buster TNT (Blanc *et al.*, 2004) qui procède par maximum de vraisemblance. Ainsi, au cours de l'affinement nous avons modifié le modèle de départ. La structure du domaine NTD disponible dans la PDB est incomplète, il manque la deuxième hélice de l'épingle à cheveux α . Dans notre carte expérimentale, on voit très clairement la densité électronique de cette deuxième hélice. Nous avons donc reconstruit cette deuxième hélice en modèle polyalanine. La boucle connectant les deux hélices de l'épingle à cheveux n'ont pas pu être construites de manière satisfaisante bien que la densité soit présente. De même la séquence de la seconde hélice n'a pas pu être placée de manière définitive. Durant l'affinement nous avons également pu placer les hélices des domaines de liaison à l'intérieur de la poche d'hexamérisation. Cette hélice était présente dans la structure du CTD isolé, mais n'était pas correctement positionnée (Stelter *et al.*, 2012). De plus la densité correspondante au domaine de liaison entre le NTD et l'hélice a pu être construite.

Dans la carte de départ, aucune densité électronique n'était visible au niveau de certaines régions du CTD isolé. Nous avons donc commencé l'affinement en enlevant ces domaines de notre structure et regardé si la densité électronique correspondante apparaissait. Cela n'a pas été le cas pour plusieurs boucles dont la première boucle de liaison à l'ADN, exposée dans le canal central de l'hexamère. En l'absence d'ADN ces boucles sont sans doute flexibles et ne sont pas présente dans notre modèle final. La structure à haute résolution du domaine CTD (Stelter *et al.*, 2012) a montré que l'insertion unique de 34 acides aminés (HPI) présente chez HpDnaB s'arrange en un motif hélice-tour-hélice. Cette insertion est effectivement visible et nous avons pu placer les HPI des deux monomères. Après affinement, il s'avère que seule une des deux hélices est présente, l'autre étant désordonnée et non visible. Nous avons donc modifié les HPI en ne laissant qu'une seule des deux hélices à chaque fois.

Pour le deuxième CTD manquant (correspondant au deuxième hexamère), des densités partielles apparaissent pour deux hélices proches des deux NTD (hélice α_{12} et α_{14}). Cependant aucun de ces domaines n'a pu être construit correctement.

Nous avons enfin pu placer un groupe d'atome $\text{Ta}_6\text{Br}_{12}^{2+}$ sur les trois présents lors du phasage. Grâce à l'ajout du groupe d'atome lourd dans le modèle final, les facteurs R/R_{free} ont

diminué pour arriver à leur valeur finale 0,25/0,29 (Tableau 8, Figure 54). Le diagramme de Ramachandran est à nuancer car il est surtout hérité des structures parents utilisées lors du remplacement moléculaire à cette résolution.

Nombre de réflexions utilisées	13223
Facteur R	0,25
Facteur R free	0,29
Nombre d'atomes du ligand	17
Nombre de résidus	1173
RMS	
Liaison (Å)	0,004
Angles (°)	0,82
Ramachandran :	
Résidus dans la région favorable	97
Résidus dans la région peu favorable	2,9
Résidus dans la région interdite	0,1
Score d'incompatibilité	11,12
Score global	1,85
Facteur B moyen	329,1

Tableau 8 : Statistiques d'affinement de la structure de 6His-HpDnaB à une résolution de 6,7 Å.

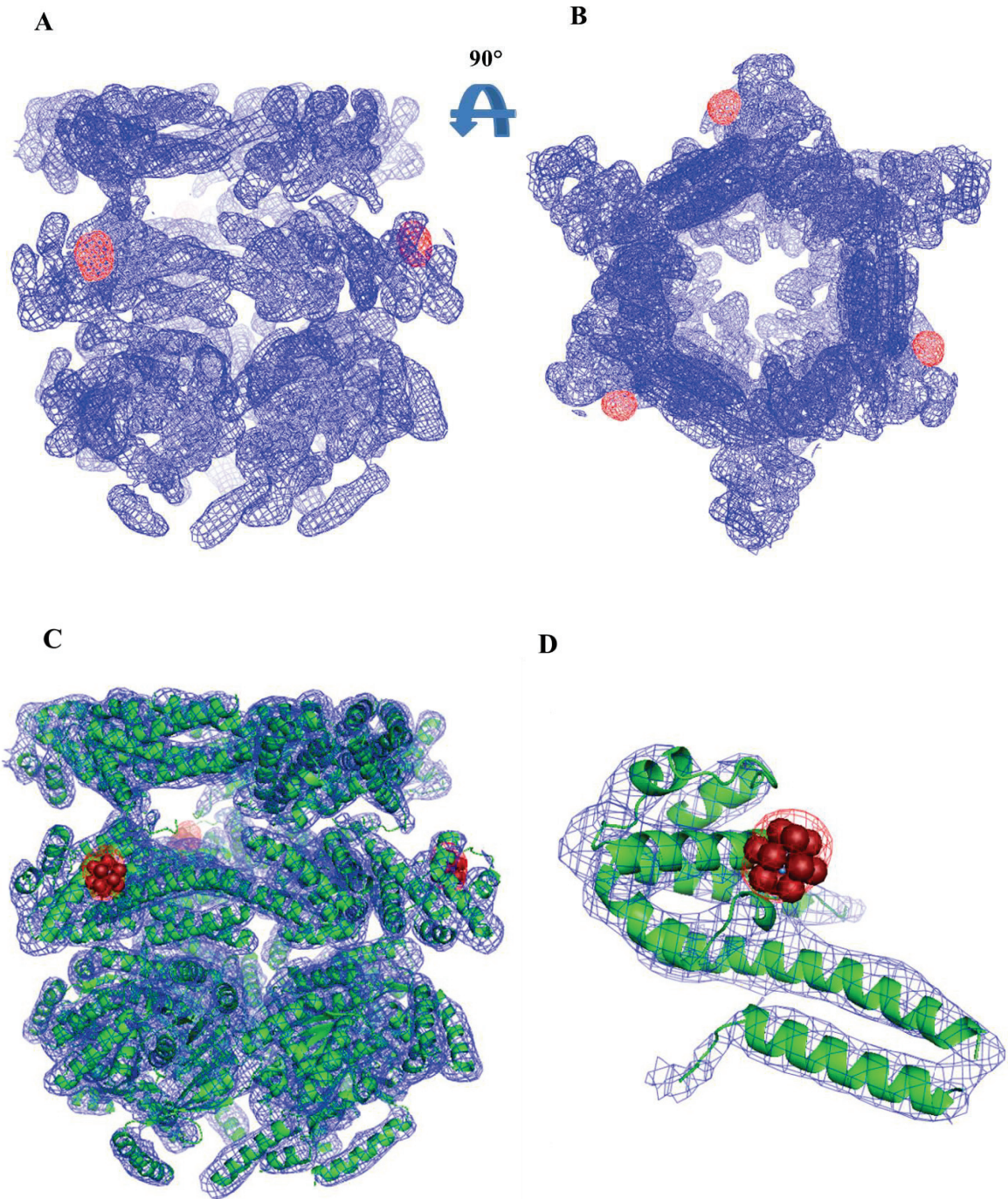


Figure 54 : Carte 2Fo-Fc de 6His-HpDnaB vue de côté (A) et vue de dessus (B). (C) Le modèle final de 6His-HpDnaB obtenu après affinement est représenté en ruban vert. La carte 2Fo-Fc est colorée en bleu. La carte de différence anormale est indiquée en rouge et les atomes de Brome du groupe d'atome lourds sont représentés en sphères rouges. (D) Image agrandie de la carte de densité électronique d'un domaine NTD.

3.2.6. Validation du modèle par l'utilisation des cristaux de 6His-HpDnaB-Se

3.2.6.1. Purification et cristallisation de 6His-HpDnaB-Se

Avant de pouvoir valider la séquence en acide aminé du modèle et la position de certaines régions, nous avons produit 6His-HpDnaB contenant des méthionines sélénées (6His-HpDnaB-Se). Le protocole de purification de la protéine est le même que pour HpDnaB. L'incorporation de la sélénométhionine n'a pas provoqué un changement des conditions de cristallisation qui restent les mêmes que pour HpDnaB : 28% à 33% PEG 3350, 0,1 M Tris pH 8,4 à 8,6 ; 0,2 M LiSO₄ et 1 mM ADP. Les cristaux ainsi obtenus sont également de forme hexagonale (Figure 55). Les cristaux ont été congelés de la même manière que les cristaux de 6His-HpDnaB

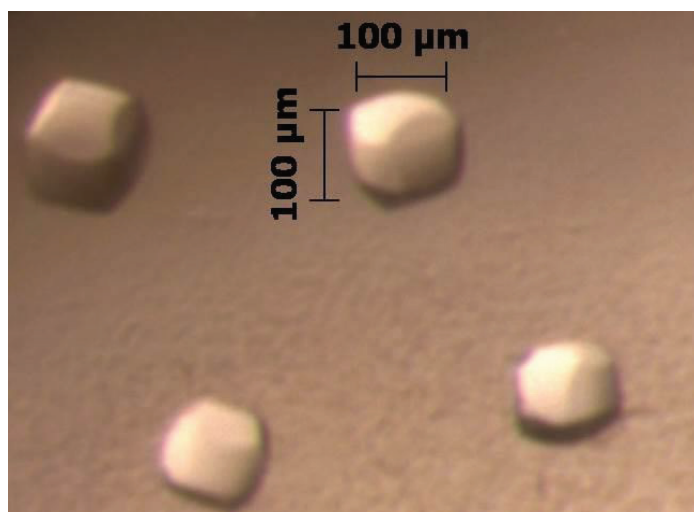


Figure 55 : Photographie des cristaux de la protéine 6His-HpDnaB-Se

3.2.6.2. Collection et traitement des données de diffraction

Un spectre de fluorescence a été réalisé à partir des cristaux sur la ligne de lumière ID14EH4 de l'ESRF (European Synchrotron Radiation Facility) à Grenoble et a confirmé la présence de la sélénométhionine. Un jeu de données a été collecté à la longueur d'onde correspondante au pic d'absorption du sélénium (0,97239 Å). Le jeu de données a été indexé et intégré avec le logiciel XDS en utilisant le même groupe d'espace que pour les cristaux trempés avec le Ta₆Br₁₂²⁺. La mise à l'échelle a ensuite été faite grâce au programme SCALA

de la suite CCP4 (Winn *et al.*, 2011). La limite de résolution a été ajustée à 8,0 Å grâce aux facteurs $CC_{1/2}$, Rmerge et le rapport signal sur bruit ($I/\sigma(I)$). Les statistiques du jeu de données sont indiquées dans le Tableau 9.

	ID14EH4-ESRF
Longueur d'onde (Å)	0,97239
Groupe d'espace	I2 ₁₃
Paramètres de la maille (Å, °)	a = b = c = 283,11
Limite de résolution (Å)	48,6-8,0 (8,43-8,0)
Nombre total de réflexions	29801 (4471)
Nombre de réflexions uniques	4111 (602)
Complétude (%)	99,5 (100)
Complétude anormale (%)	99,7 (100)
Redondance	7,2 (7,4)
Redondance anormale	18,1 (18,2)
Rmeas	0,10 (0,64)
Rpim	0,05 (0,23)
Rmerge	0,09 (0,55)
$CC_{1/2}$	99,8 (81,9)
$I/\sigma(I)$	17,4 (3,6)
Facteur B moyen (Å ²)	247

Tableau 9 : Statistiques des données de diffraction d'un cristal de 6His-HpDnaB-Se au seuil du sélénium. Les valeurs entre parenthèses correspondent à la dernière coquille de résolution.

3.2.6.3. Remplacement moléculaire

Nous avons utilisé notre modèle pour réaliser un remplacement moléculaire avec le jeu collecté avec les cristaux 6His-HpDnaB-Se. Le modèle a pu être correctement placé et nous avons utilisé les phases pour générer une carte de différence anormale. La carte nous a permis de localiser les sites de sélénium et de vérifier ainsi la position des méthionines de notre modèle. La séquence de HpDnaB, contient deux méthionines dans le domaine NTD, une méthionine dans le domaine de liaison et neuf méthionines dans le domaine CTD. Le signal anomal est assez fort pour confirmer la position des deux méthionines du domaine NTD ainsi

que pour sept méthionines du domaine CTD (Figure 56). Le signal anomal de la méthionine du domaine de liaison est faible et le signal est absent pour deux méthionines du CTD. En effet, l'une se trouve dans une boucle flexible dont la densité n'est pas visible et l'autre dans le HPI.

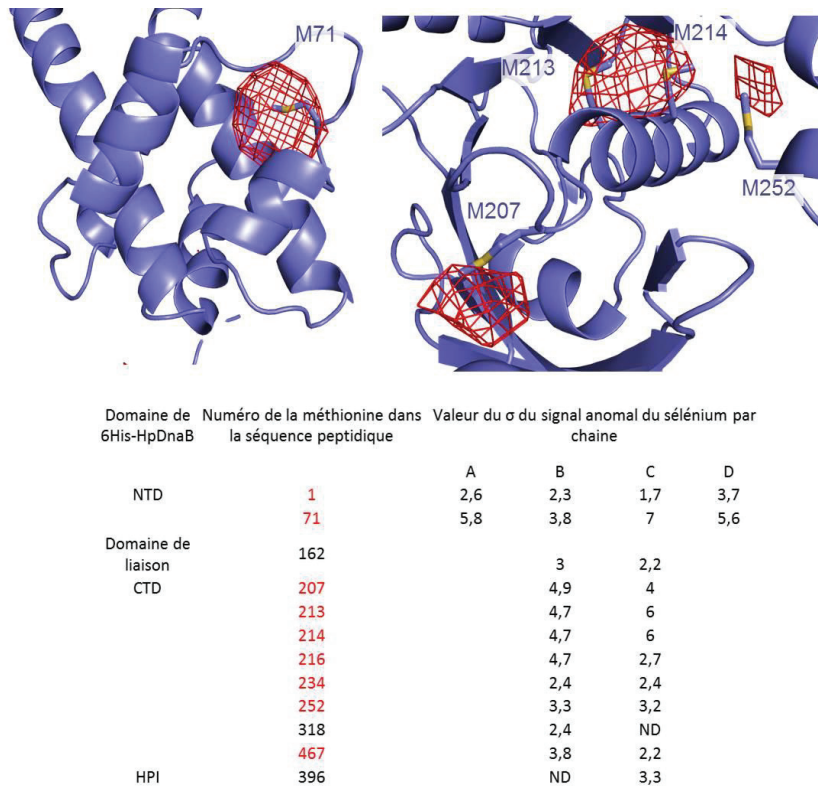


Figure 56 : En haut, carte de densité électronique de différence anormale du sélénium (en rouge) générée après remplacement moléculaire. En bas, tableau comparant les méthionines et la valeur du σ du signal anomal du sélénium. Les numéros en rouges signifient qu'un signal de sélénium élevé est présent et que la position de la méthionine est confirmée.

3.2.7. Structure générale de 6His-HpDnaB

La structure de 6His-HpDnaB est un dodécamère généré par symétrie cristallographique selon un axe trois dans le cristal (Figure 57). Les deux hexamères sont composés d'un anneau NTD mais seul l'anneau CTD du premier hexamère est visible dans la carte de densité électronique. Comme dit précédemment, l'anneau CTD du deuxième hexamère n'est pas visible car les sous-unités pourraient bouger les unes par rapport aux autres. Lorsque le dodécamère entier est reconstitué en superposant l'hexamère 1 entier sur le deuxième NTD, le modèle complet a une longueur de 190 Å et une largeur de 120 Å. Ces mesures sont différentes de celle de l'enveloppe obtenue en ME à coloration négative (155 Å

de longueur et 130 Å de largeur). Cette différence de taille peut s'expliquer par le phénomène d'aplatissement de densité en ME.

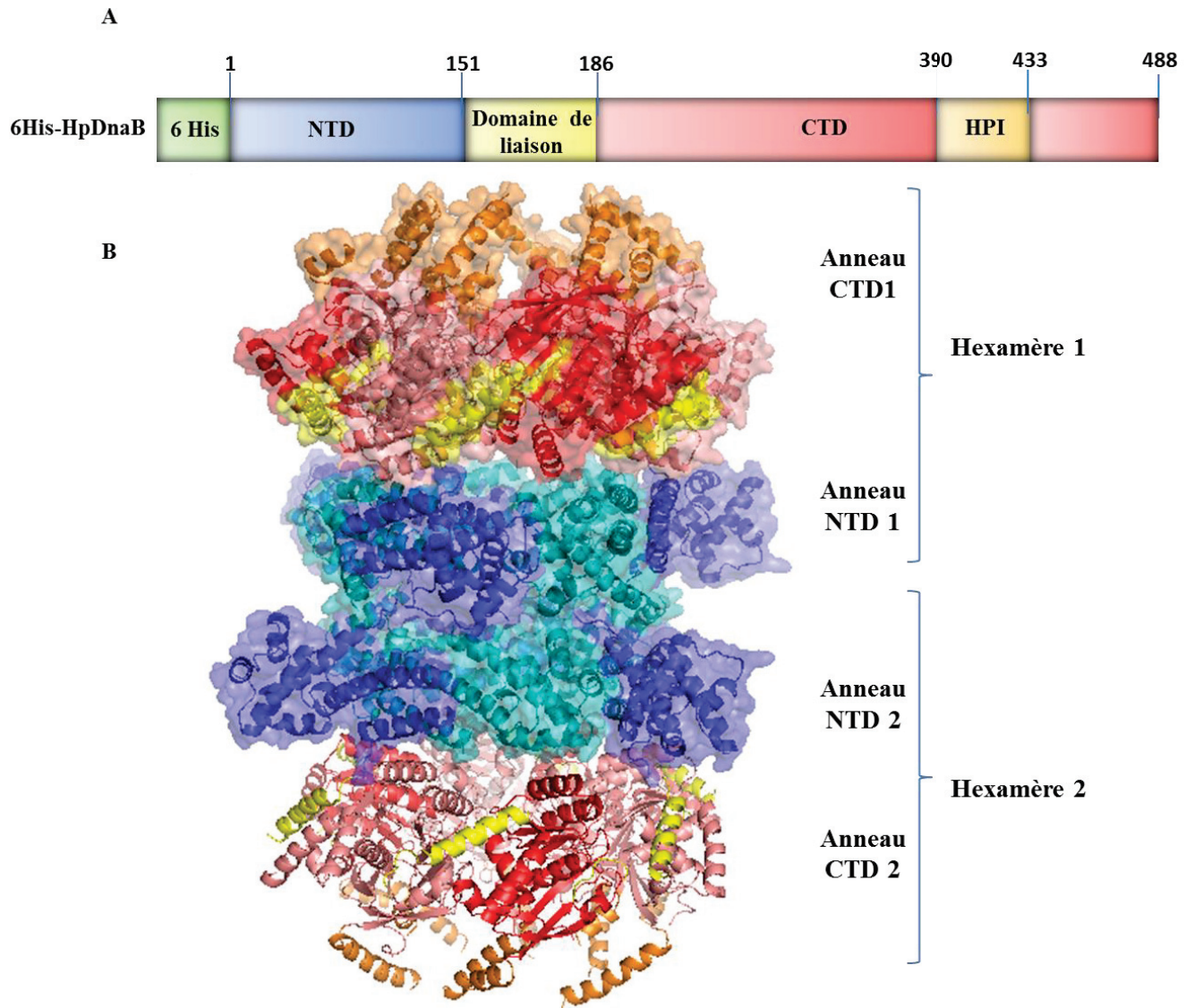


Figure 57 : (A) Représentation schématique de l'organisation des domaines du monomère de HpDnaB. (B) Structure de 6His-HpDnaB colorée en fonction des différents domaines et représentée en ruban. La partie dont la surface n'est pas représentée correspond au CTD 2 qui n'a pas pu être construit dans la carte de densité électronique.

3.2.7.1. Détails des interactions

3.2.7.1.1. Anneaux NTD et l'interface de dodécamérisation

Chaque domaine NTD interagit avec un NTD voisin pour former un dimère. L'interface d'interaction se situe au niveau des épingles à cheveux α des NTD comme observé dans les autres structures de DnaB. Les deux anneaux NTD sont identiques et sont constitués d'un trimère de dimères délimitant un canal central de 50 Å de diamètre (Figure 58A). Cette conformation des anneaux NTD est appelée conformation dilatée et a été décrite chez BstDnaB et GkDnaC (Bailey *et al.*, 2007b ; Lo *et al.*, 2009). La structure cristallographique montre que les deux hexamères interagissent ensemble au niveau des anneaux NTD comme cela avait été observé en ME. Cependant la forme dodécamérique adoptée par HpDnaB n'avait jamais été observée jusqu'à maintenant chez les hélicases répliquatives bactériennes. Notre structure cristallographique diffère de l'enveloppe obtenue en ME. En effet les anneaux NTD s'arrangent selon une symétrie trois alors que dans l'enveloppe de ME, chaque NTD du dodécamère avait été reconstruit à partir d'une symétrie C6, car celle-ci obtenait des meilleurs scores statistiques. Cette différence entre analyse microscopique et cristallographique s'explique par l'observation de la structure obtenue. En effet, dans la structure de 6His-HpDnaB, les deux anneaux NTD s'empilent l'un sur l'autre avec un décalage de 60° selon l'axe de symétrie (Figure 58B).

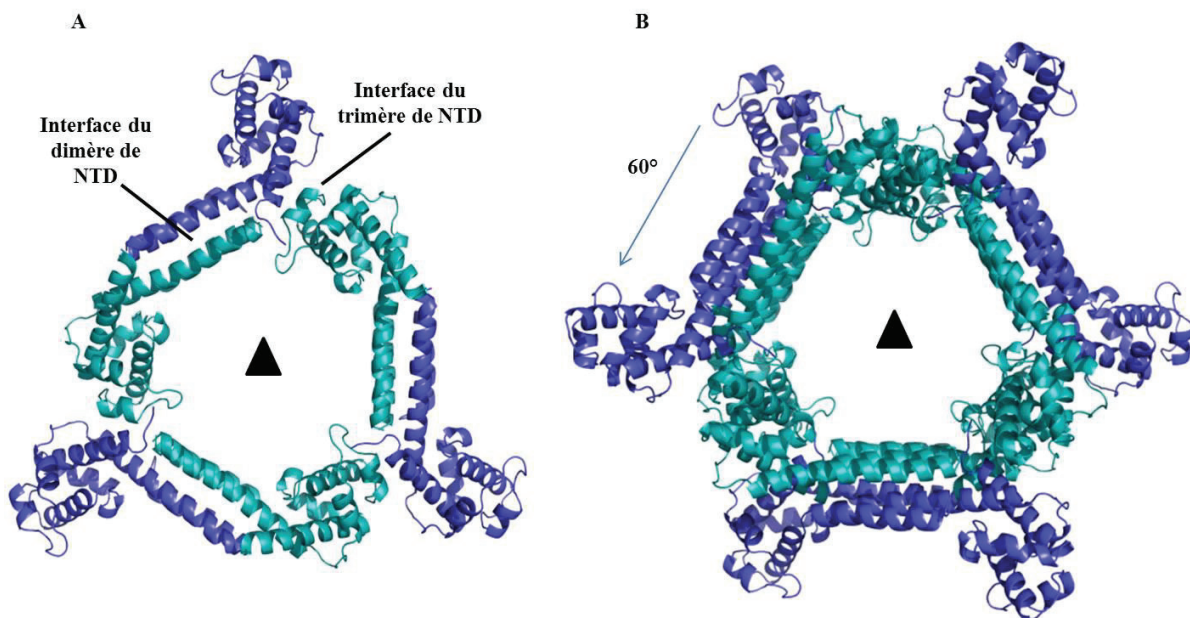


Figure 58 : (A) Représentation d'un anneau NTD de 6His-HpDnaB vue du dessus. (B) Superposition des deux anneaux NTD de 6His-HpDnaB vue du dessus.

Ainsi les deux anneaux NTD empilés l'un sur l'autre créent une symétrie pseudo-C6 artificielle vue de dessus (le long de l'axe cristallographique d'ordre trois) par projection des têtes globulaires de chaque NTD (Figure 59). De plus, les CTD étant en symétrie pseudo-C6 (voir plus loin), il n'était pas possible de déterminer une symétrie C3 à partir de l'ensemble des images.

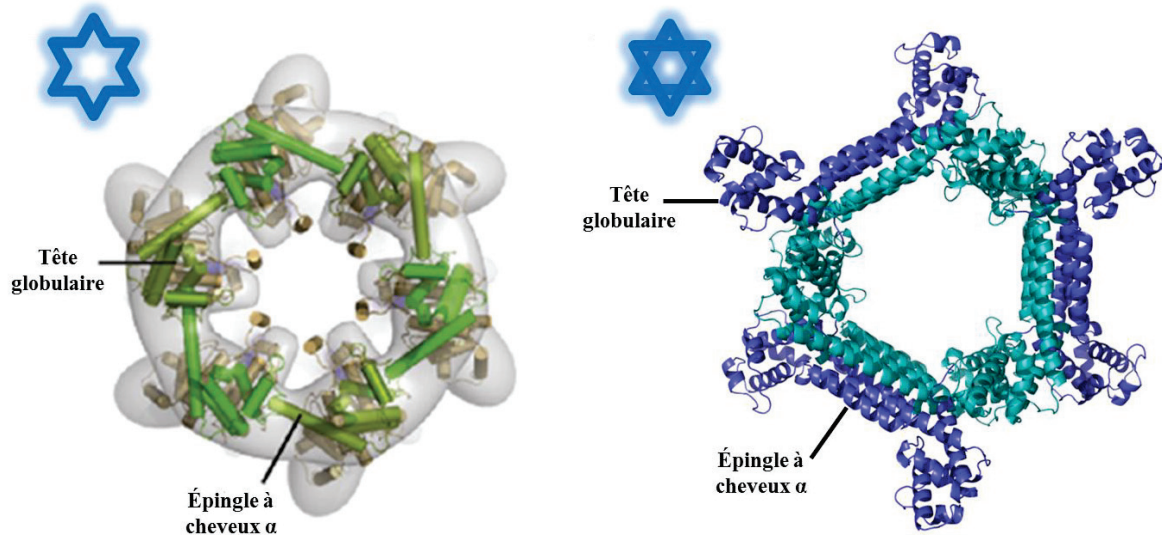


Figure 59 : Comparaison entre l'anneau NTD issu de la reconstruction en ME (à gauche) et les anneaux NTD de la structure cristallographique (à droite). Dans la structure ME, les NTD s'arrangent selon une symétrie pseudo-C6 alors que dans la structure cristallographique les NTD s'arrangent selon une symétrie C3 mais la superposition des deux anneaux NTD selon un axe 2_1 crée une symétrie C6 par projection des têtes globulaires.

Les hélicases MCM des archées et Mcm 2-7 des eucaryotes adoptent elles aussi une forme dodécamérique (Figure 60). Cependant, les deux hexamères de ces hélicases n'interagissent pas de la même façon. La structure cristallographique des domaines NTD de l'hélicase MCM de *Methanothermobacter thermautotrophicus* (MtMCM) (PDB : 1LTL, Fletcher *et al.*, 2003) montre que ces domaines interagissent entre eux par des boucles qui coordonnent chacune un atome de zinc. En ce qui concerne Mcm 2-7, il n'existe pas de structure cristallographique du double anneau et il n'est donc pas possible de comparer notre structure à cette dernière.

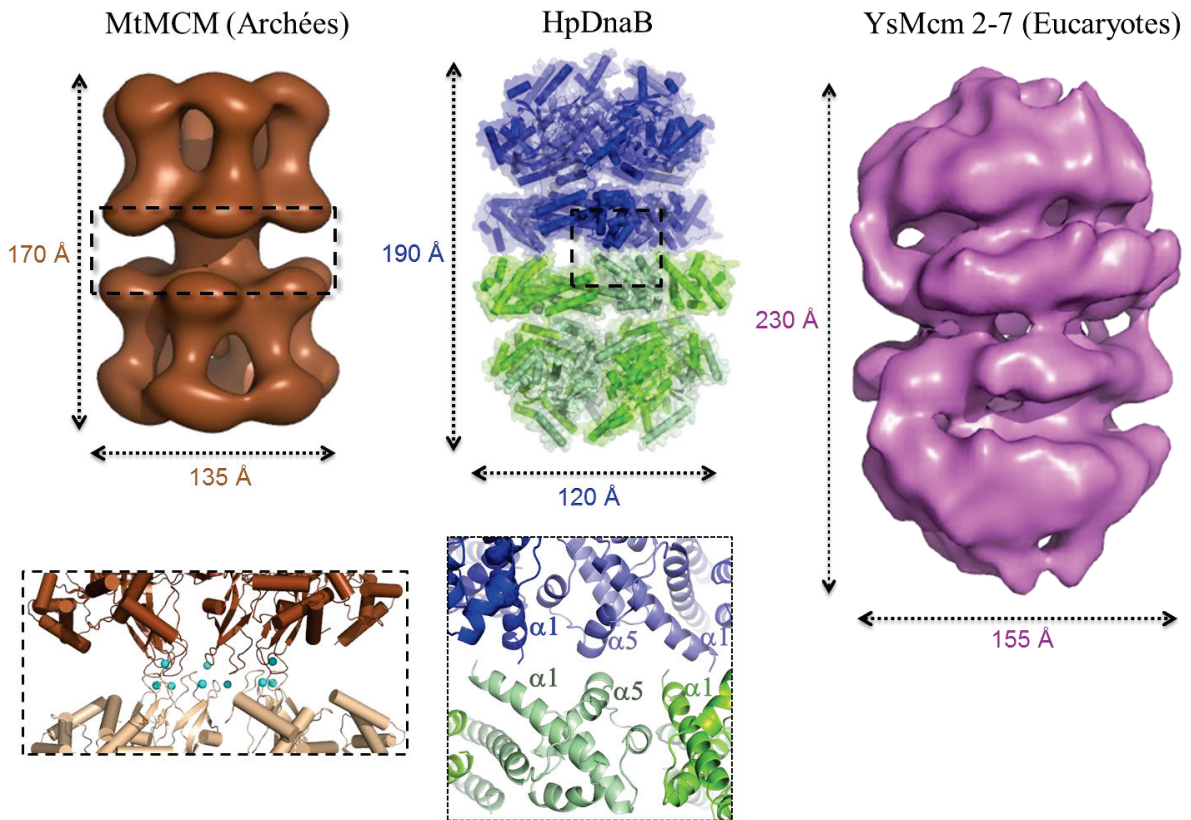


Figure 60 : Comparaison de l'enveloppe de ME de MtMCM en marron (à gauche) avec l'enveloppe de cryo-ME de YsMcm 2-7 en magenta (à droite) et le modèle de la structure de HpDnaB (au centre). En dessous sont représentés la structure cristallographique des deux anneaux NTD de MtMCM en marron avec les atomes de zinc en cyan ; et un agrandissement de l'interaction entre deux hexamères de HpDnaB (en bleu et en vert) au niveau de quatre domaines NTD.

L'interaction entre les anneaux NTD de HpDnaB s'effectue grâce aux têtes globulaires de chaque NTD individuel et l'ensemble crée une surface d'interaction totale de 1400 \AA^2 . L'analyse approfondie des deux monomères montre que les hélices $\alpha 1$ (L4 et Q8) et $\alpha 5$ (E80 et D81) du NTD de la chaîne A (anneau NTD 1) interagissent avec les hélices $\alpha 1$ de la chaîne D et $\alpha 5$ de la chaîne symétrique C' (anneau NTD 2). L'hélice $\alpha 1$ du NTD de la chaîne B (anneau NTD 1) interagit avec l'hélice $\alpha 1$ de la chaîne C (anneau NTD 2) (Figure 61).

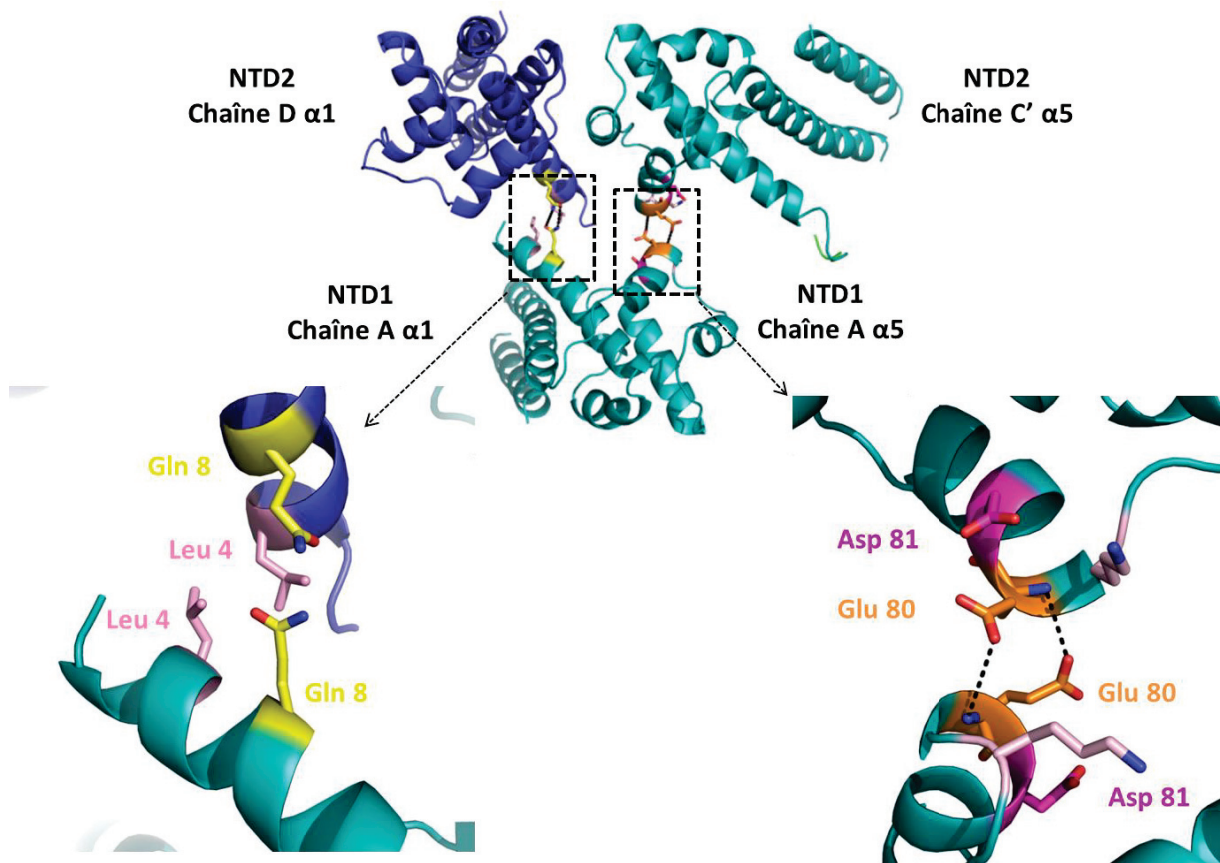


Figure 61 : Détails des interactions entre les deux anneaux NTD. En bas à gauche sont représentés les résidus impliqués dans l'interaction entre l'hélice $\alpha 1$ de la chaîne A du NTD 1 et l'hélice $\alpha 1$ de la chaîne D du NTD 2. En bas à droite sont représentés les résidus impliqués dans l'interaction entre l'hélice $\alpha 5$ de la chaîne A du NTD 1 et l'hélice $\alpha 5$ de la chaîne C du NTD 2.

3.2.7.1.2. Anneaux CTD et HPI

La structure montre que l'interaction entre les six domaines CTD est permise grâce aux domaines de liaison de chaque monomère. Les hélices $\alpha 9$, $\alpha 10$, $\alpha 11$ d'un domaine CTD forment une poche dans laquelle vient interagir l'hélice α du domaine de liaison du monomère voisin (Figure 62). Cette interaction est retrouvée dans toutes les autres structures d'hélicases bactérienne de la famille SF4.

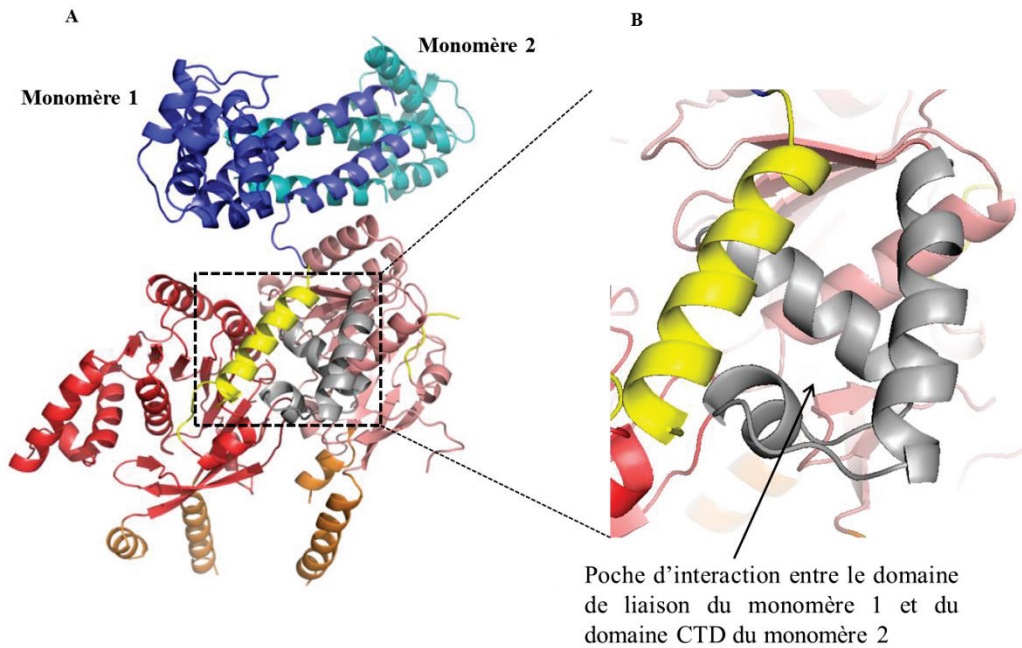


Figure 62 : (A) Arrangement de deux monomères de 6His-HpDnaB. (B) Agrandissement de la zone d'interaction. L'hélice du domaine de liaison du monomère 1 (en jaune) interagit avec la poche formée par les hélices $\alpha 9$, $\alpha 10$, $\alpha 11$ (en gris) du domaine CTD du monomère 2.

Les récentes structures de BstDnaB et AaDnaB montrent que l'anneau CTD peut adopter deux conformations différentes. La conformation de l'anneau CTD de la forme apo de BstDnaB est appelée « ouverte » et celle de l'anneau CTD d'AaDnaB en complexe avec l'ADP est appelée forme « fermée ». Dans notre structure, l'anneau CTD adopte une conformation fermée (Figure 63).

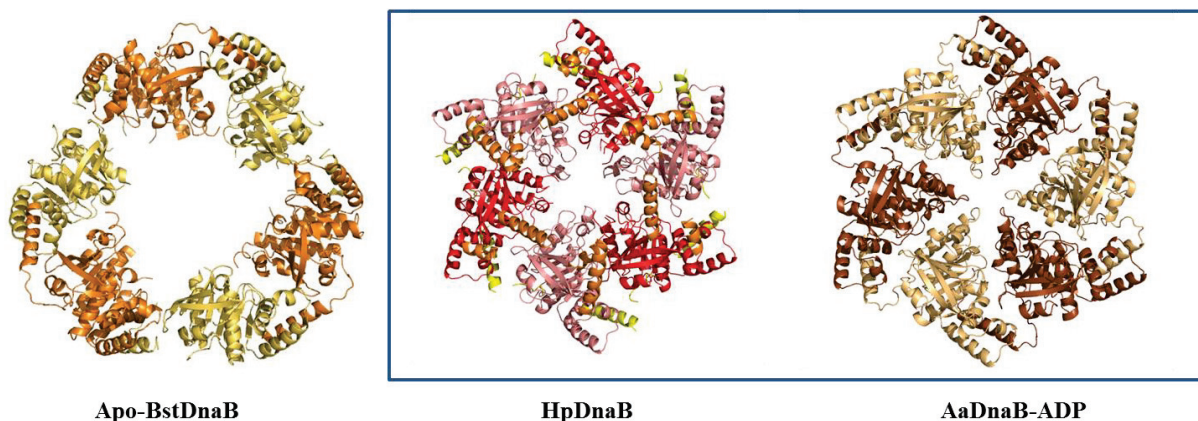


Figure 63 : Comparaison entre l'anneau CTD de HpDnaB (au milieu) avec l'anneau CTD de BstDnaB (à gauche) et l'anneau CTD d'AaDnaB-ADP (à droite).

En effet, dans la structure d'AaDnaB, la présence d'ADP au niveau du Walker A permet une rotation des six domaines CTD, ce qui réduit le diamètre du canal central (d'où la forme fermée). Les cristaux de 6His-HpDnaB ayant été obtenus en présence d'ADP, il est possible que de l'ADP soit présent dans notre structure. Néanmoins, la résolution de la structure de la protéine ne nous permet pas de voir de densité électronique correspondant à une molécule d'ADP au niveau de son site de fixation. Cependant, la conformation fermée du CTD de HpDnaB suggère que la liaison à l'ADP engendre les mêmes réarrangements structuraux que pour AaDnaB. La présence d'ADP pour AaDnaB induit également un réarrangement du trimère de dimères de l'anneau NTD, qui adopte un état resserré (ou « triskèle »). Ce n'est pas le cas dans notre structure où l'anneau NTD adopte la même conformation dilatée que BstDnaB (Figure 64). Dans la structure de 6His-HpDnaB, l'hexamère adopte donc une conformation originale, intermédiaire entre la forme apo et la forme liée à l'ADP décrites chez les autres DnaB.

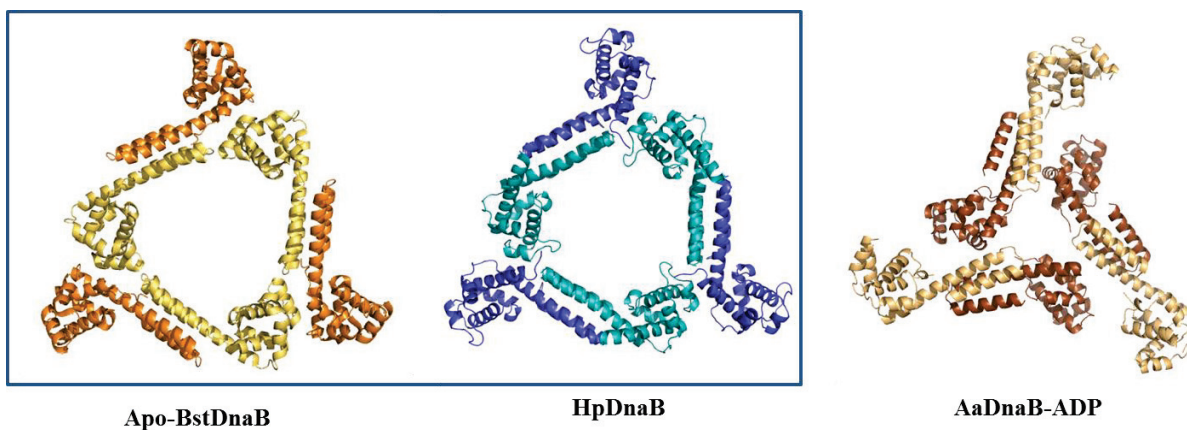


Figure 64 : Comparaison entre l'anneau NTD de HpDnaB (au milieu) avec l'anneau NTD de BstDnaB (à gauche) et l'anneau NTD d'AaDnaB-ADP (à droite).

Dans la structure cristallographique de HpDnaB^{CTD} (voir Article 1), l'insertion de 34 acides aminés (HPI) est composée de deux hélices $\alpha 15$ et $\alpha 16$ formant un motif hélice-tour-hélice. Dans la structure de 6His-HpDnaB, le HPI est différent. Dans la chaîne A, l'hélice $\alpha 16$ et seulement l'extrémité N-terminale de l'hélice $\alpha 15$ sont ordonnées. Dans la chaîne B, l'inverse est observé : l'hélice $\alpha 15$ et l'extrémité C-terminale de l'hélice $\alpha 16$ sont ordonnées (Figure 65).

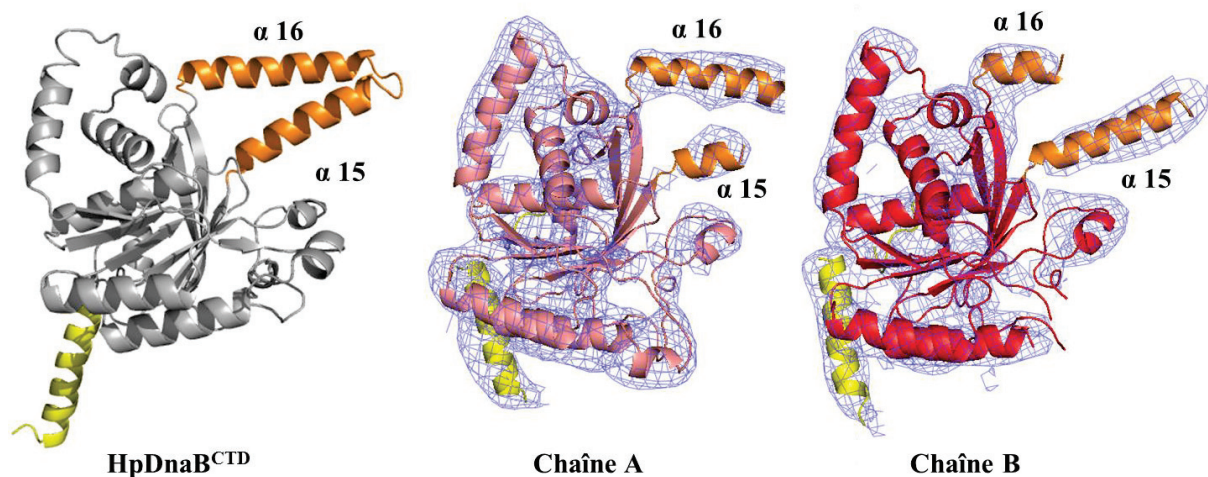


Figure 65 : Comparaison entre le HPI de la structure de HpDnaB^{CTD} et les HPI des chaînes A et B de l'unité asymétrique de 6His-HpDnaB. Dans les trois cas, le HPI est représenté en orange et l'hélice du domaine de liaison en jaune.

Dans le contexte de l'hexamère, les HPI forment un collier dépassant de l'anneau CTD et le prolongent. On ne connaît pas encore la fonction exacte de cette insertion qui n'a pour l'instant été observée chez les bactéries que dans la structure de HpDnaB. Cependant, des études ont montré qu'un mutant de la protéine ne possédant plus le HPI ne forme plus d'hexamère et est inactif (Soni *et al.*, 2003). La structure indique que l'hélice $\alpha 15$ de la chaîne A (mais pas celle de la chaîne B) interagit avec le domaine ATPase adjacent. Cette interaction permettrait d'expliquer les données selon lesquelles HpDnaB maintient ses activités ATPase et hélicase en l'absence de l'anneau NTD, contrairement aux autres DnaB (Soni *et al.*, 2003). La structure de 6His-HpDnaB suggère donc que l'interaction entre le HPI et le domaine ATPase stabilise l'anneau CTD en hexamère et ce même en l'absence de l'anneau NTD.

3.2.8. Tentative d'amélioration du modèle

Avant d'obtenir la structure de HpDnaB entière, le domaine NTD de HpDnaB avait été cloné et purifié dans le but d'obtenir des cristaux du domaine NTD entier. En effet la structure cristalline résolue ne contient que la première hélice de l'épingle à cheveux. La connaissance de la structure de cette hélice nous aurait permis d'améliorer notre modèle en incluant la structure de cette hélice (à la place du modèle polyalanine). Le domaine NTD du monomère de HpDnaB a donc été cloné et purifié.

Des essais de cristallisation ont été entrepris afin d'obtenir une autre forme cristalline que celle de l'étude de Kashav *et al.*, 2010. Nous avons pour cela criblé 288 conditions de cristallisation en utilisant les solutions commerciales PEG ION/Natrix, Cristal screen 1 et 2 et JCSG+. Une condition différente a pu être obtenue que nous avons essayé d'optimiser. Celle-ci contenait 5% à 25% de PEG 10000 et 0,1 M Hepes pH 7,5 ; néanmoins seuls des microcristaux sont apparus.

3.2.9. Trempage des cristaux avec un oligonucléotide

Nous souhaitons comprendre comment la protéine HpDnaB liait l'ADN et déterminer si cet attachement permettrait de structurer les boucles d'ADN non visibles dans la structure de 6His-HpDnaB. Les cristaux de 6His-HpDnaB obtenus en présence d'ATP, ADP ou AMP-PNP ont été utilisés pour réaliser des trempages avec un ADNsb polythymine de 18 ou 25 bases possédant un marqueur Cyanine 5 (polydT18/25-Cy5). L'utilisation du marqueur a permis de suivre la diffusion de l'ADN à l'intérieur du cristal. En effet la cyanine 5 colore la solution de cristallisation en bleu. Ainsi, les cristaux sont progressivement devenus bleu alors que la solution de cristallisation devenait incolore, indiquant que l'ADN avait pénétré dans les cristaux (Figure 66). Plusieurs périodes d'incubation ont été testées variant de 1h à 16h à 19°C.



Figure 66 : Cristaux de 6His-HpDnaB incubés avec un ADNsb polydT18-Cy5. La couleur bleu est générée par la cyanine 5 liée à l'ADN et signifie que cet ADN a diffusé à l'intérieur des cristaux.

Un jeu de données a pu être collecté (Tableau 10) et notre modèle de 6His-HpDnaB a été utilisé pour le remplacement moléculaire. Après affinement, aucune densité électronique supplémentaire n'a été observée. Plusieurs hypothèses peuvent expliquer ce résultat. Il est vraisemblable que l'ADN est bien présent dans les canaux de solvant car les cristaux sont colorés, mais celui-ci ne s'est pas fixé complètement ou de manière irrégulière. L'autre

possibilité est que la zone d'interaction soit trop dynamique pour que la densité électronique soit visible.

	ESRF-ID29
Longueur d'onde (Å)	1,28214
Groupe d'espace	I2 ₁ 3
Paramètres de la maille (Å, °)	a=b=c= 276,07; β= 90
Limite de résolution (Å)	50,40-9,0 (9,49-9,00)
Nombre total de réflexions	35239 (5387)
Nombre de réflexions uniques	2693 (386)
Complétude (%)	99,4 (100)
Redondance	13,1 (14,0)
Rmeas	0,216 (>1)
Rpim	0,06 (0,416)
Rmerge	0,207 (>1)
CC _{1/2} (%)	99,8 (61,5)
I/σ(I)	6,7 (1,8)
Facteur B moyen (Å)	336

Tableau 10 : Statistiques des données de diffraction d'un cristal de 6His-HpDnaB trempé avec un ADNsb polydT25-Cy5. Les valeurs entre parenthèses correspondent à la dernière coquille de résolution

3.3. Etude de HpDnaB en solution et en microscopie électronique

3.3.1. Etudes de HpDnaB en solution par SAXS

Les mesures SAXS de HpDnaB ont été menées dans le but de comprendre la structure de HpDnaB et son état d'oligomérisation en solution. Nous voulions également analyser les effets de l'interaction avec l'ATP, l'AMP-PNP, de l'ADNsb et de HpDnaG^{HBD} sur la structure de HpDnaB. Ainsi nous avons réalisé une série de mesure en chromatographie d'exclusion stérique couplée au SAXS (SEC-SAXS) sur différents échantillons : HpDnaB, 6His-HpDnaB seule ou en présence de nucléotides (ATP ou AMP-PNP) et également sous forme de complexes ternaires ATP ou AMP-PNP avec ADNsb ou HpDnaG^{HBD}. Pour chaque échantillon, la mesure des données a été réalisée sur le pic issu de la chromatographie d'exclusion stérique.

Une fois, les données obtenues nous avons essayé de déterminer la structure de HpDnaB dans chacune des conditions étudiées. Pour cela, nous avons utilisé notre structure cristallographique mais également généré des modèles de HpDnaB dans différentes conformation en utilisant comme modèles la structure de BstDnaB en complexe avec un ADNsb et la structure d'EcDnaB₆•EcDnaC₆. Les modèles ont été réalisés en superposant les structures cristallographiques de HpDnaB^{NTD} et de HpDnaB^{CTD} (ces dernières ayant été extraites de la structure du dodécamère) aux modèles publiés. Nous avons nommé les modèles sp-di (spirale-dilatée) lorsqu'ils ont été générés grâce à la structure de BstDnaB•ADNsb car l'hexamère est ouvert en spirale et la conformation de l'anneau NTD est dilatée. Les modèles pl-di (planaire-dilatée) sont issus de notre structure cristallographique, dans laquelle les hexamères sont plans et la conformation de l'anneau NTD est dilatée. Les modèles sp-re (spirale-resserrée) indiquent qu'ils ont été générés grâce au modèle d'EcDnaB issu du complexe EcDnaB₆•EcDnaC₆ déterminé en cryo-ME. L'hexamère est ouvert en spirale et la conformation de l'anneau NTD est resserrée (forme de « triskèle »).

Pour les modèles du complexe HpDnaB₆•HpDnaG^{HBD}₃, nous avons tout d'abord superposé les sous-unités de HpDnaB et de HpDnaG^{HBD} sur la structure du complexe BstDnaB₆•BstDnaG^{HBD}₃ (Figure 67). Pour chaque modèle, une courbe de SAXS théorique a été obtenue en utilisant le programme FoXS (Schneidman-Duhovny *et al.*, 2010 et 2013). Cette courbe a ensuite été comparée aux données expérimentales. Une valeur statistique appelée χ^2 nous a permis d'estimer la similarité entre la courbe expérimentale et une courbe théorique. Un χ^2 proche de 1, signifie que les écarts entre le modèle et les données expérimentales correspondent aux erreurs statistiques attendues. Les données SAXS nous permettent également de déterminer le rayon de giration (Rg) ainsi que la distance maximale (Dmax) de la protéine en solution. Nous avons donc pu comparer les Rg expérimentaux entre eux afin de déterminer si l'ajout de nucléotide, ADN et/ou HpDnaG^{HBD} avait un effet sur la structure de HpDnaB.

Nous avons également fait appel à la recherche des ensembles minimaux (MES) (Pelikan *et al.*, 2009). Cette méthode cherche à améliorer le χ^2 en combinant plusieurs courbes théoriques. Pour cela, plusieurs modèles sont soumis en même temps et le programme recherche les meilleures combinaisons possibles. Il détermine ensuite un pourcentage théorique pour chaque modèle permettant d'obtenir une courbe théorique similaire aux données expérimentales.

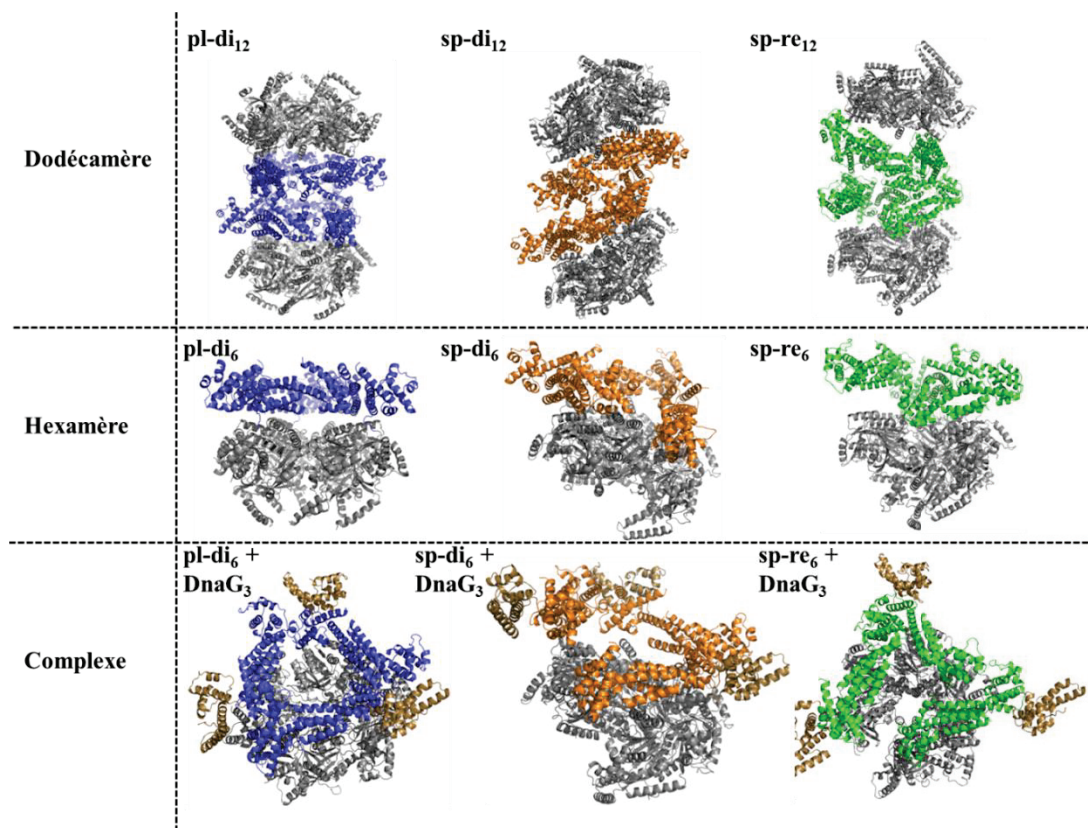


Figure 67 : Représentation des différents modèles de HpDnaB générés à partir de notre structure cristallographique pour les formes planaires, à partir de la structure de BstDnaB•GDPAIF₄ (PDB code : 4ESV) en complexe avec l'ADNsb pour les formes spirales-dilatées et à partir du modèle du complexe EcDnaB₆•EcDnaC₆ obtenu en ME pour les formes spirales-resserrées.

3.3.1.1. Influence de l'étiquette histidine sur la structure de HpDnaB

Dans un premier temps, nous avons analysé 6His-HpDnaB et HpDnaB en SEC-SAXS. Les deux protéines ont été éluées en un seul pic avec un rayon de giration stable le long du pic (Figure 68).

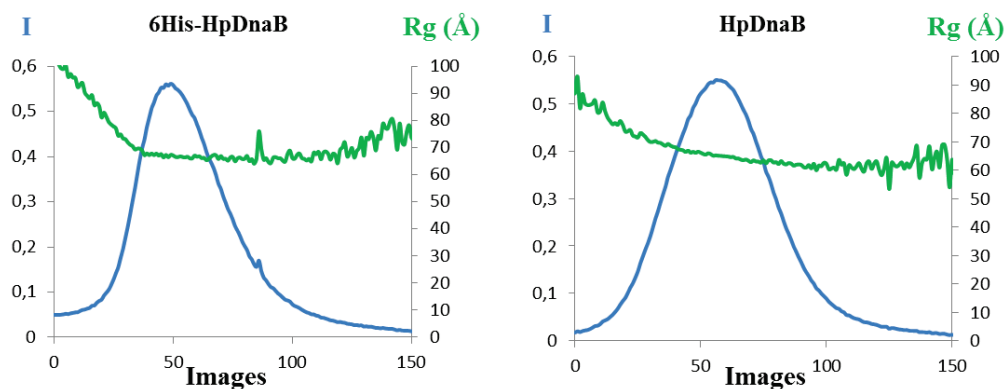


Figure 68 : Chromatogrammes de l'analyse en SEC-SAXS de 6His-HpDnaB et HpDnaB représentant l'intensité diffusée (en bleu) et le rayon de giration (en vert).

Dix expositions correspondant au pic ont été moyennées (Figure 68). L'intensité diffusée par le tampon seul avant l'élution de la protéine a ensuite été soustraite aux données des protéines (Figure 69).

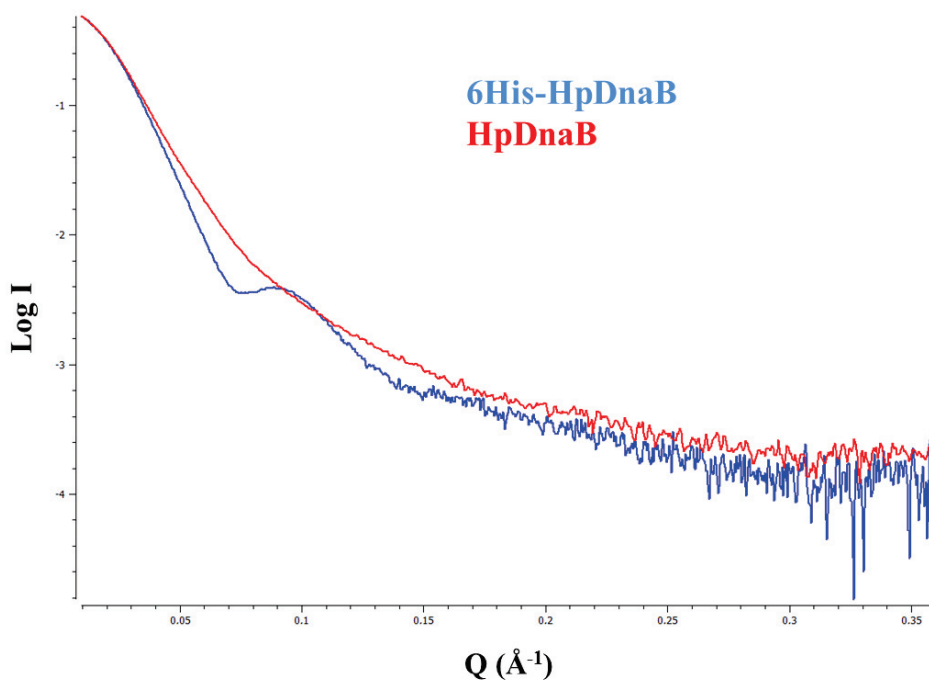


Figure 69 : Courbes SAXS expérimentales de 6His-HpDnaB et HpDnaB.

La courbe expérimentale de 6His-HpDnaB présente un épaulement par rapport à la courbe de HpDnaB. Ceci peut indiquer que HpDnaB est plus dynamique en solution que 6His-HpDnaB. Nous avons ensuite analysés les données par le logiciel PRIMUS (Konarev *et al.*, 2003). L'analyse de la région de Guinier des courbes a permis de déterminer le rayon de giration des deux protéines. 6His-HpDnaB possède un rayon de giration de 66,7 Å et une Dmax de 234 Å

alors que HpDnaB possède un rayon de giration de 65,3 Å et une Dmax de 228 Å. Les résultats indiquent que les rayons de giration et les Dmax de 6His-HpDnaB et de HpDnaB sont très similaires (Figure 70).

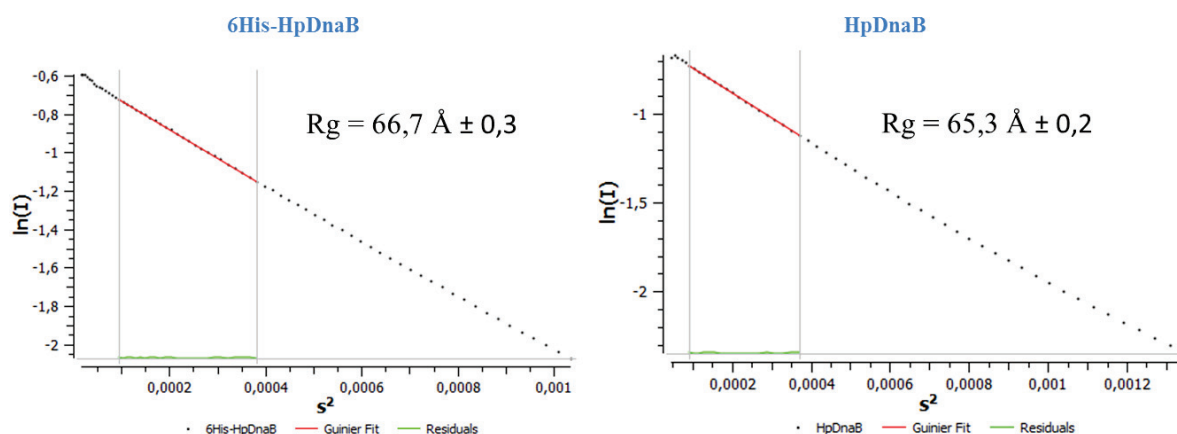


Figure 70 : Calculs des rayons de giration de 6His-HpDnaB et HpDnaB à partir de la région de Guinier des données SAXS expérimentales.

Nous avons ensuite utilisé les modèles pour générer des courbes SAXS théoriques et les comparer aux données expérimentales en utilisant FoXS (Tableau 11).

Modèles	Données SAXS			
	6His-HpDnaB		HpDnaB	
	FoXs χ^2	FoXs MES $\chi^2 = 6,5$	FoXs χ^2	FoXs MES $\chi^2 = 8,2$
sp-re ₁₂	15,5		18,6	
pl-di ₁₂	11,7	36%	23,6	
sp-di ₁₂	8,5	64%	9	94%
pl-di ₆	80,2		59,8	6%
sp-re ₆	84		60,8	
pl-re ₆	101,4		74,8	
sp-di ₆	90,8		65,1	
sp-re ₆ /DnaG ₃	57,2		37,7	
sp-di ₆ /DnaG ₃	65,8		43,8	
pl-di ₆ /DnaG ₃	58,6		44,9	

Tableau 11 : Tableau représentant l'ensemble des χ^2 des différents modèles testés ainsi que les résultats de l'expérience de MES. Les calculs ont été réalisés sur le serveur web FoXs.

Le modèle qui correspond le plus aux données expérimentales de 6His-HpDnaB et HpDnaB est le modèle sp-di₁₂ avec un χ^2 de 8,5 avec les données de 6His-HpDnaB et un χ^2 de 9 avec

les données de HpDnaB. L'analyse MES indique qu'un mélange de deux conformations permet d'améliorer le χ^2 . La combinaison de la conformation sp-di₁₂ (64%) et de la conformation pl-di₁₂ (36%) permet d'obtenir une courbe théorique qui correspond à la courbe expérimentale avec un χ^2 de 6,5 (Figure 71). L'analyse des données expérimentales de HpDnaB indique que contrairement au cas précédent, le mélange de plusieurs conformations différentes n'améliore que très peu le χ^2 (8,2 au lieu de 9). Le MES indique qu'un mélange de la conformation sp-di₁₂ (94%) avec la conformation pl-di₆ (6%) permet d'obtenir une courbe théorique comparable aux données expérimentales (Figure 71). Etant donné, la faible amélioration du χ^2 par le MES, la courbe théorique générée avec la conformation sp-di₁₂ seule et celle issue du MES sont les mêmes.

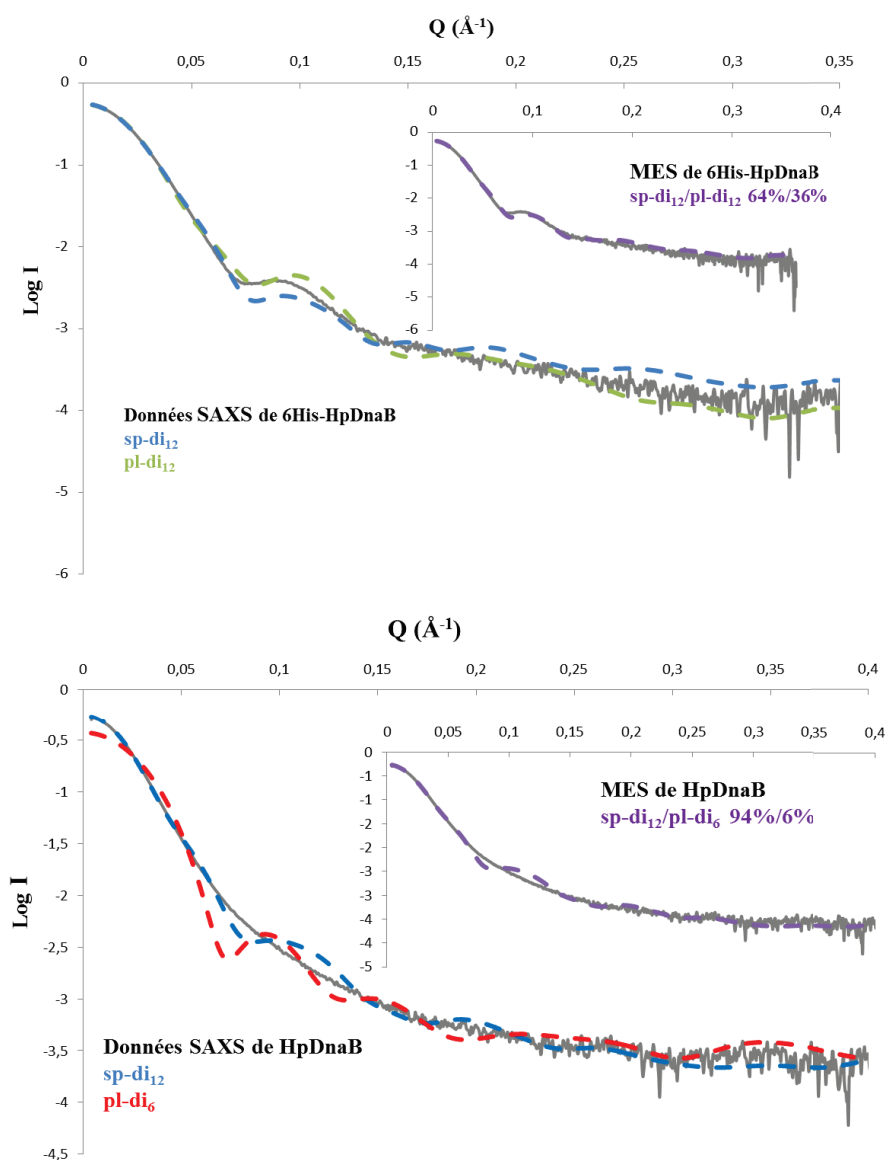


Figure 71 : Comparaison des données SAXS expérimentales de 6His-HpDnaB avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₆ (rouge) et la courbe du MES (violet).

En conclusion, les expériences de SEC-SAXS indiquent que le rayon de giration ainsi que la D_{max} de 6His-HpDnaB et de HpDnaB correspondent aux valeurs d'un dodécamère. Cependant, la conformation du dodécamère semble différer en fonction de la présence ou non de l'étiquette histidine. En effet, 6His-HpDnaB adopterait une conformation planaire-dilatée comme dans le cristal mais également une conformation spirale-dilatée. Au contraire, la conformation de HpDnaB semble proche de la conformation spirale-dilatée mais que cette dernière est dynamique. Nous en avons ainsi déduit que la présence de l'étiquette histidine permettrait de favoriser une conformation planaire du dodécamère qui serait également moins dynamique. Ceci pourrait expliquer que la protéine ait cristallisé dans cette conformation.

3.3.1.2. Effet des nucléotides sur la structure de HpDnaB

Nous avons étudié ensuite l'influence de l'ATP et de l'AMP-PNP sur HpDnaB en SEC-SAXS. Comme précédemment, les deux protéines sont éluées en un seul pic avec un rayon de giration stable dans le cas de 6His-HpDnaB ou variant légèrement pour HpDnaB + ATP et HpDnaB + AMP-PNP (Figure 72).

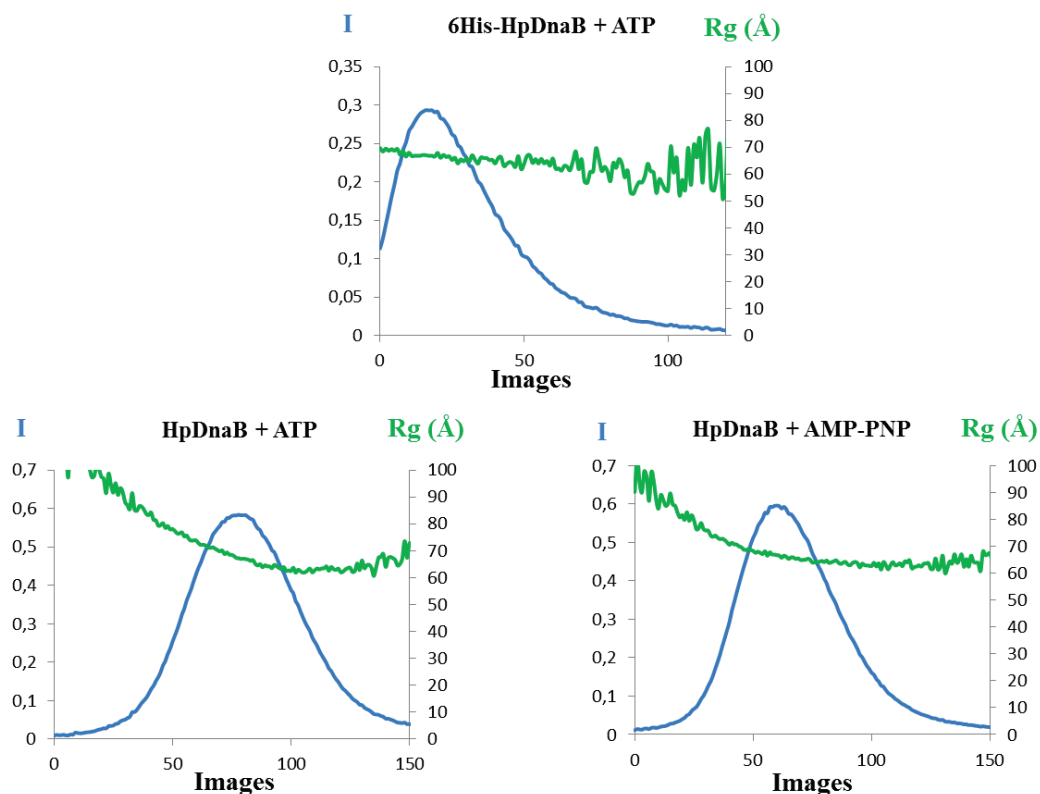


Figure 72 : Chromatogrammes de l'analyse en SEC-SAXS de 6His-HpDnaB + ATP et HpDnaB + ATP ou AMP-PNP représentant l'intensité diffusée (bleu) et le rayon de giration (vert).

Les courbes SAXS générées de 6His-HpDnaB et de HpDnaB ont une forme différente. En effet, les courbes expérimentales de HpDnaB avec ATP ou AMP-PNP sont similaires alors que celle de 6His-HpDnaB + ATP présente un épaulement. Comme précédemment, ceci peut indiquer que HpDnaB en présence de nucléotide est plus dynamique que 6His-HpDnaB (Figure 73).

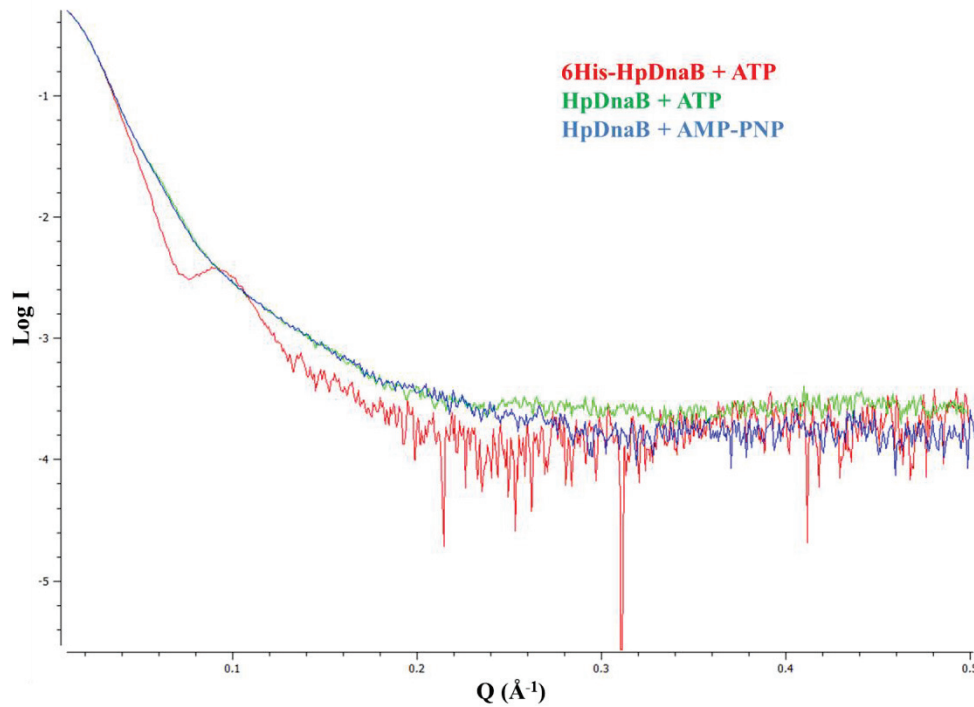


Figure 73 : Courbes SAXS expérimentales de 6His-HpDnaB + ATP en rouge, HpDnaB + ATP en vert et HpDnaB + AMP-PNP en bleu.

L'analyse de la région de Guinier de la courbe de 6His-HpDnaB + ATP indique un R_g de 66,6 Å avec une D_{max} de 228 Å. De même, pour HpDnaB + ATP ou AMP-PNP, les R_g s sont respectivement de 66,8 Å et 66,7 Å et la D_{max} est de 234 Å dans les deux cas (Figure 74).

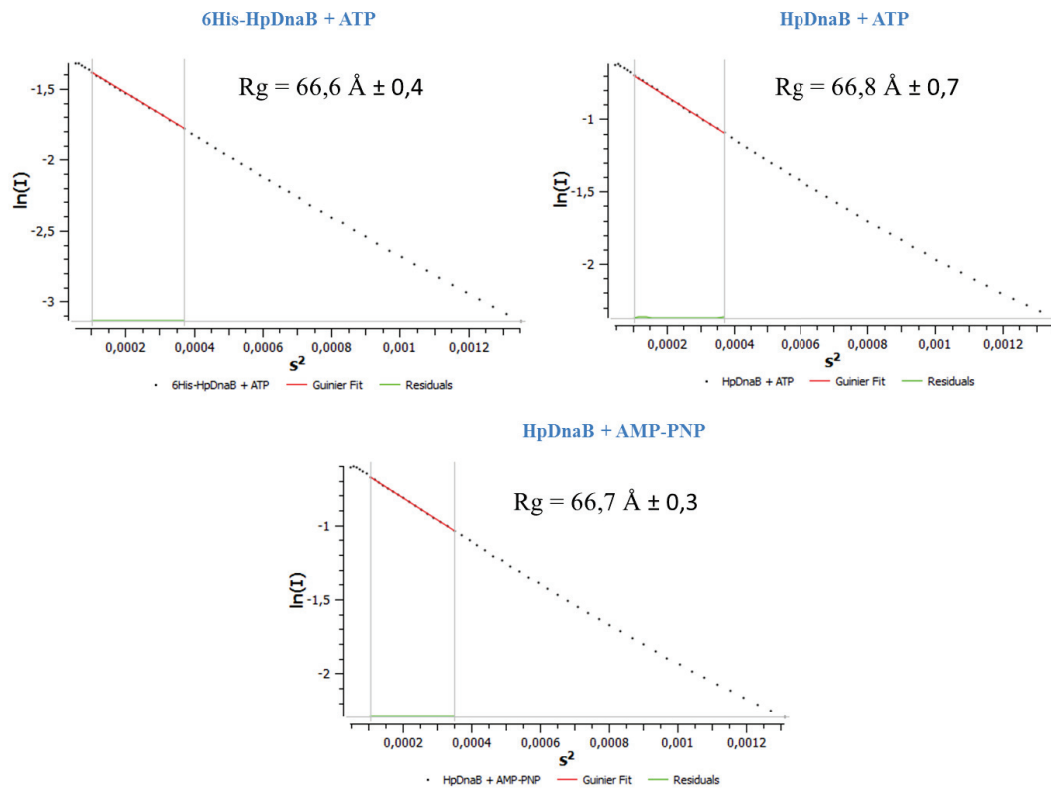


Figure 74 : Calculs des rayons de giration de 6His-HpDnaB + ATP, HpDnaB + ATP et HpDnaB + AMP-PNP à partir de la région de Guinier des données SAXS expérimentales

Ces expériences montrent que l'ATP n'a pas d'effet sur l'état oligomérique de HpDnaB ou de 6His-HpDnaB. Comme précédemment, il y a donc peut être plusieurs conformations du dodécamère présentes en solution. Nous avons donc continué l'analyse en essayant de trouver quelles étaient les conformations du dodécamère dans les deux cas en utilisant nos modèles théoriques (Tableau 12).

Modèles	Données SAXS					
	6His-HpDnaB + ATP		HpDnaB + ATP		HpDnaB + AMP-PNP	
	FoXs χ^2	FoXs MES $\chi^2 = 4,2$	FoXs χ^2	FoXs MES $\chi^2 = 12,8$	FoXs χ^2	FoXs MES $\chi^2 = 12,0$
sp-re ₁₂	9,6		19,3		18,8	
pl-di ₁₂	7,2	40%	28,8	7%	28,3	4%
sp-di ₁₂	5,7	60%	12,9	93%	12	96%
pl-di ₆	51		76,3		75,2	
sp-re ₆	53,5		75,3		74,6	
pl-re ₆	64,5		87,8		87,9	
sp-di ₆	57,8		78,1		77,8	
sp-re ₆ /DnaG ₃	36,8		52,9		51,1	
sp-di ₆ /DnaG ₃	42,2		57,9		56,6	
pl-di ₆ /DnaG ₃	37,5		63,7		61,6	

Tableau 12 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de 6His-HpDnaB + ATP, HpDnaB + ATP et HpDnaB + AMP-PNP. Les calculs ont été réalisés sur le serveur web FoXs.

Pour 6His-HpDnaB + ATP, la comparaison des différents modèles laisse supposer que la dodécamère adopte une conformation sp-di₁₂ avec un $\chi^2 = 5,7$. L'analyse en MES indique que le mélange de deux conformations différentes permet d'améliorer le χ^2 passant de 5,7 à 4,2 (Tableau 12). La courbe théorique générée par MES résulte de la combinaison de 60% sp-di₁₂ et de 40% pl-di₁₂ (Figure 75, Tableau 12).

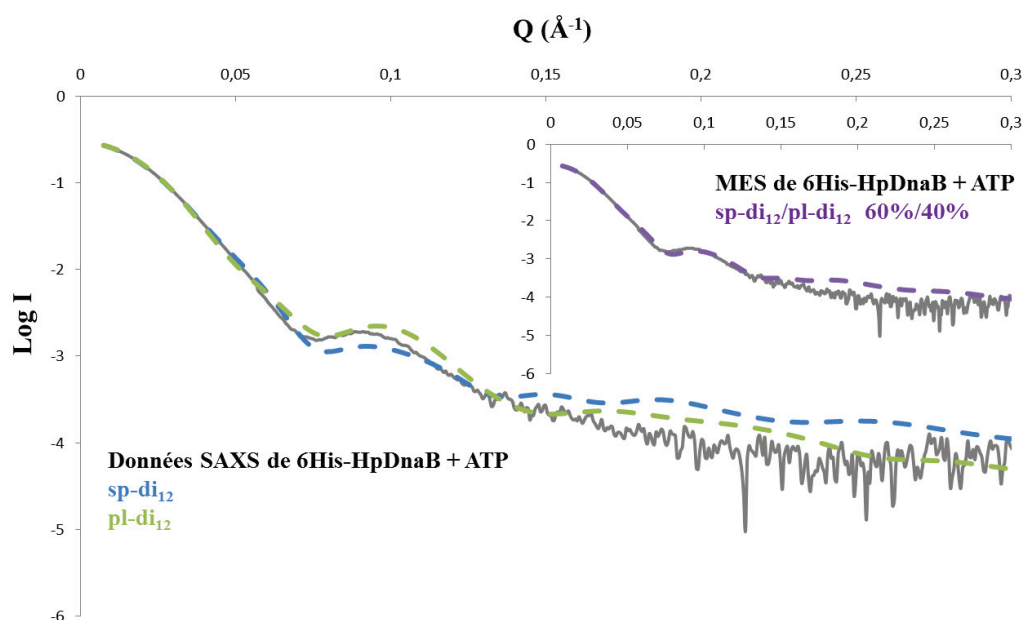


Figure 75 : Comparaison des données SAXS expérimentales de 6His-HpDnaB + ATP avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₁₂ (vert) et la courbe du MES (violet).

Pour HpDnaB + ATP, le modèle théorique qui correspond le mieux aux données SAXS est sp-di₁₂ avec un χ^2 égal à 12,9 (Tableau 12). L'analyse MES n'améliore que très peu le χ^2 et montre que le dodécamère en présence d'ATP est majoritairement sous forme sp-di₁₂ à 93% et à 7 % sous une forme pl-di₁₂. Les χ^2 entre le modèle sp-di₁₂ seul et le MES étant très proches, leurs courbes théoriques sont similaires et sont quasiment identiques (Figure 76, Tableau 12).

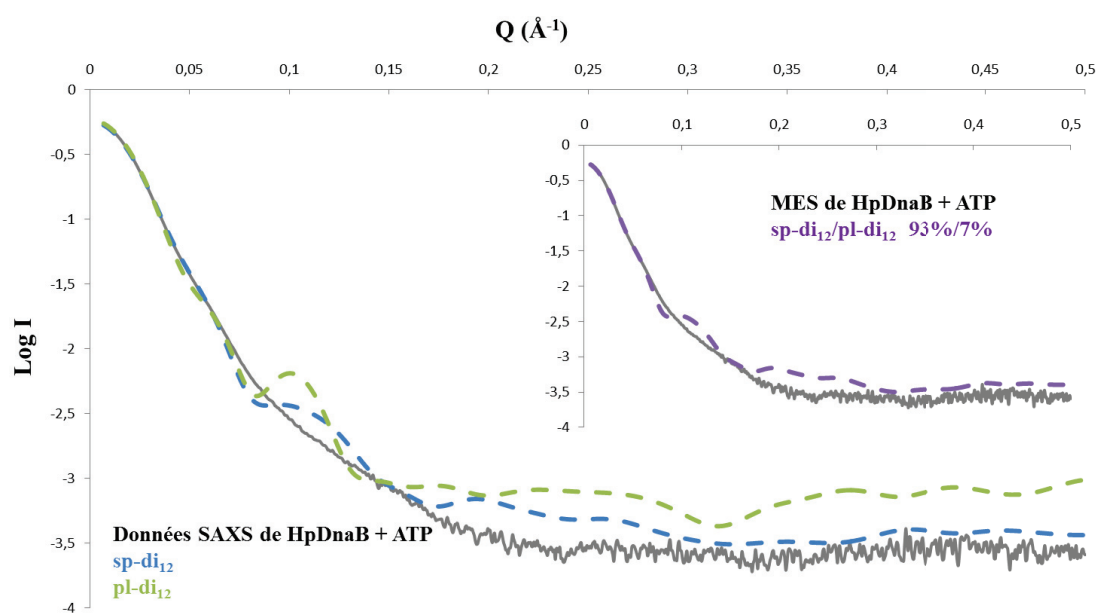


Figure 76 : Comparaison des données SAXS expérimentales de HpDnaB + ATP avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₁₂ (vert) et la courbe du MES (violet).

Le modèle qui décrit le mieux les données expérimentales est le modèle en sp-di₁₂ avec un χ^2 égal à 12 (Tableau 12). Comme précédemment, l'analyse en MES n'améliore pas le χ^2 et indique que HpDnaB + AMP-PNP est majoritairement sp-di₁₂ (96%) et minoritairement pl-di₁₂ (4%) (Figure 77, Tableau 12).

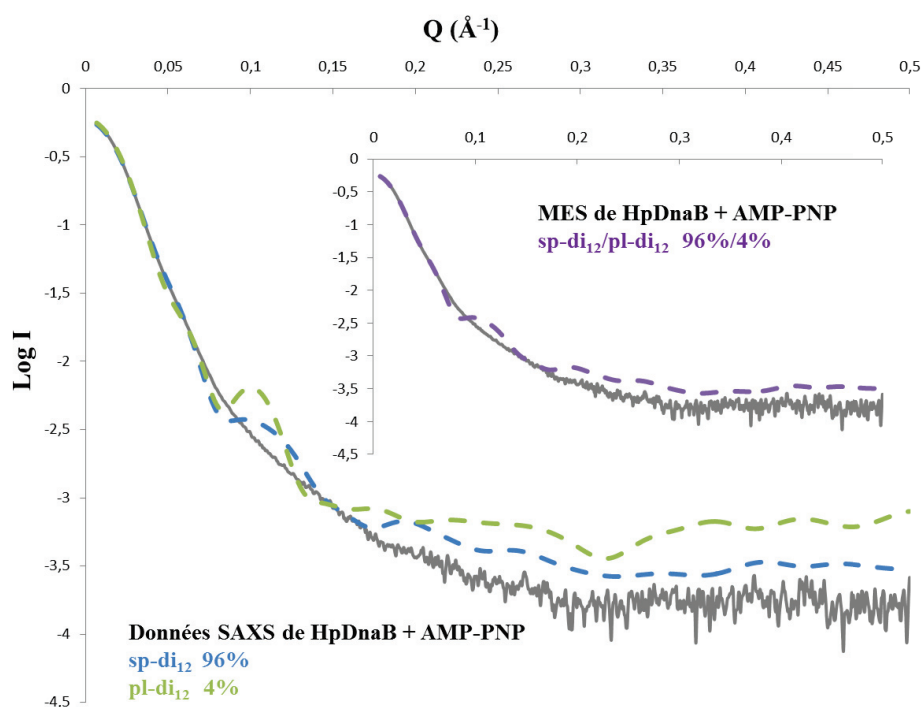


Figure 77 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP avec les courbes théoriques calculées à partir des modèles sp-di_{12} (bleu), pl-di_{12} (vert) et la courbe du MES (violet).

La présence d'AMP-PNP semble donc avoir le même effet que l'ATP dont l'hydrolyse ne semble pas engendrer de larges mouvements. En effet, la présence ou non de nucléotide ne semble pas provoquer de changements de conformation important. Que ce soit pour 6His-HpdnaB ou HpDnaB, l'état oligomérique majoritaire reste le dodécamère avec une conformation proche de sp-di_{12} .

3.3.2. Microscopie électronique de HpDnaB

La structure de 6His-HpDnaB d'après l'étude en ME avait montré que la protéine formait des dodécamères. Dans le but de s'assurer que la conformation de HpDnaB est bien un dodécamère, j'ai effectué une étude de ME dans l'équipe du Dr Irina Gutsche à l'UVHCI de Grenoble. Nous avons préparé des grilles de carbone contenant HpDnaB qui ont été trempées dans l'acétate d'uranyl (voir partie 2.6). L'homogénéité et la qualité de l'échantillon ont été vérifiées sur un microscope JEOL 1200 EX II à 100keV. Des images ont été collectées sur film photographique à un grossissement de 40000 fois puis scannées à 21 \AA grâce à un scanner Zeiss/Intergraph Photoscan TD de l'équipe d'accueil. Avec un grossissement 40000 fois et un scan à 21 \AA , cela représente une résolution de 5,25 $\text{\AA}/\text{pixel}$. Les images scannées

permettent de sélectionner les particules de protéine grâce au programme Boxer du logiciel EMAN (Ludtke *et al.*, 1999) dans des boîtes de 57x57 pixels. Les images ont ensuite été corrigées par la fonction de transfert de contraste (CTF) et des classes ont été générées par le logiciel IMAGIC (van Heel *et al.*, 1996). Les classes sont utilisées comme série de référence pour l'alignement de chaque particule sélectionnée avec le programme Spider (Frank *et al.*, 1996; Shaikh *et al.*, 2008). De nouvelles classes sont générées, les meilleures sont sélectionnées et permettent d'améliorer l'alignement des particules. Ce processus est itératif et a pour but d'obtenir les meilleures classes possibles, c'est-à-dire d'obtenir le maximum d'orientations possibles. A partir de cette étape, il est possible de déterminer les angles d'Euler pour définir la symétrie de l'objet.

Nous avons pu ainsi obtenir des classes pour HpDnaB. Les classes obtenues sont essentiellement des vues du dessus mais on observe également quelques vues de côté (Figure 78). Ces dernières montrent que HpDnaB forme des particules similaires à 6His-HpDnaB. En effet il est possible d'observer que les 4 anneaux (2 anneaux NTD et 2 anneaux CTD) sont présents et qu'ils s'associent par les NTD. Cependant, il n'a pas été possible d'obtenir une reconstruction 3D dû fait de l'hétérogénéité de l'échantillon.

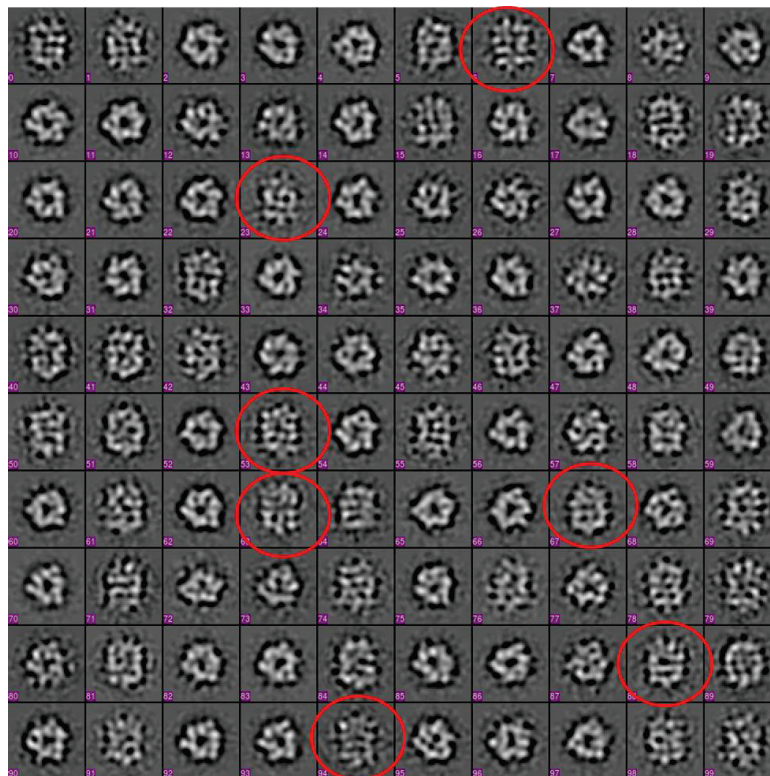


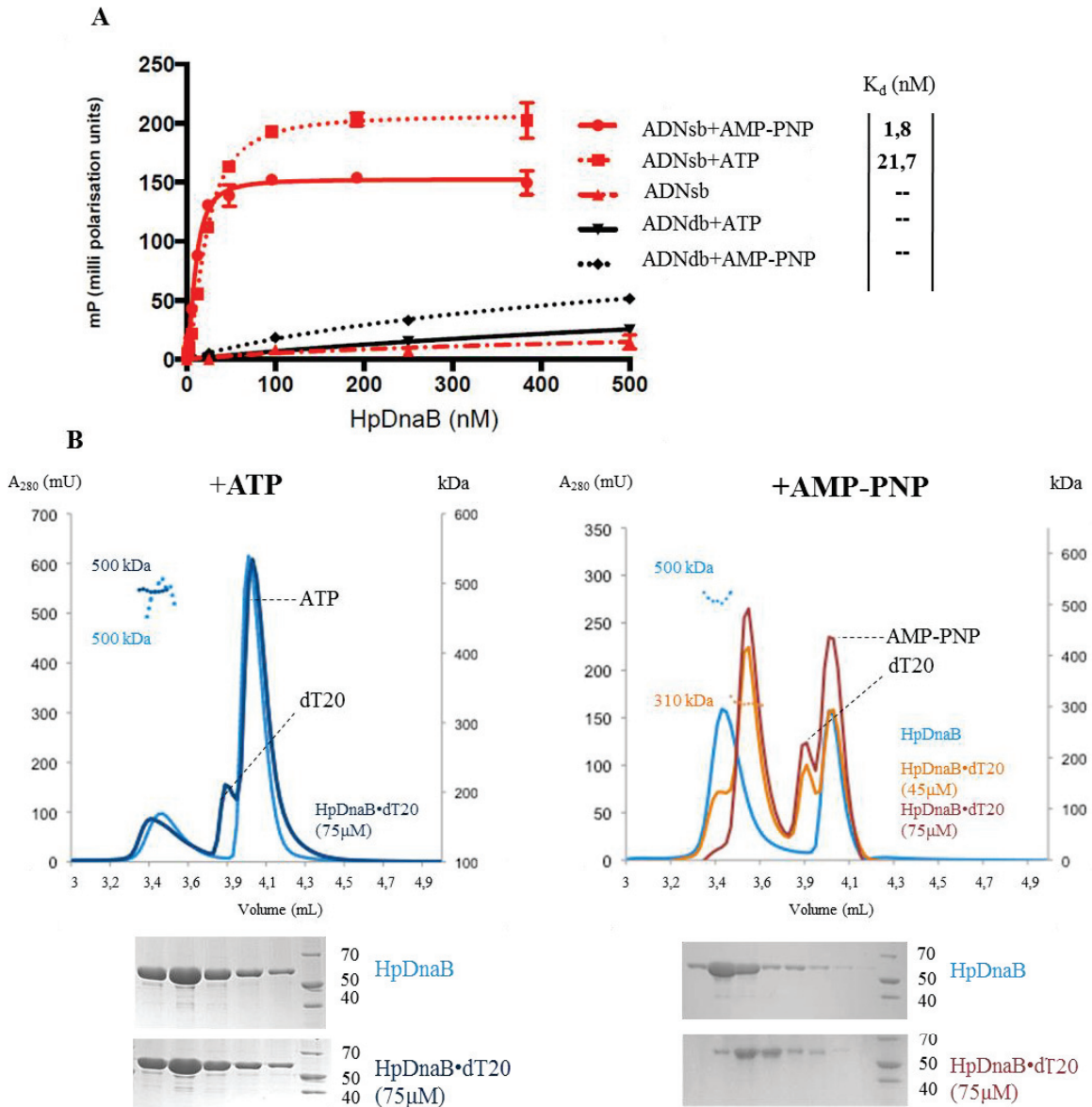
Figure 78 : Classes de HpDnaB obtenues en ME à coloration négative

3.3.3. Etude de l'interaction de HpDnaB avec l'ADN

Nos expériences de SAXS ont montré que HpDnaB adoptait une conformation en dodécamère en présence d'un oligonucléotide polydT20 (oligodT20). Nous avons donc analysé cette interaction par anisotropie de fluorescence en collaboration avec Joanna Timmins à l'IBS de Grenoble. Ses expériences ont déterminé que HpDnaB interagit avec l'ADN avec une constante de dissociation (K_d) de 21,7 nM en présence d'ATP et 1,9 nM en présence d'AMP-PNP. Contrairement à ces affinités élevées pour l'ADNsb, HpDnaB n'a que très peu d'affinité pour l'ADNdb. De plus, l'affinité pour l'ADNsb diminue très fortement si le nucléotide n'est plus présent (Figure 79A).

Nous avons ensuite analysé HpDnaB avec ATP en présence ou non d'un oligodT20 en SEC-MALS. Le profil d'élution à 280 et 260 nm montre que HpDnaB en présence d'ATP à une masse calculée de 500 kDa ce qui ne correspond pas à la masse théorique d'un dodécamère. Cependant, le même échantillon analysé en SEC-SAXS (voir partie 3.3.1.2) possède un rayon de giration de 66,6 Å et une D_{max} de 234 Å ce qui correspond aux paramètres d'un dodécamère. En effet, la masse calculée à partir des données SAXS grâce au serveur SAXS Mow (Fischer *et al.*, 2010) est de 632,8 kDa. De plus, la masse calculée grâce au volume de Porod est de 554,8 kDa ce qui est en accord avec la masse calculée en SEC-MALS. Lorsque l'oligodT20 est rajouté, la présence de l'ADN ne change ni la masse calculée ni le volume d'élution de HpDnaB ce qui montre que l'ADNsb n'a pas d'effet sur le dodécamère en présence d'ATP (Figure 79B)

La même expérience réalisée en remplaçant l'ATP par l'AMP-PNP montre qu'en présence de l'oligodT20 en concentration croissante, le volume d'élution du pic à 280 nm correspondant à HpDnaB est augmenté signifiant que la protéine est plus petite (Figure 79B). En effet, la masse calculée par SEC-MALS est de 500 kDa dans les conditions sans ADNsb alors que la masse est de 310 kDa lorsque l'ADNsb est rajouté. En bloquant ainsi l'hydrolyse de l'ATP, HpDnaB se lierait à l'ADN sous forme d'hexamère, ce qui est en accord avec nos analyses en SAXS (voir partie 3.3.1.3).



3.3.3.1. Etude par SAXS de l'effet de l'ADN sur HpDnaB

Nous avons ensuite étudié la structure de HpDnaB en présence d'un ADNsb polyT de 20 nucléotides (oligodT20), et ce en présence soit d'ATP soit d'AMP-PNP. Dans les deux cas la protéine est éluée en un seul pic avec un rayon de giration plus ou moins stable en fonction de la présence d'ATP ou d'AMP-PNP.

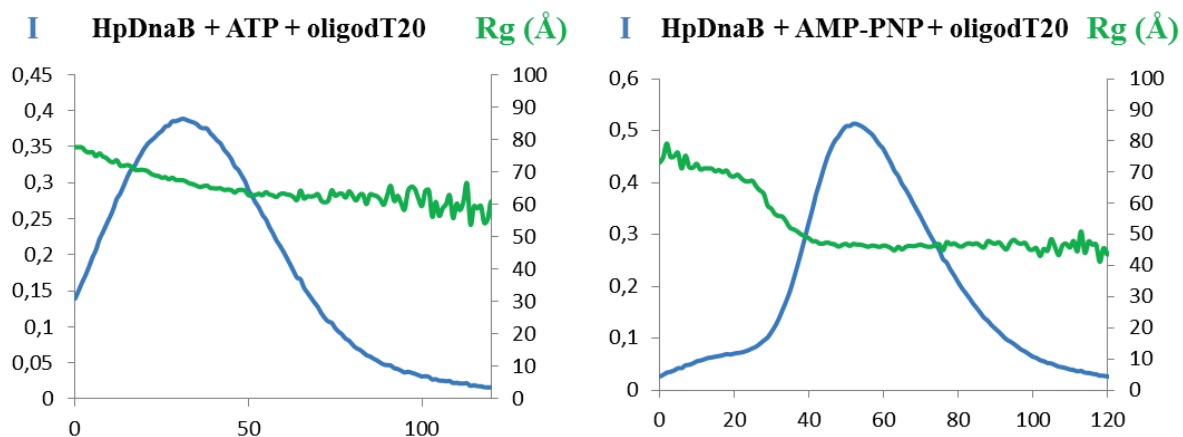


Figure 80 : Chromatogrammes de l'analyse en SAXS de HpDnaB + ATP ou AMP-PNP + oligodT20 représentant l'intensité diffusée (bleu) et le rayon de giration (vert).

Les courbes SAXS indiquent que les deux échantillons présentent des différences structurales (Figure 81). En effet le $I(0)$ n'est pas identique si on compare les deux courbes, ce qui laisse supposer une différence de poids moléculaire. De plus, la courbe en présence d'AMP-PNP présente un épaulement comparé à la courbe avec ATP indiquant un certain dynamisme.

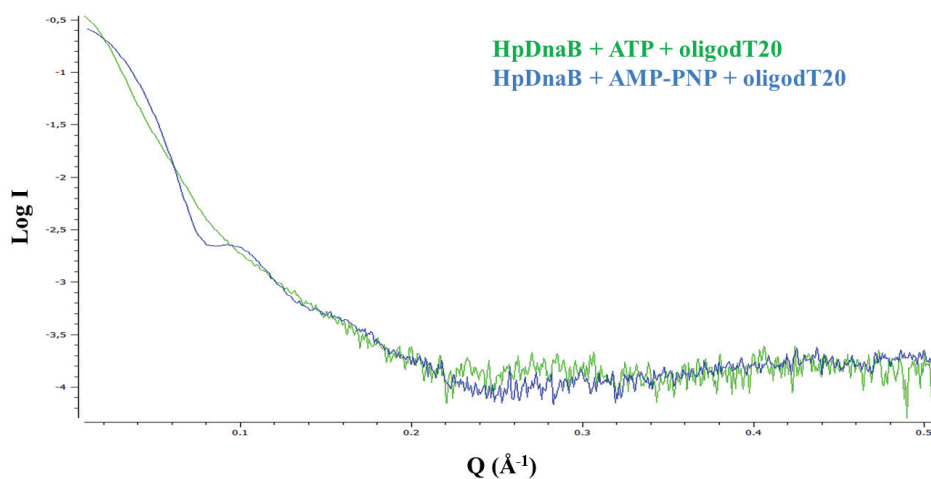


Figure 81 : Courbes SAXS expérimentales de HpDnaB + ATP + oligo-dT20 en vert, HpDnaB + AMP-PNP + oligo-dT20 en bleu.

Les différences se confirment lorsque les régions de Guinier des deux courbes sont analysées. En effet, l'échantillon de HpDnaB+ATP+oligodT20, présente un R_g de 67,1 Å et une D_{max} de 235 Å (Figure 82). Alors que la courbe SAXS de HpDnaB+AMP-PNP+oligodT20 qui a

une forme différente de HpDnaB+ATP et de HpDnaB+ATP+oligodT20, présente un R_g calculé de 46,8 Å et une D_{max} de 146 Å (Figure 82).

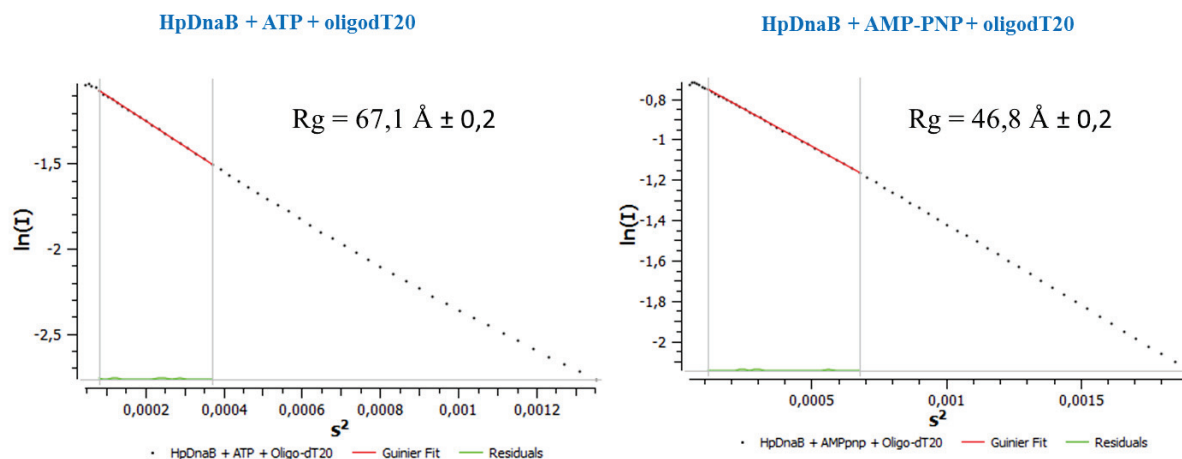


Figure 82 : Calculs des R_g de HpDnaB + ATP + oligodT20, HpDnaB + AMP-PNP+ oligodT20 à partir de la région de Guinier des données SAXS expérimentales

La courbe expérimentale de HpDnaB + ATP + oligodT20 a donc des valeurs de R_g et de D_{max} qui sont comparables à l'échantillon HpDnaB + ATP sans ADN. Alors que dans le cas de HpDnaB + AMP-PNP + oligodT20, les valeurs de R_g et de D_{max} ne correspondent plus à celles d'un dodécamère mais à celles d'un hexamère. Nous avons donc continué notre analyse en comparant les courbes théoriques de nos différents modèles avec les données expérimentales de ces deux échantillons (Tableau 13).

Modèles	Données SAXS			
	HpDnaB + ATP + oligodT20		HpDnaB + AMP-PNP + oligodT20	
	FoXs χ^2	FoXs MES $\chi^2 = 8,4$	FoXs χ^2	FoXs MES $\chi^2 = 7,8$
sp-re ₁₂	13,6		102,1	
pl-di ₁₂	21,3	5%	122	
sp-di ₁₂	8,4	95%	94,6	
pl-di ₆	55,8		20,1	
sp-re ₆	55,4		14,2	
pl-re ₆	65,3		11,5	49%
sp-di ₆	57,9		11,2	51%
sp-re ₆ /DnaG ₃	38,1		35,9	
sp-di ₆ /DnaG ₃	42,2		27	
pl-di ₆ /DnaG ₃	45,7		32,6	

Tableau 13 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de HpDnaB + ATP + oligo dT20, HpDnaB + AMP-PNP+ oligo dT20. Les calculs ont été réalisés sur le serveur web FoXs.

Ces comparaisons montrent que dans le cas de HpDnaB + ATP + oligodT20, la courbe théorique qui correspond le plus aux données expérimentales est celle d'un dodécamère sp-di₁₂ avec un χ^2 de 8,4 (Tableau 13). L'analyse en MES confirme qu'en présence d'ATP et de l'oligodT20, HpDnaB adopte majoritairement cette forme puisque la meilleure combinaison obtenue est 95% de sp-di₁₂ et 5% de pl-di₁₂ (Figure 83, Tableau 13).

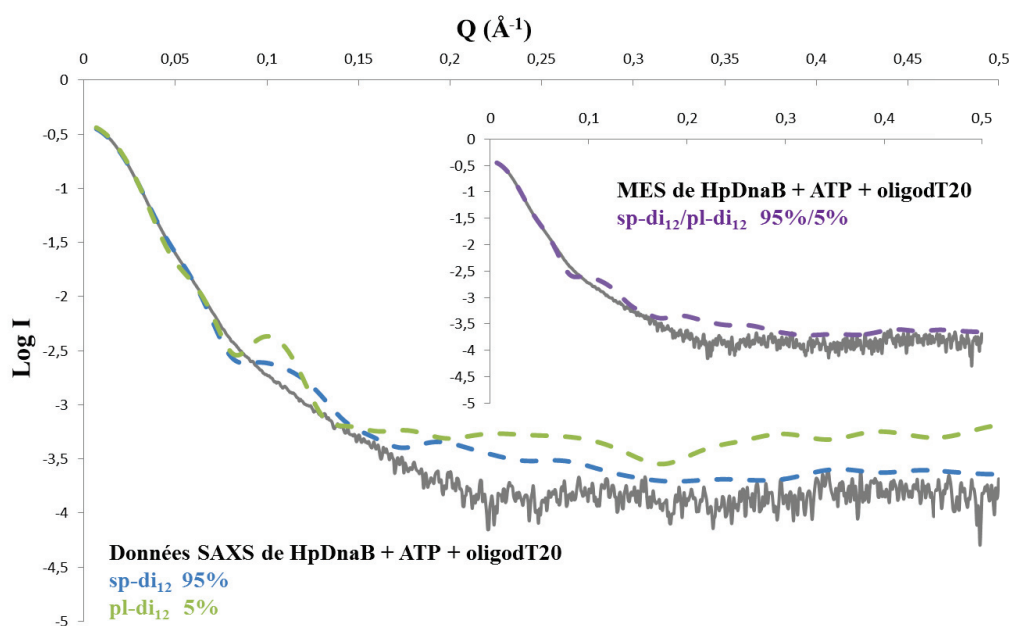


Figure 83 : Comparaison des données SAXS expérimentales de HpDnaB + ATP + oligodT20 avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₁₂ (vert) et la courbe du MES (violet).

En présence d'AMP-PNP, c'est la courbe théorique d'un hexamère en $sp\text{-}di_6$ qui correspond le mieux aux données expérimentales avec un χ^2 de 11,2 (Tableau 13). La combinaison de plusieurs modèles en MES améliore le χ^2 (7,8). La courbe théorique qui en résulte est générée par la combinaison du modèle d'hexamère $sp\text{-}di_6$ (51%) et du modèle $sp\text{-}re_6$ (49%) (Figure 84, Tableau 13).

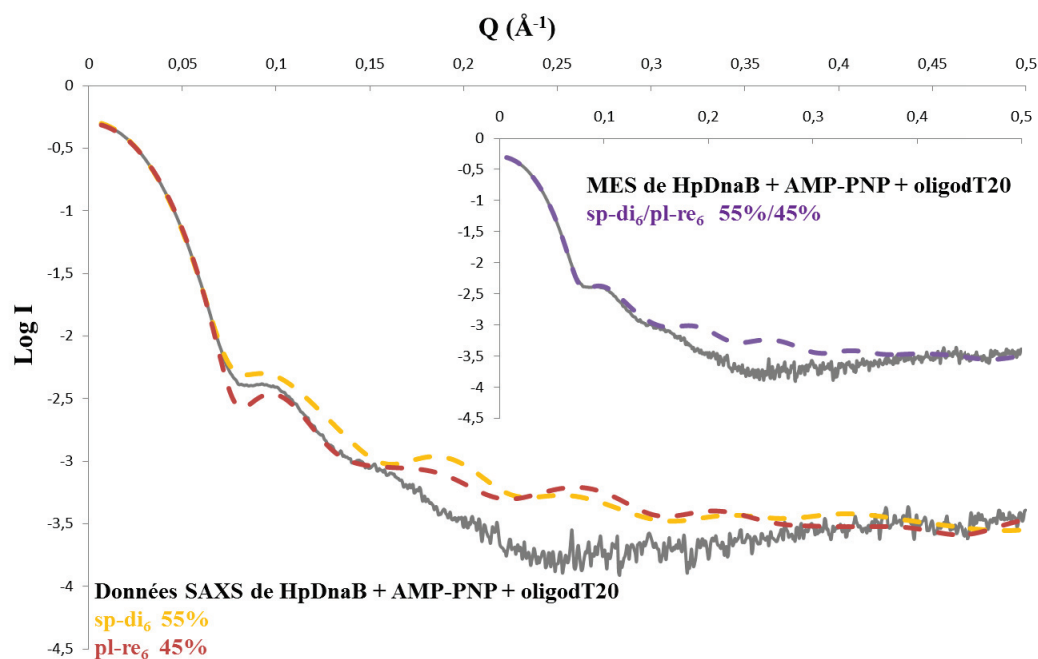


Figure 84 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP + oligodT20 avec les courbes théoriques calculées à partir des modèles $sp\text{-}di_6$ (jaune), $pl\text{-}re_6$ (rouge) et la courbe du MES (violet).

HpDnaB reste donc en dodécamère lorsque la protéine est en présence d'ADN et d'ATP. En présence d'AMP-PNP et en interagissant avec l'ADNsb, la conformation de HpDnaB semble être un hexamère. Les données SAXS indiquent également que ces hexamères pourraient adopter plusieurs conformations. Ces dernières pourraient être $pl\text{-}re_6$ et $pl\text{-}di_6$, bien que d'autres conformations non envisagées dans nos modèles soient possibles. De manière intéressante, la forme spirale-dilatée correspond à la structure de BstDnaB en complexe avec l'ADN. Il est donc vraisemblable que HpDnaB sous forme d'hexamère et lié à l'ADN adopte une conformation similaire.

3.4. Etude biochimique de l'interaction HpDnaB•HpDnaG^{HBD}

Afin d'étudier la formation du primosome chez *H. pylori*, nous avons étudié le complexe HpDnaB•HpDnaG^{HBD} en SEC-SAXS et en SEC-MALS.

3.4.1. Etude par SAXS de l'effet de HpDnaG^{HBD} sur HpDnaB

Nous avons commencé par étudier l'effet de HpDnaG^{HBD} sur HpDnaB en SEC-SAXS. Pour cela, nous avons mélangé les deux protéines dans des proportions équimolaires en présence ou non d'ATP ou d'AMP-PNP. Pour chaque échantillon, l'analyse a été réalisée sur les pics obtenus à l'issue de la chromatographie d'exclusion stérique (Figure 85).

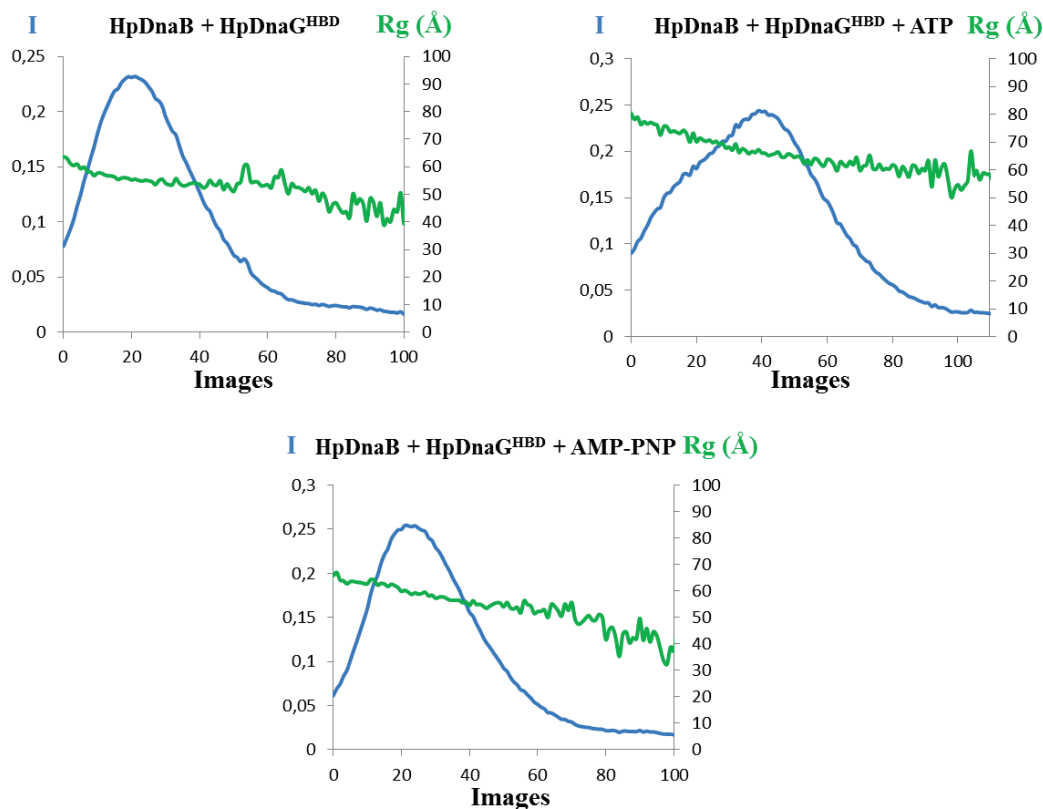


Figure 85 : Chromatogrammes de l'analyse en SEC-SAXS de HpDnaB + HpDnaG^{HBD} seules ou en présence d'ATP ou AMP-PNP, représentant l'intensité diffusée (bleu) et le rayon de giration (vert).

Nous avons ainsi pu générer comme précédemment des courbes SAXS que nous avons ensuite analysées (Figure 86).

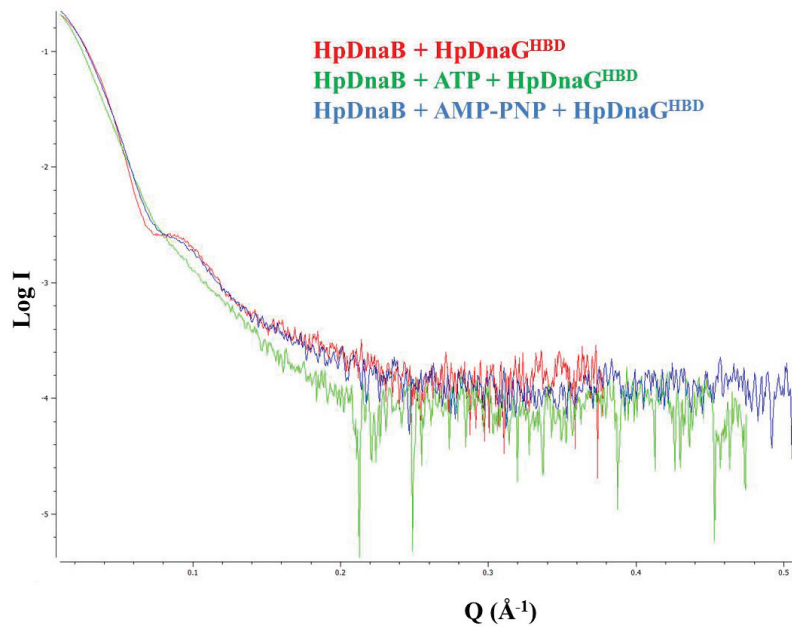


Figure 86 : Courbes SAXS expérimentales de HpDnaB + HpDnaG^{HBD} en rouge, HpDnaB+ ATP+ HpDnaG^{HBD} en vert et HpDnaB + AMP-PNP + HpDnaG^{HBD} en bleu.

Dans le cas de HpDnaB + HpDnaG^{HBD}, le Rg calculé est égal à 55,2 Å avec une Dmax de 193 Å. Lorsque de l'ATP est présent le Rg du complexe est de 65,8 Å et la Dmax est de 230 Å. Puis quand l'ATP est remplacé par de l'AMP-PNP, le Rg est égale à 60,3 Å et la Dmax est de 211 Å (Figure 87).

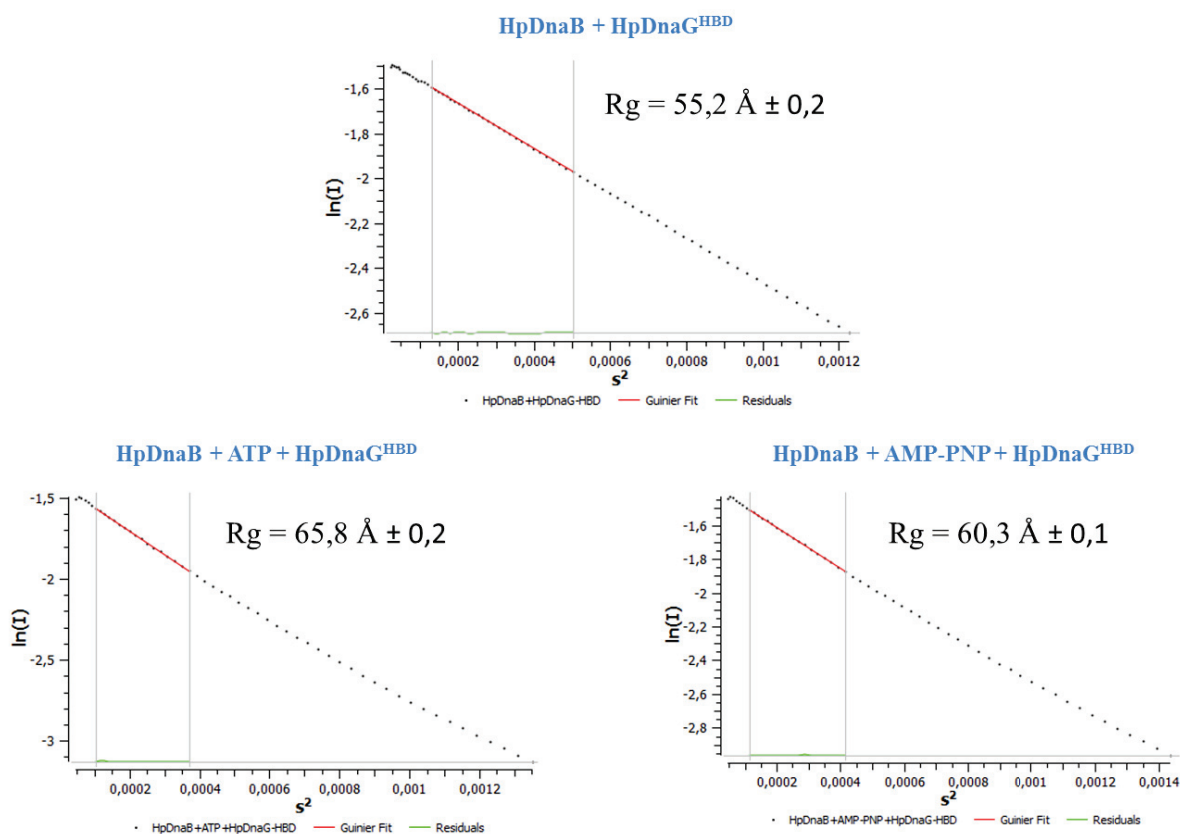


Figure 87 : Calculs des rayons de gyration de de HpDnaB + HpDnaG^{HBD}, HpDnaB+ ATP+ HpDnaG^{HBD} et HpDnaB + AMP-PNP + HpDnaG^{HBD} à partir de la région de Guinier des données SAXS expérimentales

Les valeurs de R_g et de D_{max} de HpDnaB + HpDnaG^{HBD} seul ou en présence d'AMP-PNP ne correspondent pas aux valeurs d'un hexamère ni à celles d'un dodécamère mais se rapproche plus de valeurs d'un complexe HpDnaB•HpDnaG^{HBD}. Cependant lorsque l'AMP-PNP est remplacé par de l'ATP, les valeurs de R_g et de D_{max} changent et sont comparables à celles d'un dodécamère. Afin d'identifier la conformation de HpDnaB en présence de HpDnaG^{HBD}, nous avons poursuivi notre analyse en comparant comme précédemment les courbes théoriques de nos modèles avec les courbes expérimentales des trois échantillons (Tableau 14).

Modèles	Données SAXS					
	HpDnaB + HpDnaG ^{HBD}		HpDnaB + ATP+HpDnaG ^{HBD}		HpDnaB + AMP-PNP +HpDnaG ^{HBD}	
	FoXs χ^2	FoXs MES $\chi^2 = 2,3$	FoXs χ^2	FoXs MES $\chi^2 = 5,5$	FoXs χ^2	FoXs MES $\chi^2 = 3,5$
sp-re ₁₂	35,1		12,8		25,1	40%
pl-di ₁₂	36,8		19,9		30,2	
sp-di ₁₂	28,4	16%	7,3	90%	18,9	
pl-di ₆	16,1		35,3	10%	25,4	
sp-re ₆	18,5		35,3		25,8	
pl-re ₆	31		43,2		36,4	
sp-di ₆	23,2		37,4		29,1	
sp-re ₆ /DnaG ₃	3,1		22,5		11,6	
sp-di ₆ /DnaG ₃	6,4		25,4		14,5	60%
pl-di ₆ /DnaG ₃	5,8	84%	28,2		17,9	

Tableau 14 : Tableau récapitulatif des χ^2 obtenus à partir des différents modèles testés sur les données SAXS de HpDnaB + HpDnaG^{HBD}, HpDnaB + ATP + HpDnaG^{HBD} et HpDnaB + AMP-PNP+ HpDnaG^{HBD}. Les calculs ont été réalisés sur le serveur web FoXs.

Dans le cas de HpDnaB + HpDnaG^{HBD}, les analyses montrent que les χ^2 obtenus en comparant les courbes théoriques de nos différents modèles en dodécamères et en hexamères avec la courbe expérimentale sont très élevés. Alors que la courbe théorique d'un hexamère en complexe avec trois molécules de HpDnaG^{HBD} superposée aux données expérimentales permettent d'obtenir un χ^2 de 3,1 (Tableau 14). L'analyse MES améliore légèrement le χ^2 qui est égale à 2,3. C'est la combinaison du modèle de HpDnaB en complexe avec HpDnaG^{HBD} pl-di₆/DnaG₃ (86%) et celui de HpDnaB sp-di₁₂ (16%) qui génère la courbe théorique correspondant le mieux à la courbe expérimentale (Figure 88, Tableau 14).

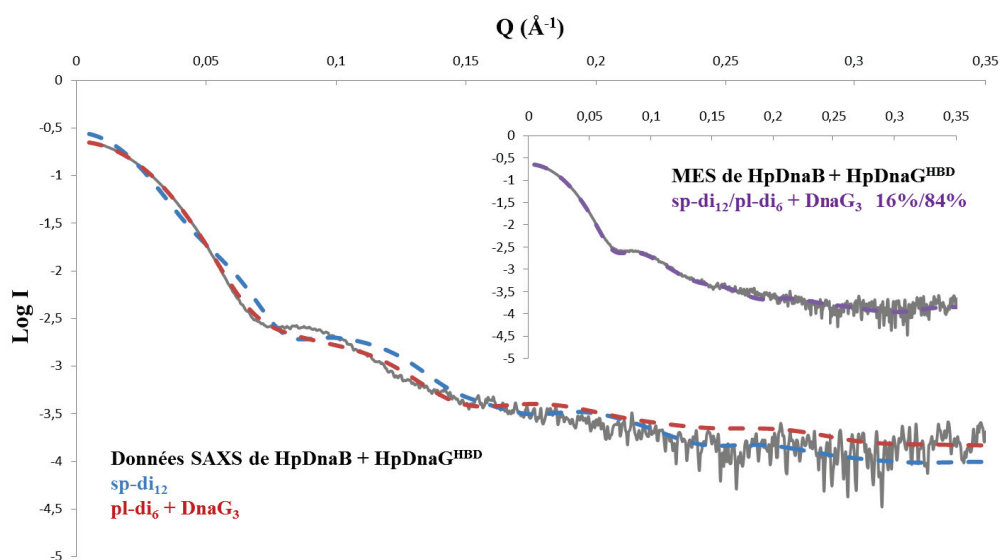


Figure 88 : Comparaison des données SAXS expérimentales de HpDnaB + HpDnaG^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₆ + DnaG₃ (rouge) et la courbe du MES (violet).

La même analyse a été réalisée avec HpDnaB + HpDnaG^{HBD} + ATP. Le modèle qui correspond le mieux aux données expérimentales est le modèle du dodécamère sp-di₁₂ avec un χ^2 égal à 12,9 (Tableau 14). L'analyse en MES n'améliore pas le χ^2 . Ces données sont donc similaires aux données de HpDnaB + ATP (Figure 89, Tableau 14)

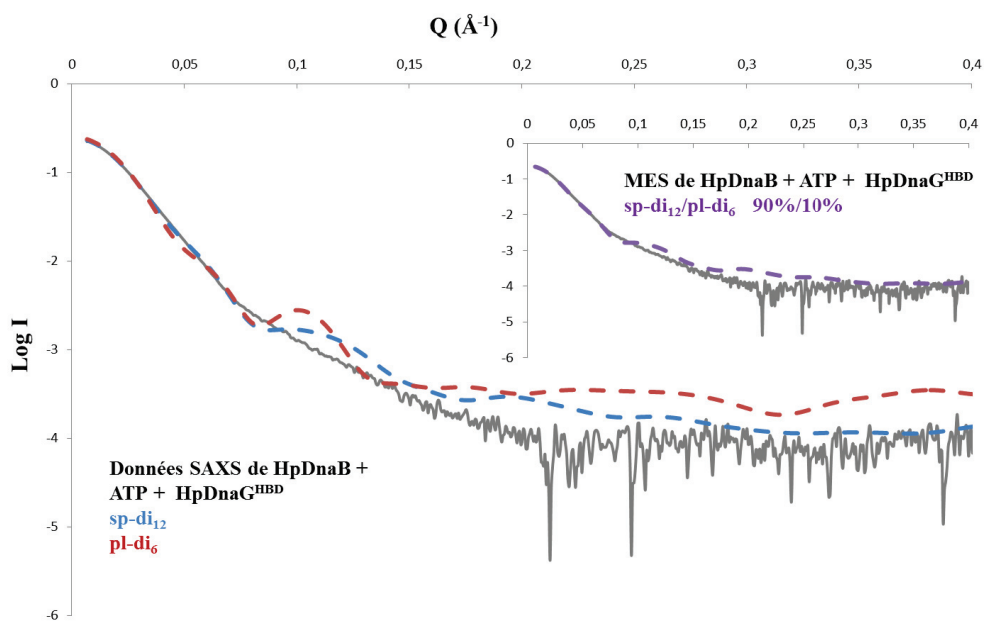


Figure 89 : Comparaison des données SAXS expérimentales de HpDnaB + ATP + HpDnaG^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di₁₂ (bleu), pl-di₆ (rouge) et la courbe du MES (violet).

Enfin dans le cas de HpDnaB + AMP-PNP + HpDnaG^{HBD}, le curve théorique qui correspond le mieux aux données expérimentales est celle du modèle de HpDnaB construit sur DnaB issu du complexe EcDnaB₆•EcDnaC₆ avec un χ^2 de 11,6 (Tableau 14). Il correspond à un hexamère en spirale avec un anneau NTD qui n'est plus dilaté mais resserré en complexe avec trois molécules de HpDnaG^{HBD} (sp-re₆-DnaG₃). Cependant le MES améliore significativement le χ^2 qui passe à 3,5 et indique que la curve théorique correspondant le mieux aux données est générée avec un mélange d'un complexe sp-di₆/DnaG₃ (60%) et d'un dodécamère sp-re₁₂, (40%) (Figure 90, Tableau 14).

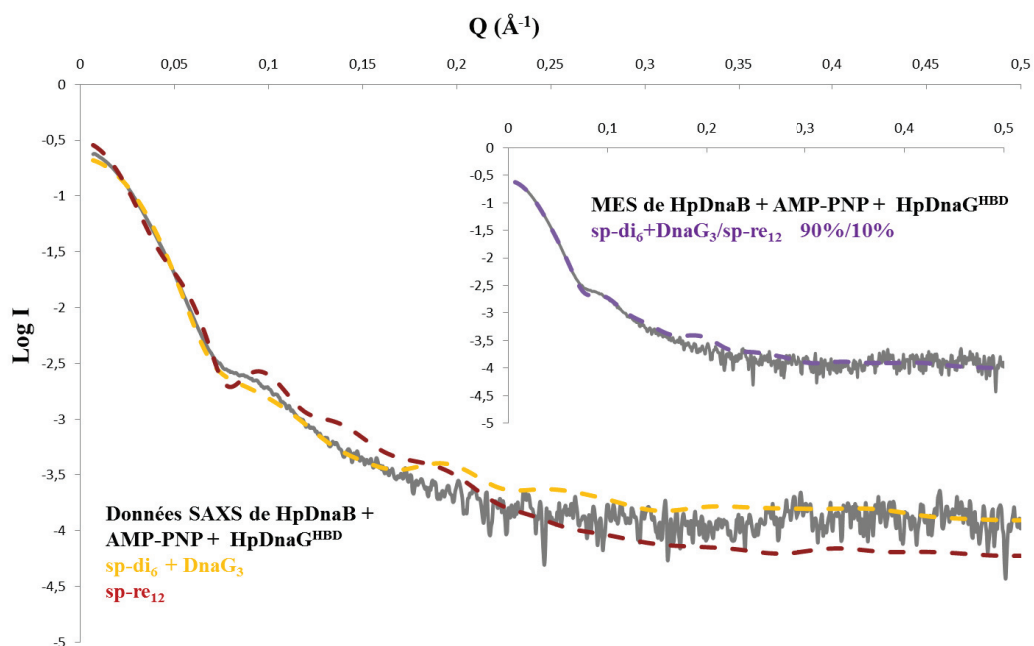


Figure 90 : Comparaison des données SAXS expérimentales de HpDnaB + AMP-PNP + HpDnaG^{HBD} avec les courbes théoriques calculées à partir des modèles sp-di₆ + DnaG₃ (jaune), pl-re₆ (rouge) et la courbe du MES (violet).

Ces résultats montrent donc que HpDnaG^{HBD} seule est capable de se fixer à HpDnaB. Cette interaction permettrait de passer d'un dodécamère sp-di₁₂ à un hexamère pl-di₆ en complexe avec HpDnaG^{HBD}. Il semblerait également que l'ATP joue un rôle dans cette interaction. En effet, HpDnaB en présence de HpDnaG^{HBD} et d'ATP reste un dodécamère majoritairement en sous forme sp-di₁₂. Enfin lorsque l'hydrolyse de l'ATP n'est plus possible, on retrouve une dissociation du dodécamère par HpDnaG^{HBD}.

3.4.2. Etude par SEC-MALS de l'interaction HpDnaB/HpDnaG^{HBD}

Nous avons ensuite étudié la masse de HpDnaB et observé son état d'oligomérisation avant et après interaction avec HpDnaG^{HBD}. La masse de HpDnaB seule a été mesurée en premier, puis HpDnaB et HpDnaG^{HBD} ont été mélangées (HpDnaB+HpDnaG^{HBD}), incubées ensemble 10 min sur la glace et injectées. Enfin un échantillon du complexe issu de la co-expression de HpDnaB et HpDnaG^{HBD} après purification sur colonne de nickel (HpDnaB•HpDnaG^{HBD}) a été injecté et la masse de HpDnaB mesurée.

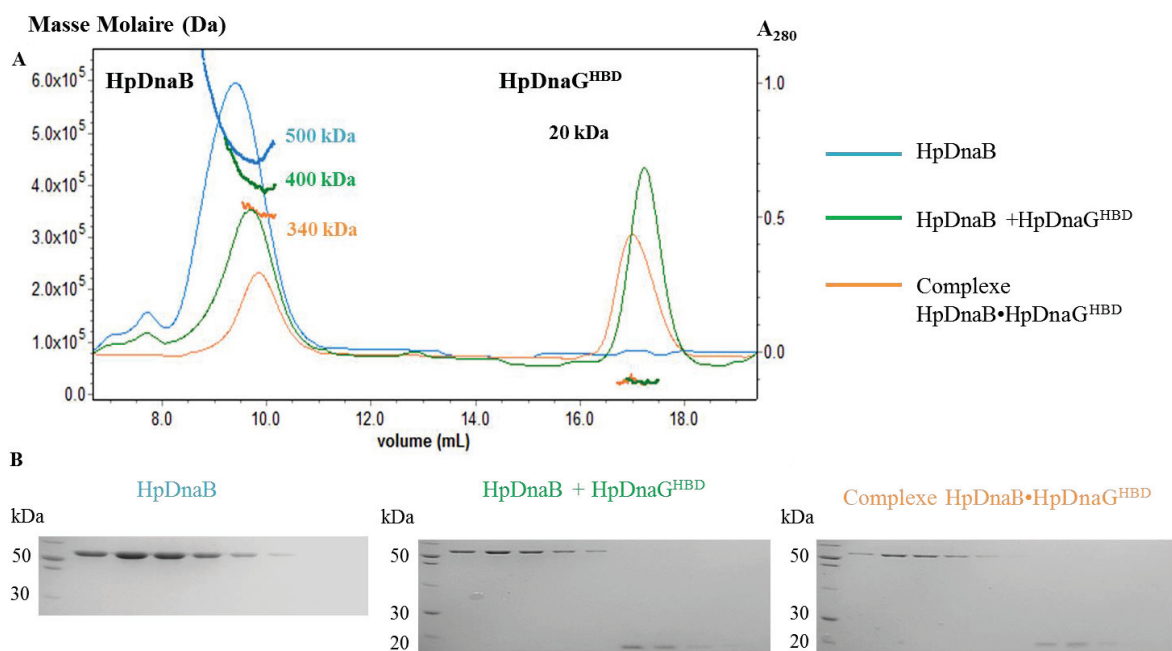


Figure 91 : (A) Mesures de la masse (en kDa) de HpDnaB seule, mélangée ou co-exprimée avec HpDnaG^{HBD} par SEC-MALS (Superdex S200). (B) Analyse des fractions des pics de protéines sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel

Les résultats montrent que la masse de HpDnaB est d'environ 500 kDa et cette masse ne correspond pas à un dodécamère (668 kDa). Cependant, lorsque l'on regarde l'évolution de la masse tout au long du pic, on voit que celle-ci débute à une masse correspondant à un dodécamère (680 kDa) et diminue jusqu'à 500 kDa (Figure 91). Cette variation peut s'expliquer par le fait que la masse donnée par le MALS est une moyenne de l'ensemble des oligomères. La masse serait donc plus faible du fait de la présence d'hexamères en solution et serait la conséquence d'une dissociation partielle du dodécamère et d'un équilibre hexamère-

dodécamère. Cependant d'après les expériences de SAXS, HpDnaB en solution adopte majoritairement une conformation de dodécamère.

Lorsque HpDnaB est mélangée à HpDnaG^{HBD}, on constate que le volume d'éluion augmente par rapport à celui de la protéine seule. On observe également un second pic correspondant à HpDnaG^{HBD}. La masse du pic HpDnaB diminuant à 400 kDa en moyenne (Figure 91), HpDnaG^{HBD} semble donc avoir un effet sur la masse de HpDnaB. L'équilibre dodécamère-hexamère pourrait se déplacer vers l'augmentation des formes hexamériques.

Enfin lorsqu'un échantillon du complexe purifié est injecté sur SEC-MALS, le complexe est complètement séparé par la colonne. La masse mesurée de HpDnaB est de 375 kDa, ce qui est proche de la masse d'un hexamère (334 kDa) et cette masse est stable tout au long du pic de protéine (Figure 91).

Cependant, ces résultats sont en désaccord avec les résultats obtenus en SEC-SAXS. Dans ces derniers, un pic majoritaire contenant un complexe HpDnaB₆•HpDnaG^{HBD}₃ avait été observé. Dans la mesure où les expériences avaient été réalisées avec des colonnes d'exclusion stérique différentes, nous avons donc réalisé de nouvelles expériences de SEC-MALS en utilisant la colonne Shodex utilisée pour le SAXS.

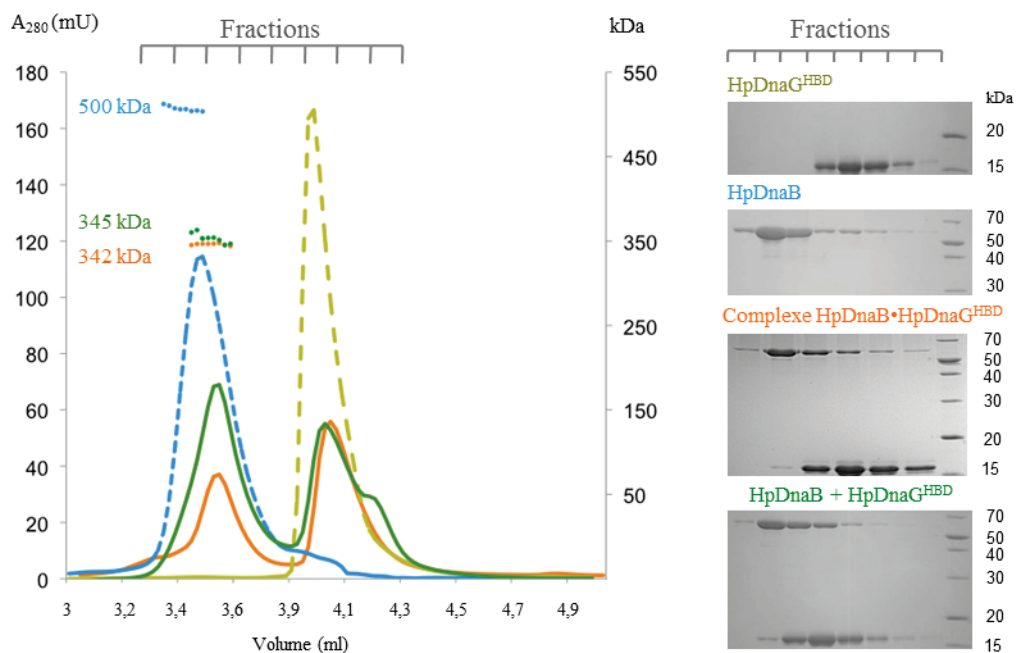


Figure 92 : (A) Mesures SEC-MALS de HpDnaB, mélangée ou co-exprimée avec HpDnaG^{HBD} par SEC-MALS (Shodex KW405-4F) (B) Analyse des fractions des pics de protéines sur gels SDS colorés au bleu de Coomassie. Le marqueur de poids moléculaire est indiqué sur la dernière piste.

Les résultats montrent que le complexe est élué en deux pics (Figure 92). Les gels SDS indiquent que le premier pic contient le complexe HpDnaB•HpDnaG^{HBD} alors que le second pic contient exclusivement HpDnaG^{HBD}. Les calculs de masses montrent que le complexe possède une masse de 345 kDa. Ceci suggère que le premier pic correspond à HpDnaB sous forme d'hexamère liée à une ou plusieurs molécules de HpDnaG^{HBD} (17 kDa). Ces résultats sont donc différents de ceux observés avec la colonne Superdex S200. Ceci suggère que cette dernière impose des contraintes stériques qui déstabilisent le complexe même lorsque celui-ci est issu de la co-expression de HpDnaB et HpDnaG^{HBD}.

3.4.3. Mesure de l'affinité de l'interaction HpDnaB-HpDnaG^{HBD} par titration colorimétrique isotherme (ITC)

Afin de déterminer les paramètres thermodynamique de l'interaction entre HpDnaB et HpDnaG^{HBD}, des expériences de titrations calorimétriques isothermes ont été réalisées. La protéine HpDnaB a été titrée par injection de HpDnaG^{HBD} et la variation de chaleur a été mesurée par le microcalorimètre (ITC200, GE Healthcare). On suit ainsi la variation de chaleur au cours du temps qui augmente pour atteindre un plateau. Les aires des pics sont ensuite analysées avec le logiciel Origin en partant du principe qu'il n'existe qu'un seul site de fixation de HpDnaG^{HBD}. On obtient un graphique représentant la quantité de chaleur par mole d'HpDnaG^{HBD} injectée en fonction du ratio molaire. Cela nous permet de déterminer la constante d'association et le ratio molaire de la réaction. L'expérience a été réalisée avec HpDnaB et 6His-HpDnaB. Dans le cas de 6His-HpDnaB, la variation de chaleur est positive avec une constante de dissociation d'environ 20 μ M et un ratio molaire de 0,3. Dans le cas de HpDnaB, la variation de chaleur est négative, la constante de dissociation est d'environ 60 μ M et le ratio molaire est d'environ 0,6 (Figure 93). Néanmoins dans les deux expériences, la variation d'enthalpie libre est négative ce qui signifie que la réaction est spontanée.

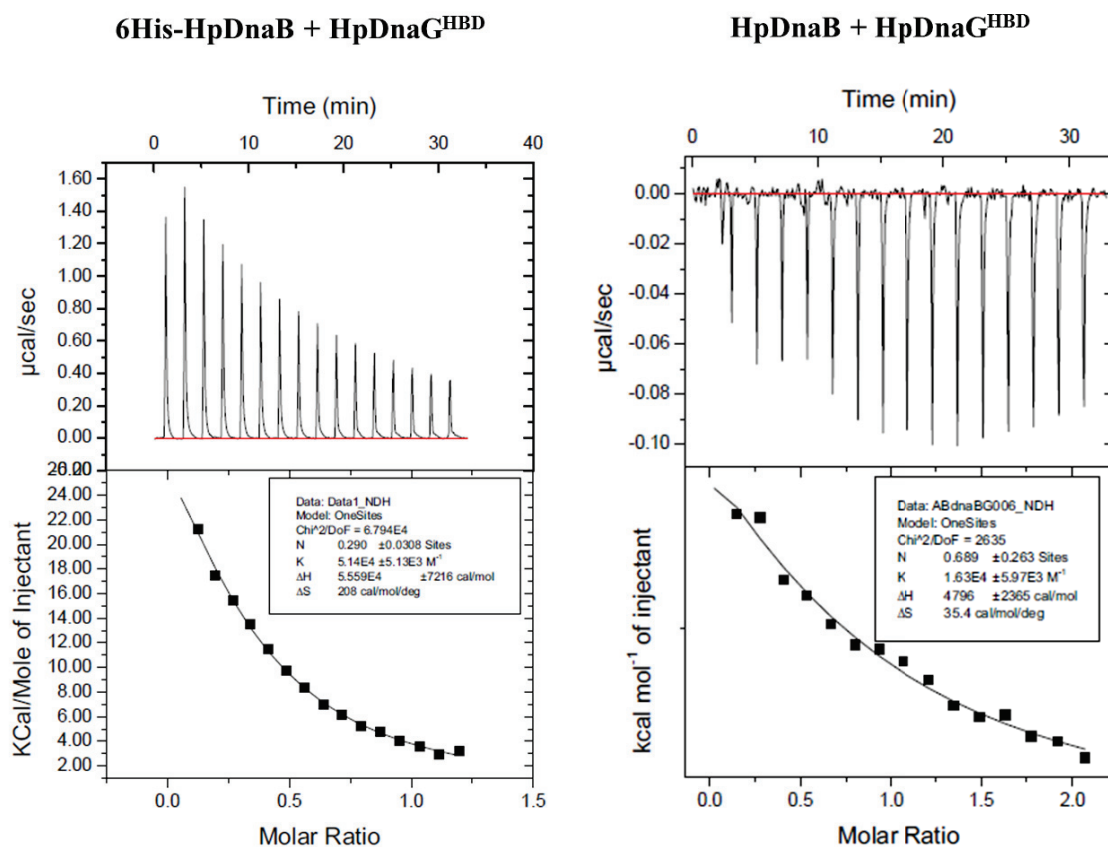


Figure 93 : Titration calorimétrique isotherme de 6His-HpDnaB (à gauche) et de HpDnaB (à droite), titrées par HpDnaG^{HBD}.

Nous constatons que l'affinité entre les deux protéines est assez faible de l'ordre de 10 μM . Cette affinité correspond à ce qui est observé avec les protéines équivalentes d'*E. coli* (Oakley *et al.*, 2005 ; Mitkova *et al.*, 2003). Dans le cas de 6His-HpDnaB le rapport molaire est de 0.3 correspondant à 1 HpDnaG^{HBD} pour 3 HpDnaB. Or, la structure du complexe BstDnaB•BstDnaG^{HBD} montre que le rapport pourrait atteindre 1 HpDnaG^{HBD} pour 2 HpDnaB. Les résultats indiquent que tous les sites de HpDnaB ne sont pas saturés par HpDnaG^{HBD}. Dans le cas de HpDnaB, la stœchiométrie observée en ITC est de 0,6 soit proche du rapport 1 HpDnaG^{HBD} pour 2 HpDnaB. Dans ce cas, il semble que tous les sites de liaisons soient saturés. Dans le cas de HpDnaB, l'expérience d'ITC suggère également que l'interaction n'est pas simple car on observe plusieurs évènements lors de la titration. La variation d'enthalpie ne correspondant pas à celle de l'expérience avec 6His-HpDnaB, il est vraisemblable que l'on mesure également la dissociation du dodécamère en deux hexamères en même temps que la liaison de HpDnaG^{HBD} à HpDnaB. Cette différence suggère également que 6His-HpDnaB n'est pas dissociée en présence de HpDnaG^{HBD}.

Pour vérifier cette hypothèse, nous avons étudié le mélange 6His-HpDnaB + HpDnaG^{HBD} par SEC-MALS. Les résultats montrent que les deux protéines sont séparées et que la masse de l'hélicase reste inchangée (678 kDa) (Figure 94). Ceci confirme que l'interaction avec HpDnaG ne modifie pas la structure de 6His-HpDnaB.

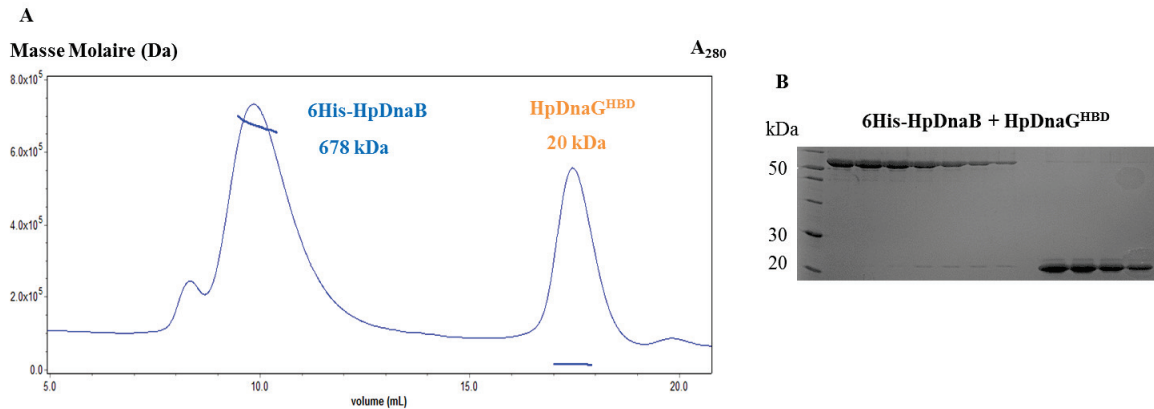


Figure 94 : (A) Mesures de la masse (en kDa) de 6His-HpDnaB mélangée avec HpDnaG^{HBD} par SEC-MALS (Superdex S200). (B) Analyse des fractions des pics de protéines sur gel SDS coloré au bleu de Coomassie. Le marqueur de poids moléculaire (en kDa) est indiqué sur la première piste du gel.

3.4.4. Mesure de l'activité hélicase de HpDnaB et du complexe HpDnaB•HpDnaG^{HBD} et de leurs affinités pour l'ADNsb

Nos expériences de SAXS ont montré qu'en présence d'ATP, HpDnaB est un dodécamère avec un ADNsb. Le dodécamère étant composé de deux hexamères qui se font face, un seul peut être positionné dans le bon sens pour dérouler l'ADNdb puisque la protéine possède une polarité 5'-3'. Nous avons montré précédemment en SEC-MALS et en SEC-SAXS que l'interaction avec HpDnaG^{HBD} séparait les deux hexamères. Pour étudier l'effet d'HpDnaG^{HBD} sur l'activité hélicase de HpDnaB, notre collaboratrice Joanna Timmins a effectué un test hélicase en hybridant deux oligonucléotides dont l'un porte une sonde fluorescente et l'autre une molécule qui supprime la fluorescence de la sonde lorsque les deux oligonucléotides sont hybridés. Ainsi si la protéine a une activité hélicase, les deux oligonucléotides sont séparés et la fluorescence augmente avec le temps.

Dans cette expérience, HpDnaB ne possède qu'une faible activité hélicase avec seulement 10% d'ADNdb séparé (Figure 95A). Le complexe HpDnaB•HpDnaG^{HBD} possède une activité hélicase plus élevée avec 60% d'ADNdb séparé (Figure 95A). Pour mesurer l'affinité de HpDnaB et du complexe HpDnaB•HpDnaG^{HBD} pour l'ADNsb, des expériences d'anisotropie de fluorescence ont été réalisées. Les résultats indiquent qu'en présence d'ATP, HpDnaG^{HBD}

augmente l'affinité de HpDnaB pour l'ADNsb avec un K_d passant de 21,7 nM pour HpDnaB seule à un K_d de 8,3 nM pour le complexe HpDnaB•HpDnaG^{HBD}. Au contraire, en présence d'AMP-PNP, HpDnaG^{HBD} n'a que peu d'effet sur l'affinité de HpDnaB pour l'ADNsb. En effet, le K_d est de 1,8 nM pour HpDnaB seule alors qu'avec HpDnaB•HpDnaG^{HBD}, le K_d est de 1 nM. Or, les résultats de SEC-SAXS et de SEC-MALS (voir partie 3.3.1.2 et 3.3.2) montrent que HpDnaB est sous forme d'hexamère en présence d'AMP-PNP et d'ADNsb. Ces résultats suggèrent donc que HpDnaG^{HBD} n'a un effet sur HpDnaB que si la protéine est sous forme de dodécamère (Figure 95B).

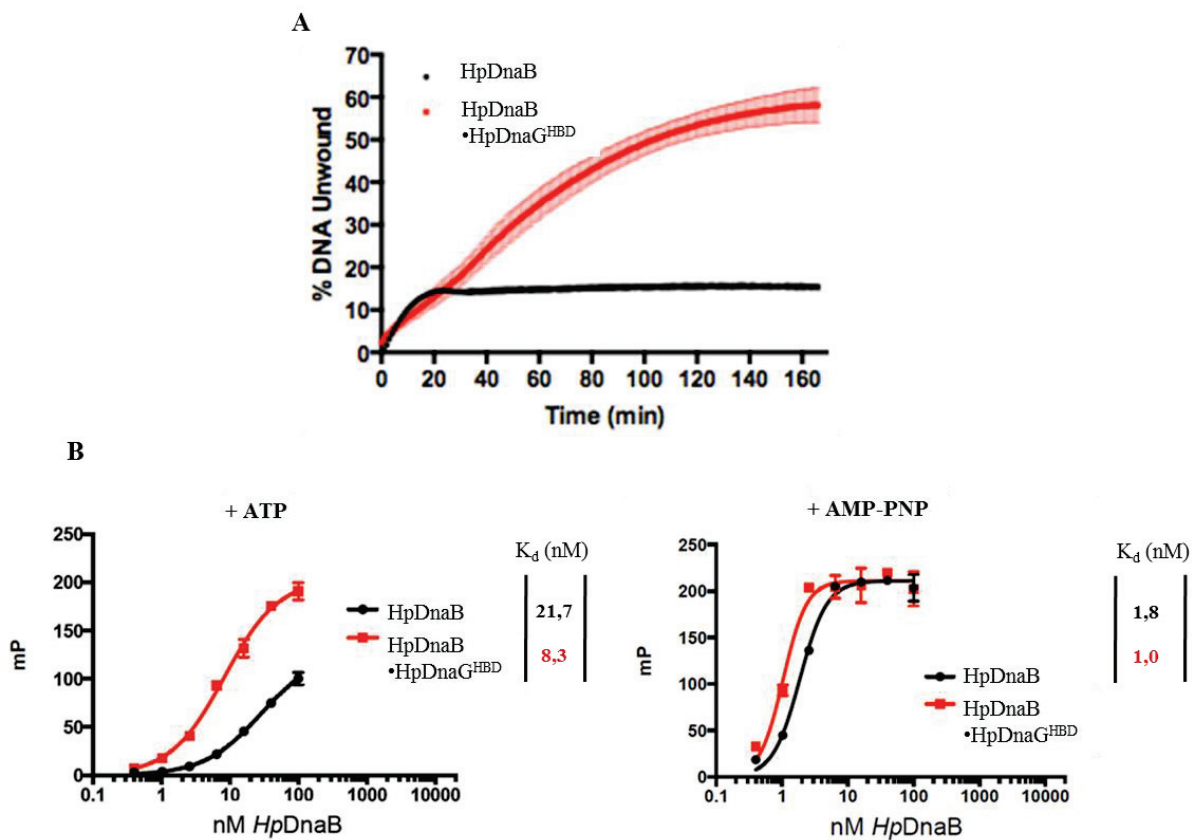


Figure 95 : (A) Mesures des activités hélicase représentant le pourcentage d'ADNdb séparé en fonction du temps pour HpDnaB seule (noir) ou du complexe HpDnaB•HpDnaG^{HBD} (rouge). (B) Expériences d'anisotropie de fluorescence pour la détermination de l'affinité de HpDnaB seule (noir) ou du complexe HpDnaB•HpDnaG^{HBD} (rouge) pour l'ADNsb en présence d'ATP (à gauche) ou d'AMP-PNP (à droite).

3.4.5. Mesure de l'activité ATPase de HpDnaB et des complexes du primosome

6His-HpDnaB possède une activité ATPase de 0.069 μM d'ATP/sec (soit une vitesse de 0.863 s^{-1}) et 0.279 μM d'ATP/sec (soit 3.48 s^{-1}) en présence d'un ADNsb de 50 nucléotides (voir Article 1). Dans l'hypothèse où HpDnaG^{HBD} sépare le double anneau, l'activité ATPase de la protéine native seule a été testée, en présence d'un ADNsb de 20 nucléotides (oligodT20), en présence de HpDnaG^{HBD} et en présence des deux. HpDnaB est utilisée à une concentration de 500 nM pour pouvoir mesurer la vitesse initiale de la réaction car à plus faible concentration, l'activité est insuffisante pour être détectée par le spectrophotomètre. HpDnaB seule montre une activité ATPase de 0,045 μM d'ATP/sec (soit une vitesse de 0.09 s^{-1}). Le taux d'erreur est important du fait de la limite de détection du spectrophotomètre. Cette activité est inférieure à celle observée pour 6His-HpDnaB. Cependant, lorsque l'on rajoute l'oligodT20, l'activité ATPase est de 0,132 μM d'ATP/sec (soit 0.264 s^{-1}) et elle est de 0,158 μM d'ATP/sec (soit 0.316 s^{-1}) quand HpDnaB est incubée avec HpDnaG^{HBD}. Enfin lorsque l'ADN et HpDnaG^{HBD} sont ajoutés à HpDnaB, l'activité ATPase de l'hélicase est de 0,349 μM d'ATP/sec (soit 0.698 s^{-1}). En présence d'ADN ou d'HpDnaG^{HBD}, l'activité ATPase de HpDnaB est multipliée par trois et en présence des deux, l'activité ATPase est multipliée par sept. L'action de l'ADN et celle de HpDnaG^{HBD} semblent donc être complémentaire. (Figure 96).

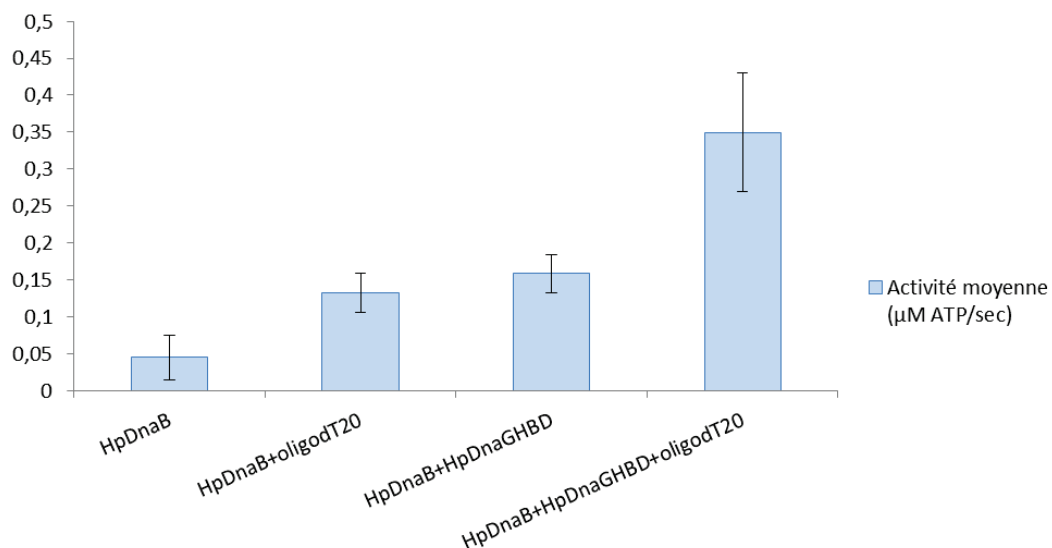


Figure 96 : Mesures de l'activité ATPase de la protéine HpDnaB seule et/ou en présence d'un oligonucléotide-dT20 et/ou de HpDnaG^{HBD}.

3.4.6. Etude par microscopie électronique de HpDnaB après interaction avec HpDnaG^{HBD}

Le complexe HpDnaB•HpDnaG^{HBD} a également été observé en ME à coloration négative. Normalement il est préférable de passer l'échantillon sur une colonne d'exclusion stérique pour faire une analyse en ME. Sachant qu'une telle colonne sépare le complexe, nous avons tenté de faire des grilles sans passer par l'étape de gel filtration. Nous avons également injecté le complexe sur colonne de gel filtration et récupéré uniquement HpDnaB afin d'observer son état d'oligomérisation après interaction avec HpDnaG^{HBD}. Les méthodes utilisées pour la préparation des grilles, la collecte et le traitement des images sont les mêmes que précédemment. Dans le cas du complexe, les images et les classes sont très bruitées alors que pour HpDnaB après interaction avec HpDnaG^{HBD}, les classes sont meilleures. Contrairement à HpDnaB native, nous avons pu obtenir plus de vues de côté. Une reconstruction 3D a donc pu être initiée. Pour cela, le logiciel IMAGIC a été utilisé pour imposer la symétrie et calculer les angles d'Euler nécessaires pour la reconstruction 3D. Dans les deux cas, une ébauche d'enveloppe a pu être obtenue (Figure 97).

Les enveloppes de HpDnaB•HpDnaG^{HBD} après ou avant purification sur colonne d'exclusion stérique sont des hexamères et les deux anneaux NTD et CTD sont identifiables. Dans le cas de HpDnaB•HpDnaG^{HBD} après Superdex S200, les symétries C3 de l'anneau NTD et C6 de l'anneau CTD de HpDnaB se distinguent. Cependant en retirant l'étape de purification sur la colonne d'exclusion stérique, la symétrie des deux anneaux semble identique. En effet, les classes étant très bruitées dans ce cas, l'enveloppe qui en est issue est de moins bonne qualité. Il ne faut cependant pas exclure la possibilité que trois molécules de HpDnaG^{HBD} puissent être fixées pouvant créer par projection avec l'anneau NTD une symétrie C6.

Ce travail nécessite donc d'être poursuivi et approfondi en collectant plus d'images dans toutes les orientations possibles et en approfondissant le traitement des données afin d'améliorer la résolution. Des enveloppes à plus haute résolution nous permettrait de pouvoir les comparer et de placer à l'intérieur de celles-ci les différents domaines de HpDnaB après interaction et de construire un modèle du complexe. Il serait également intéressant d'utiliser la colonne Shodex pour purifier le complexe avant de collecter les images du complexe car nous savons maintenant que cette colonne ne sépare pas le complexe contrairement à la Superdex S200. Ceci permettrait d'éliminer les molécules de HpDnaG^{HBD} en excès et ainsi de réduire le bruit de fond.

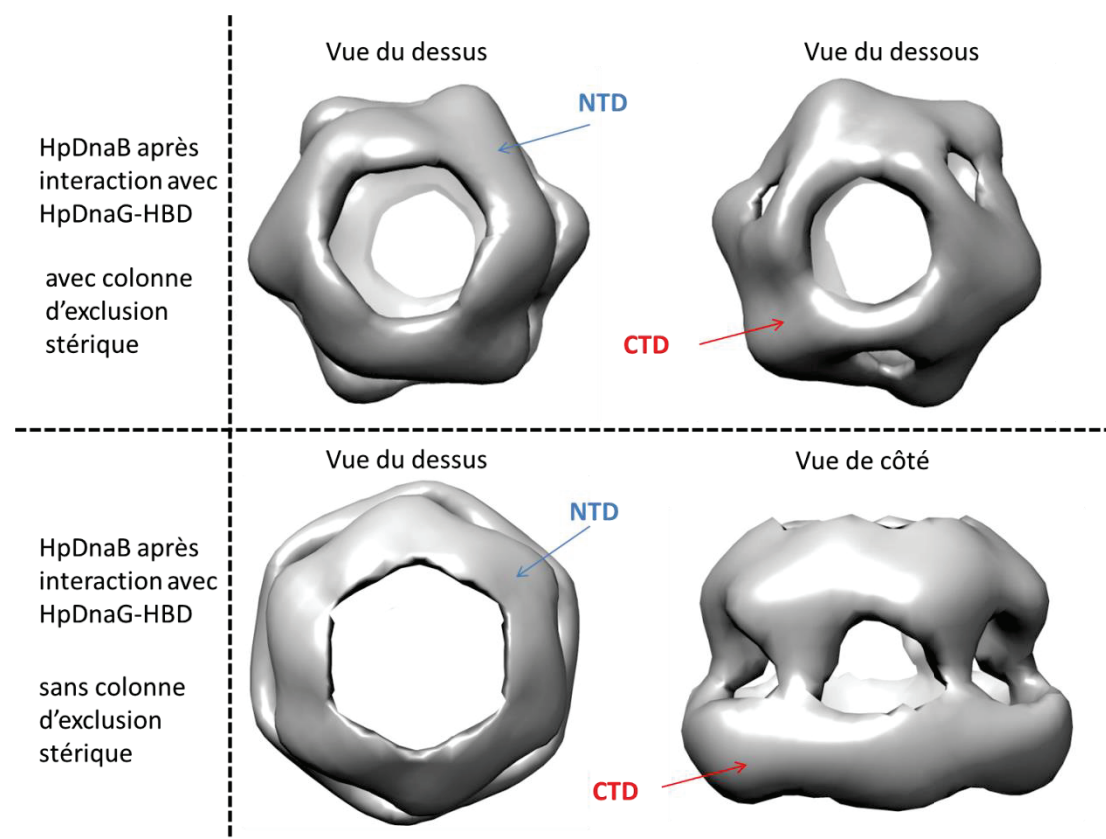


Figure 97 : Reconstruction en 3D de HpDnaB après interaction avec HpDnaG^{HBD} avec ou sans purification sur colonne d'exclusion stérique.

3.5. Etude fonctionnelle du chargement de HpDnaB sur l'origine de réplication

3.5.1. Expérience d'ouverture de l'origine de réplication

Afin de comprendre l'ouverture de l'origine de réplication de *H. pylori*, et le chargement de l'hélicase, un essai combinant un plasmide contenant l'origine de réplication et la nucléase P1 a été réalisé. L'origine de réplication est composée des deux régions *oriC1* et *oriC2* séparées par le gène *dnaA* exactement comme dans le génome de *H. pylori* (Donczew *et al.*, 2012). Le plasmide a été incubé avec HpDnaA seule ou avec les autres protéines. Il a ensuite été mis en présence de la nucléase P1 qui hydrolyse uniquement l'ADNsb. Après purification, la plasmide est coupé une deuxième fois avec l'enzyme de restriction *EcoRI* qui excise la région contenant l'origine de réplication et permet de déterminer s'il y a eu ouverture. La mise au point de cette expérience a été réalisée dans le laboratoire d'Anna Zawilak-Pawlik (Institute of Immunology and Experimental Therapy) à Wrocław en Pologne.

Plusieurs bandes sont observées à l'issue de l'expérience sans protéine (Figure 98). La bande à environ 4507 pb (bande 1) correspond à la taille du vecteur digéré par *EcoRI* seulement. La bande aux alentours de 2838 pb (bande 2) est le résultat de la coupure par l'enzyme de restriction *EcoRI* et la bande à 700 pb (bande 3) signifie que l'origine de réplication pBR322 du plasmide a été coupée par la nucléase P1. Enfin les deux bandes à 2310 pb et 1604 pb (bandes 4) montrent que la nucléase P1 a coupé au niveau de l'origine de réplication *oriC2* et que l'enzyme de restriction *EcoRI* a également coupé le plasmide. Cet essai permet de montrer que d'une part l'origine de réplication du plasmide s'ouvre toute seule naturellement et donc la nucléase P1 a pu couper à cet endroit. D'autre part, *oriC2* s'ouvre également toute seule. Lorsque HpDnaA est rajoutée avec des concentrations croissantes, les bandes aux alentours de 2310 pb et 1604 pb augmentent en intensité. Cela signifie que HpDnaA s'est fixée sur l'origine de réplication de *H. pylori* et qu'il y a eu ouverture de celle-ci. A contrario, plus la concentration de HpDnaA augmente, moins la bande correspondant au vecteur non digéré est intense. La protéine HpDnaA purifiée au laboratoire est donc bien active et a un effet sur l'ADN à partir de 200 ng.

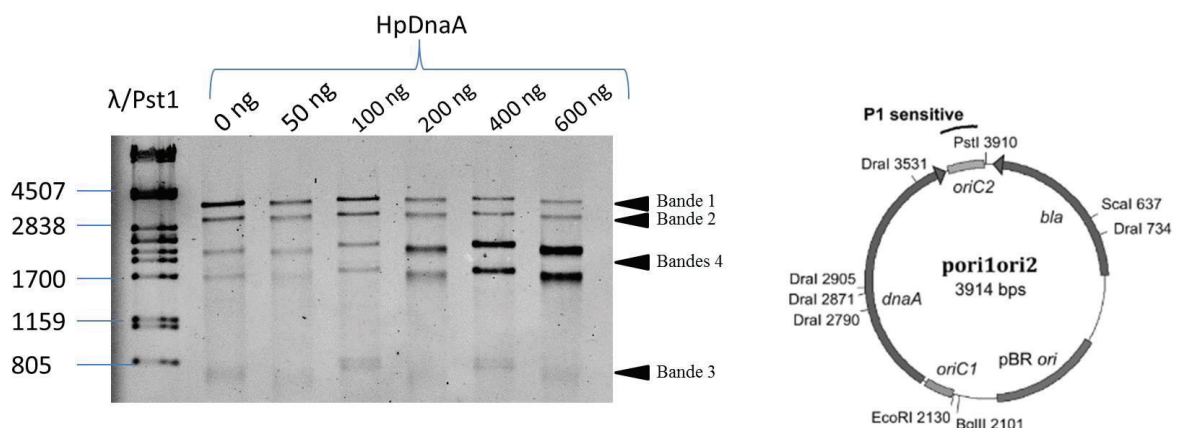


Figure 98 : Expérience d'ouverture de *HporiC* insérée dans le plasmide *porIori2* (à droite). Le résultat est visualisé sur un gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de taille est indiqué sur la première piste du gel.

Les protéines HpDnaB, SSB et HpDnaG^{HBD} ont été testées en combinaison avec HpDnaA dans le but d'observer un chargement de HpDnaB sur l'ADNsb. En effet, si HpDnaB se charge et qu'elle exerce son activité hélicase, alors la nucléase P1 devrait hydrolyser un fragment plus grand d'ADNsb et les bandes observées sur gel d'agarose seront donc plus petites après la réaction.

La même expérience que précédemment a donc été réitérée avec un échantillon sans protéine, un échantillon avec 150 ng d'HpDnaA et trois échantillons avec 150 ng de HpDnaA avec une concentration croissante de HpDnaB (100ng/200ng/400ng) (Figure 99). L'échantillon contrôle contenant uniquement HpDnaA montre que celle-ci a bien le même effet que précédemment (bandes 1 à 4). Si on rajoute HpDnaB, on constate que les fragments d'ADN à l'issue de la réaction sont toujours de même taille mais leur intensité est plus faible en comparaison avec l'échantillon contenant seulement HpDnaA. HpDnaB s'est donc peut être chargée sur l'ADNs_b empêchant ainsi l'action de la nucléase P1 mais la protéine étant en double anneau sur chaque ADNs_b de l'origine, son activité hélicase n'est pas assez forte. Cependant, si la nucléase P1 ne peut plus couper, alors le plasmide ne devrait être coupé que par *EcoR1* or nous n'observons pas d'augmentation d'intensité de la bande correspondant au plasmide linéaire (bande 1 à fragment à 4507 pb).

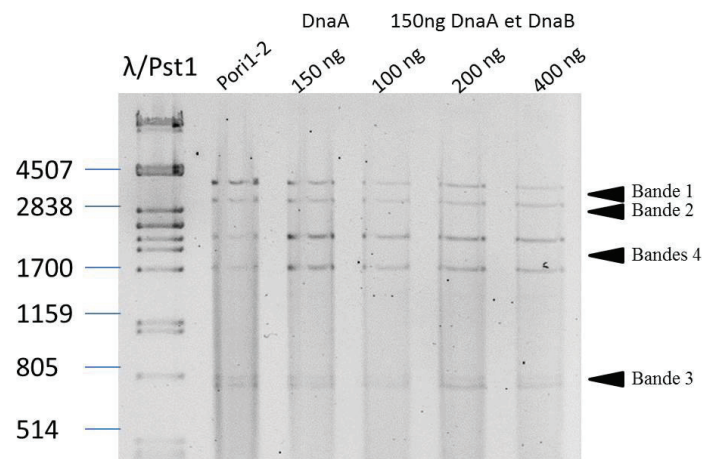


Figure 99 : Expérience d'ouverture de *HporiC* insérée dans le plasmide *pori1ori2* en présence de HpDnaA et HpDnaB. Le résultat est visualisé sur un gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de taille est indiqué sur la première piste du gel.

L'expérience a donc été réitérée en testant également l'action de HpDnaG^{HBD}. L'objectif était de vérifier si HpDnaG^{HBD} sépare le double anneau pour que l'hexamère de HpDnaB puisse étendre la région simple brin à partir de l'origine. Pour cela, les protéines HpDnaA, HpSSB, HpDnaB et HpDnaG^{HBD} ont été utilisées. Nous avons comparé le plasmide seul avec le plasmide en présence de HpDnaA, de HpDnaB ou de HpDnaG^{HBD} puis testé l'action des trois protéines en même temps avec ou sans HpSSB (Figure 100). Les résultats montrent que HpDnaA ouvre l'origine de répllication (bandes 4) mais que HpDnaB et HpDnaG^{HBD} seules n'ont aucun effet. Cependant, le rajout de HpDnaB à HpDnaA n'a pas d'effet sur l'ouverture

d'*oriC2*. Enfin en présence des trois protéines en même temps et avec une concentration croissante de HpDnaG^{HBD}, il n'y a aucune différence par rapport à HpDnaA seule ou HpDnaA+HpDnaB. HpDnaG^{HBD} n'a donc pas eu l'effet escompté. Enfin, le fait de rajouter HpSSB à la réaction ne permet pas non plus à l'hélicase de séparer l'ADNdb.

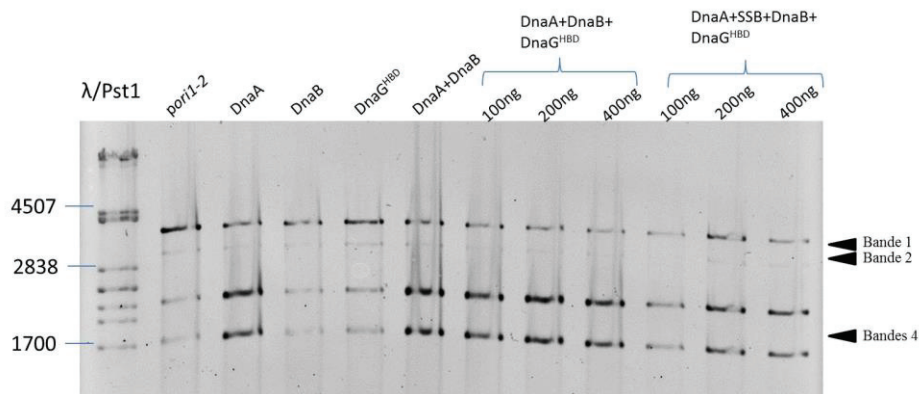


Figure 100 : Essai d'ouverture de *HporiC* insérée dans le plasmide *pori1ori2* en présence de HpDnaA, HpDnaB et HpDnaG^{HBD}. Le résultat est visualisé sur un gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de taille est indiqué sur la première piste du gel.

Ces expériences ont donc permis d'observer l'effet de HpDnaA sur un plasmide contenant l'origine de répllication. Néanmoins, nous n'avons pas pu observer le chargement de HpDnaB ni l'effet de HpDnaG^{HBD} sur l'hélicase, même en présence de plusieurs protéines intervenant dans l'initiation de la répllication. Cette expérience n'est peut-être pas adaptée pour observer la progression de l'hélicase le long de l'ADN. Nous avons donc décidé de mettre en place une expérience qui a permis, dans d'autres travaux d'observer le chargement d'une hélicase sur un plasmide et la formation du primosome.

3.5.2. Etude du chargement de l'hélicase

Chez *E. coli*, les protéines DnaA, DnaB, DnaC et HU sont requises pour la formation du primosome sur l'ADN (Baker *et al.*, 1986). *In vitro*, l'addition de la gyrase et de SSB au primosome provoque un changement de conformation de l'ADN. Il passe d'une forme super enroulée (F1) à une forme circulaire fermée et très déroulée (F1*) qui se distingue par électrophorèse sur gel d'agarose (Baker *et al.*, 1986). Pour réaliser cette expérience, nous avons utilisé le protocole de l'essai réalisé dans une étude antérieure (Konieczny et Helinski, 1997 ; Jiang *et al.*, 2005) (Figure 101).

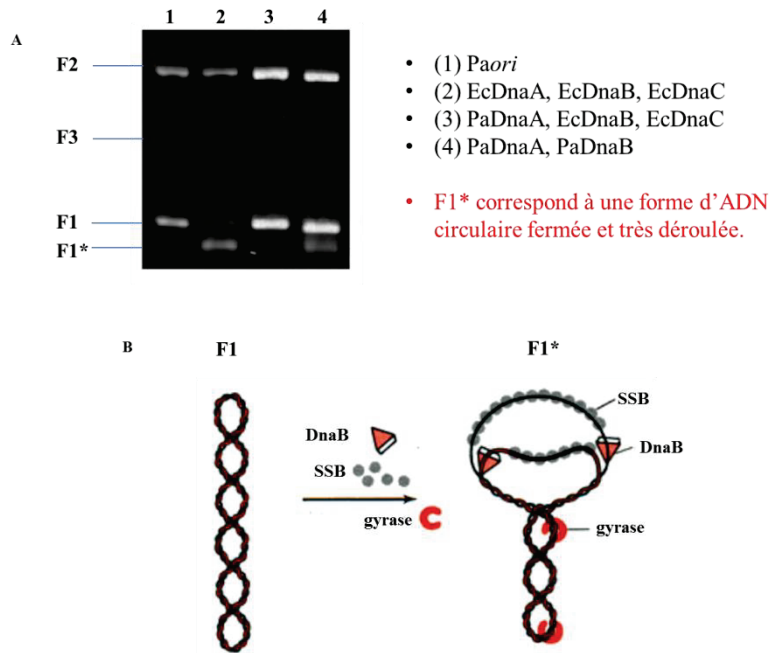


Figure 101 : (A) Expérience de chargement de l'hélicase DnaB de *Pseudomonas aeruginosa* (PaDnaB) (Jiang *et al.*, 2005). (B) Schéma représentant le passage de la forme F1 à F1* par l'action de la gyrase et de SSB une fois le primosome formé (par souci de clarté seule DnaB a été représentée) (Kornberg et Baker, 1992).

Les premiers essais effectués ne montrent pas d'apparition de la forme F1* en présence de toutes les protéines nécessaires à la formation du primosome chez *H. pylori* (HpDnaA, HpDnaB, HobA, HpSSB, HU et la gyrase). Le fragment correspondant à la forme superenroulée (F1) n'est plus visible et les concentrations en forme « nicked » (F2) et linéaire (F3) du plasmide augmentent. Les résultats sont les mêmes en utilisant 6His-HpDnaB au lieu de HpDnaB. Il est étrange que la forme F3 soit présente puisque celle-ci correspond à la forme linéaire du plasmide or aucune des protéines utilisées n'a la capacité de couper l'ADN (Figure 102).

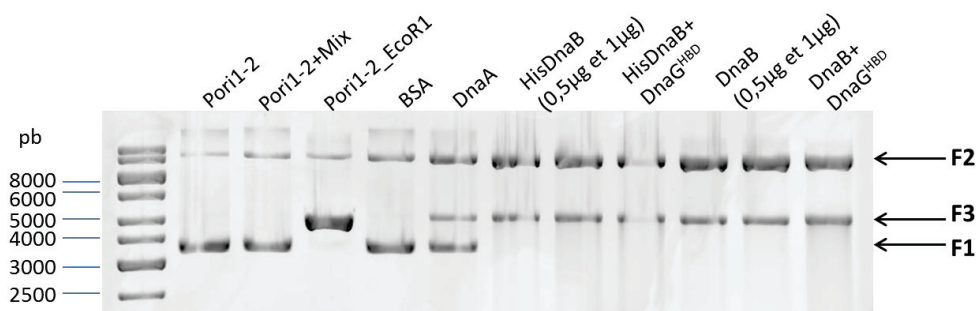


Figure 102 : Expérience de chargement de l'hélicase 6His-HpDnaB et HpDnaB en utilisant toutes les protéines nécessaires à la formation du primosome. Le résultat est visualisé sur un gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de poids moléculaire est indiqué sur la première piste du gel.

Pour identifier la ou les protéines responsables de l'apparition des formes F2 et F3, la même expérience a été réalisée en utilisant différentes combinaisons de protéines. Les résultats montrent qu'avec la protéine HpDnaA seule, les formes F2 et F3 commencent à apparaître. Lorsque 6His-HpDnaB est ajoutée à HpDnaA, la concentration de la forme F2 augmente. Enfin dès que la protéine HobA est ajoutée, la forme superenroulée F1 disparaît complètement au profit de la forme « nicked » F2 et linéaire F3 (Figure 103).

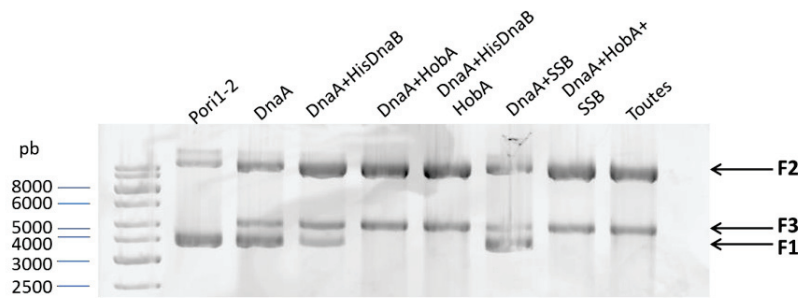


Figure 103 : Expérience de chargement de l'hélicase 6His-HpDnaB ou HpDnaB en testant différentes combinaisons de protéines. Les résultats sont visualisés sur gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de poids moléculaire est indiqué sur la première piste du gel.

La même expérience a été ensuite renouvelée sans HobA pour déterminer l'effet de 6His-HpDnaB et de HpDnaB sur les plasmides *pori1-2*, *pori2* et *pori1*. Les résultats montrent que 6His-HpDnaB semble avoir le même effet que HobA. Le plasmide naturellement sous une forme superenroulée (F1) passe à une conformation majoritairement « nicked » (F2) et linéaire (F3). En présence de HpDnaB, l'effet est moindre et similaire à ce qui a été observé précédemment avec HpDnaA seule (Figure 104).

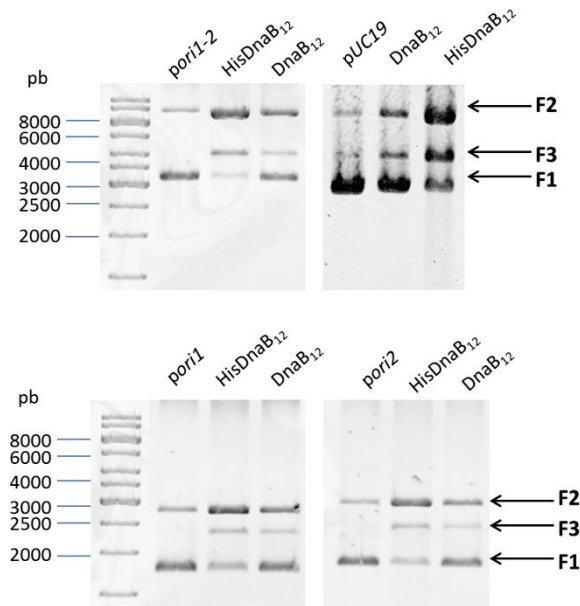


Figure 104 : Expérience de chargement de l'hélicase 6His-HpDnaB et HpDnaB sans la protéine HobA. Le résultat est visualisé sur gel d'agarose 1% coloré au bromure d'éthidium. Le marqueur de poids moléculaire est indiqué sur la première piste du gel.

En conclusion, nos résultats préliminaires n'ont pas permis de mettre en évidence l'apparition de la forme F1*. Dès que le plasmide est incubé avec HpDnaA et HpDnaB, celui-ci change de conformation passant d'une forme superenroulée (F1) à une forme majoritairement « nicked » (F2). Il est donc possible que le primosome se forme sur l'origine de réplication. En effet, nous savons que HpDnaA est active grâce à nos essais d'ouverture. En s'oligomérisant sur l'origine, celle-ci provoque son ouverture, réduisant ainsi le nombre de supertours de l'ADN. Le chargement de HpDnaB et dans une plus grande mesure l'action de HobA changeraient alors complètement la conformation du plasmide passant d'une forme superenroulée à « nicked ». Les études antérieures ayant montré que la forme F1* est le résultat de l'addition *in vitro* de SSB et de la gyrase sur la conformation de l'ADN une fois le primosome formé. Il est donc possible que la gyrase et/ou la protéine SSB ne soient pas actives dans nos conditions d'expérimentation. De ce fait, les supertours de l'ADN nécessaires au maintien de la forme F1 et à l'apparition de la forme F1* ne seraient plus présents. L'étude du chargement de l'hélicase HpDnaB doit donc être poursuivie en optimisant les conditions expérimentales.

4. Discussion

Lors de la réplication chromosomique bactérienne, la séparation de l'ADNdb nécessaire à la progression des deux fourches de réplication repose sur l'hélicase répllicative DnaB. Notre travail sur l'hélicase bactérienne de *H. pylori* (HpDnaB) avait établi que celle-ci forme des dodécamères, contrairement à celles caractérisées jusqu'alors. Le modèle issu de l'enveloppe obtenue par ME suggérait que les anneaux NTD s'organisaient selon une symétrie C6. Nous avons donc supposé que le dodécamère de HpDnaB était chargé sur l'origine de réplication, ce qui permettrait d'orienter deux hexamères dans des directions opposées en l'absence de facteur de chargement. Dans ce modèle, l'interaction avec la primase HpDnaG pourrait séparer le double anneau une fois l'hélicase chargée sur l'origine de réplication. Cette séparation induirait une transition des anneaux NTD de la conformation C6 à une conformation C3 ce qui permettrait la fermeture de l'anneau NTD et le passage de l'ADNsb à l'intérieur de l'hélicase.

Durant cette thèse, notre objectif était d'améliorer la résolution du modèle de HpDnaB en résolvant sa structure cristallographique par diffraction des rayons X et d'étudier ses interactions avec l'ADN et avec la primase HpDnaG, étapes clés de la formation du primosome.

- ***Structure cristallographique du double anneau***

La structure du dodécamère de HpDnaB à 6,7 Å démontre clairement que la protéine est composée de deux hexamères interagissant par leurs domaines NTD. La structure de l'hexamère indique que l'anneau CTD adopte la même conformation qu'AaDnaB-ADP bien qu'aucune molécule d'ADP ne soit visible. Ceci est vraisemblablement dû à la faible résolution. L'anneau NTD n'adopte cependant pas la même conformation que celui d'AaDnaB-ADP. En effet, la structure révèle que les anneaux NTD s'organisent selon une conformation dilatée, similaire à la forme apo de BstDnaB, ce qui diffère de nos observations en ME. HpDnaB adopte donc une conformation intermédiaire avec un anneau NTD en conformation dilatée et un anneau CTD orienté pour lier l'ADP. Cette organisation planaire semble permettre une flexibilité des anneaux CTD. Ces derniers pourraient alors changer de conformation en fonction de l'hydrolyse de l'ATP et/ou de la liaison à l'ADN. HpDnaB conserve une flexibilité remarquable dans le cristal et en solution, suggérant que la protéine peut accommoder des réarrangements structuraux importants tout en restant dodécamérique.

Un autre exemple du dynamisme du domaine CTD réside dans la structure de l'insertion HPI. La structure de HpDnaB^{CTD} a montré que l'insertion est organisée en un motif hélice-tour-hélice (HTH). Dans la structure du dodécamère, les HPI forment également un collier au-dessus de l'anneau CTD ; cependant l'une des hélices de chaque HPI est désordonnée. Il semblerait donc que les HPI puissent passer d'une conformation ordonnée à désordonnée qui pourrait être dépendante de l'état d'hydrolyse de l'ATP et/ou de la liaison à l'ADN. De manière analogue, la structure de MtMCM obtenue en présence d'ADN montre que les motifs HTH présents dans les domaines CTD de la protéine forment un collier au-dessus de l'anneau CTD, similaire à celui formé par le HPI de MtMCM (Costa *et al.*, 2006b). Des études plus approfondies sont nécessaires pour comprendre le rôle de cette flexibilité sur l'activité ATPase, la liaison à l'ADN et l'activité hélicase de la protéine en réalisant des mutations sur le HPI. La ressemblance avec MtMCM suggère que cet anneau pourrait accomplir une fonction associée spécifiquement aux hélicases dodécamériques.

- ***Implications évolutives***

La topologie en double anneau de HpDnaB n'avait jamais été observée pour une hélicase bactérienne à ce jour. Néanmoins cette conformation est similaire aux hélicases répliquatives du virus SV40 (LTag), de certaines archées (MCM) ou d'eucaryotes (Mcm 2-7) (Bae *et al.*, 2009; Costa *et al.*, 2006a, 2006b; Gómez-Llorente *et al.*, 2005; Li *et al.*, 2003; Pape *et al.*, 2003). En effet, les monomères de l'hélicase du virus SV40 (LTag) s'assemblent sur l'origine de réplication virale pour former deux hexamères qui se font face et interagissent au niveau de leurs domaines N-terminaux (Valle *et al.*, 2000). Il a également été montré que les hélicases Mcm 2-7 et MtMCM formaient des doubles hexamères sur l'origine de réplication ou des fragments d'ADNdb (Costa *et al.*, 2006a ; Evrin *et al.*, 2009 ; Remus *et al.*, 2009).

L'analyse de la structure du dodécamère des anneaux NTD de MtMCM montre que les deux hexamères interagissent grâce à des boucles qui, complexées à un atome de zinc, établissent des liaisons hydrogènes pour maintenir les deux hexamères ensemble (Fletcher *et al.*, 2003). Chez HpDnaB, les contacts entre les domaines NTD sont différents. En effet, l'association des deux anneaux NTD semble dépendre d'interactions entre les hélices $\alpha 1$ et $\alpha 5$ des têtes globulaires de chaque monomère. Ainsi, malgré une ressemblance topologique, aucun élément structural ne suggère que les hélicases Mcm 2-7, MCM, LTag et HpDnaB possèdent un mode d'assemblage similaire, pouvant provenir d'une protéine ancestrale commune.

La comparaison des DnaBs de différentes espèces bactériennes par alignement de séquences montre que seuls quelques acides aminés du domaine NTD sont strictement conservés (Figure 105). Les acides aminés impliqués dans l'interaction entre les deux anneaux NTD (L4, Q8, E80 et D81) identifiés dans la structure de HpDnaB ne le sont pas (Figure 105 ; Figure 106). Le domaine NTD est en effet le moins conservé des deux domaines principaux. Cependant, on observe que l'hélice $\alpha 1$ du NTD de HpDnaB est plus longue que celle de la structure de BstDnaB. Or c'est dans l'extension de cette hélice $\alpha 1$ que résident des résidus clés de la formation du dodécamère (L4 et Q8). Certaines espèces bactériennes possèdent des NTD ayant parfois une quinzaine d'acides aminés supplémentaires par rapport à HpDnaB (Figure 105). Il est possible que ces acides aminés forment une hélice $\alpha 1$ plus longue, et soient donc potentiellement capables d'interagir avec un autre hexamère. Il serait intéressant d'étudier les hélicases de ces espèces bactériennes afin de déterminer si elles forment des dodécamères.

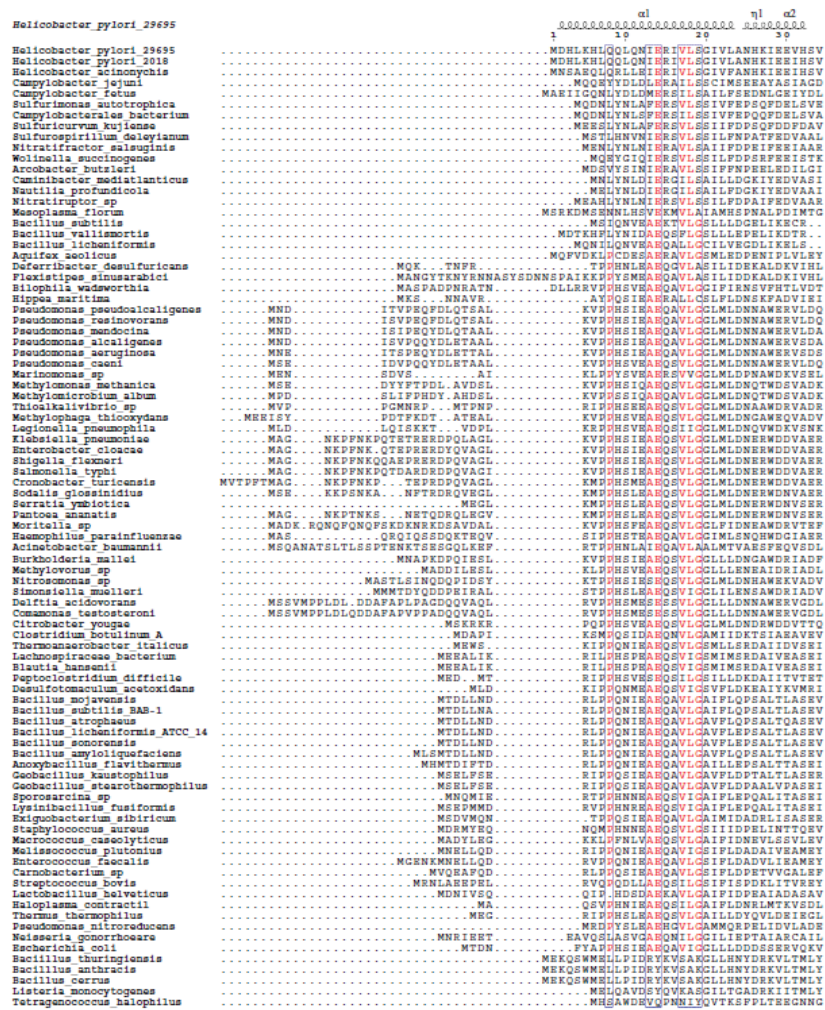


Figure 105 : Alignement de la séquence de HpDnaB (résidu 1 à 34) avec celles d'autres espèces bactériennes. Les séquences ont été obtenues grâce au serveur blastp (Altschul *et al.*, 1990) puis alignées grâce à Clustal W (Larkin *et al.*, 2007 ; Goujon *et al.*, 2010). Les similarités de séquence et les informations de structures secondaires ont été générées à partir de l'alignement par le serveur ESPript (<http://esprict.libcp.fr>) (Robert et Gouet, 2014).

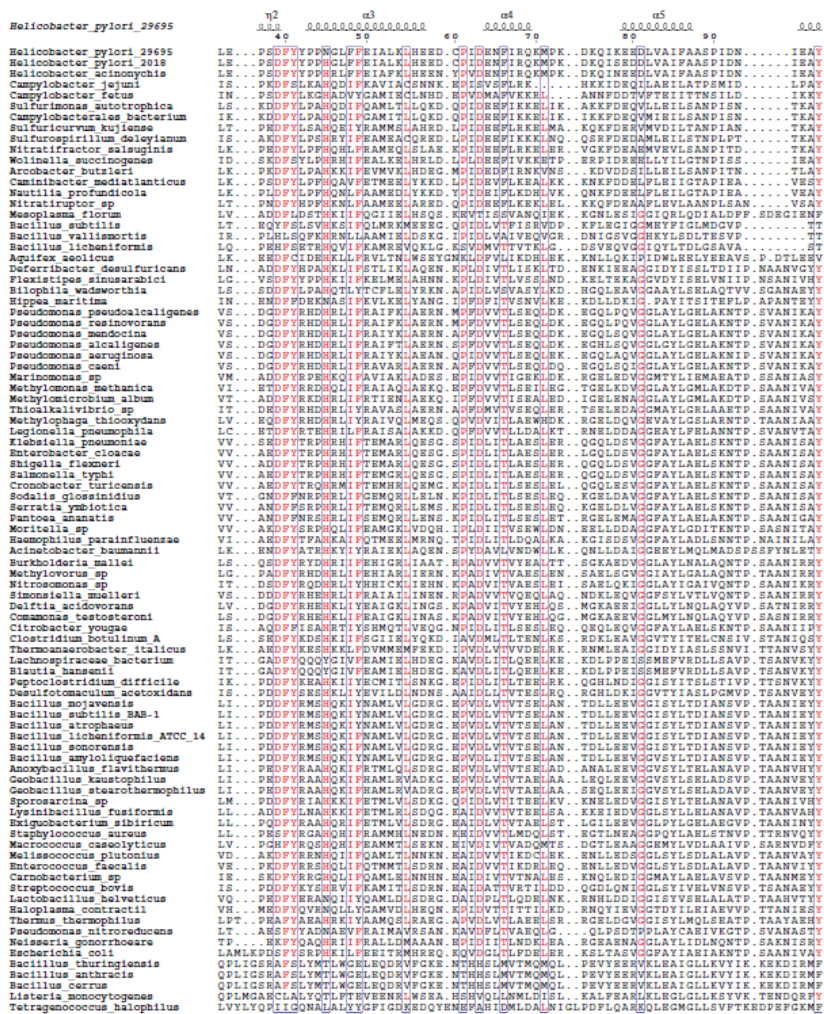


Figure 106 : Alignement de la séquence de HpDnaB (résidus 35 à 98) avec celles d'autres espèces bactériennes. Les séquences ont été obtenues grâce au serveur blastp (Altschul *et al.*, 1990) puis alignées grâce à Clustal W (Larkin *et al.*, 2007 ; Goujon *et al.*, 2010). Les similarités de séquence et les informations de structures secondaires ont été générées à partir de l'alignement par le serveur ESPrnt (<http://esprnt.ibcp.fr>) (Robert et Gouet, 2014).

Des études bio-informatiques préliminaires réalisées avec le Professeur Combet suggèrent également que HpDnaB n'est pas la seule hélicase bactérienne possédant une insertion dans son domaine ATPase (Figure 107). Ainsi, certaines caractéristiques de HpDnaB (hélice $\alpha 1$ longue et HPI) sont partiellement présentes dans d'autres DnaBs. Analyser ces protéines par SEC-MALS et SAXS permettrait d'étudier leur état d'oligomérisation et ainsi d'établir si une conformation dodécamérique est corrélée à la présence ou non de ces caractéristiques.

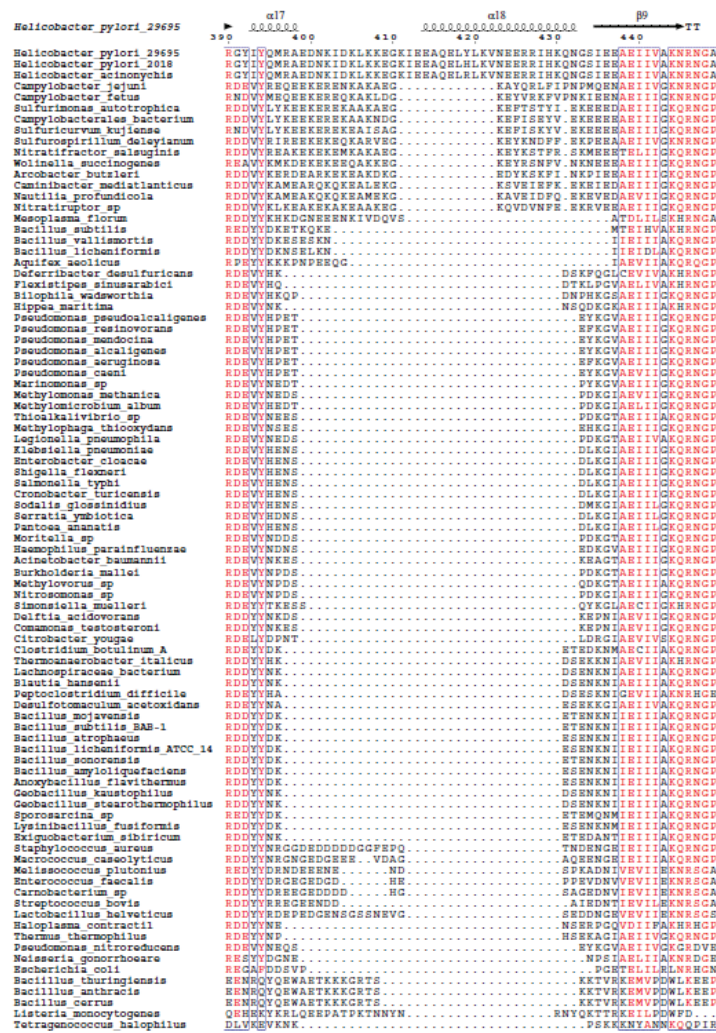


Figure 107 : Alignement de la séquence de HpDnaB (résidus 390 à 449) avec celles d'autres espèces bactériennes. Les séquences ont été obtenues grâce au serveur blastp (Altschul *et al.*, 1990) puis alignées grâce à Clustal W (Larkin *et al.*, 2007 ; Goujon *et al.*, 2010). Les similarités de séquence et les informations de structures secondaires associées ont été générées à partir de l'alignement par le serveur ESPrnt (<http://esprnt.ibcp.fr>) (Robert et Gouet, 2014).

L'existence d'une hélicase dodécamérique dans tous les règnes du vivant pose certaines interrogations : (i) La forme dodécamérique a-t-elle apporté un avantage évolutif comparé à une forme hexamérique ? Si oui, des DnaB dodécamériques devraient être présentes chez d'autres bactéries (ii) Existe-t-il un facteur de chargement inconnu, différent de DnaC et DnaI, qui reste encore à identifier et qui permet de charger ces hélicases ?

- ***Chargement de HpDnaB sur l'origine de réplication***

Il a déjà été observé qu'un dodécamère pouvait se charger sur une origine de réplication notamment pour les hélicases de types MCM (Remus *et al.*, 2009 ; Evrin *et al.*, 2009) et l'hélicase LTag du virus SV40 (Valle *et al.*, 2000). Néanmoins dans le cas de HpDnaB, il n'est pas certain que la protéine se charge sur l'ADNdb comme c'est le cas pour les hélicases MCM, ou sur l'ADNsb comme c'est le cas pour l'hélicase Twinkle de la mitochondrie (Jemt *et al.*, 2011).

Nos expériences de liaison de HpDnaB avec l'ADN ont montré que l'hélicase se liait préférentiellement à l'ADNsb plutôt qu'avec l'ADNdb et ce, seulement en présence de nucléotide. La faible affinité de HpDnaB pour l'ADNdb suggère que le chargement de l'hélicase est réalisé sur de l'ADNsb. Nos expériences de SEC-MALS confirment qu'en présence d'un ADNsb, HpDnaB reste sous forme de dodécamère. Le fait que le dodécamère ne se dissocie pas en présence d'ADNsb et d'ATP suggère que HpDnaB se chargerait sur l'ADNsb *in vivo*. Ce simple brin pourrait correspondre à la zone DUE, une fois celle-ci ouverte par l'action de HpDnaA.

Cependant la structure de HpDnaB montre que la conformation planaire du dodécamère ne permet pas le passage de l'ADN dans le canal central car les anneaux NTD et CTD sont fermés. Le passage de l'ADN de l'extérieur de l'anneau vers l'intérieur requiert donc une ouverture des anneaux NTD et CTD. Dans le cas d'EcDnaB, la liaison du facteur de chargement sur l'hélicase provoque un changement de conformation de l'hexamère qui forme une spirale, ouvrant ainsi les anneaux NTD et CTD et pourrait permettre le passage de l'ADNsb (Arias-Palomo *et al.*, 2013).

Nos expériences de SAXS ont montré que la conformation de la protéine en solution diffère de celle de la structure cristallographique. Nos différents modèles construits à partir de structures cristallographiques connues suggèrent que la conformation qui correspond le mieux aux données expérimentales de HpDnaB est un dodécamère en spirale-dilatée en présence ou non de nucléotide. Ces données montrent également que la protéine est dynamique en solution. Or cette forme spiralée du dodécamère laisse une ouverture latérale sur chaque hexamère (Figure 108). Cette ouverture pourrait être suffisamment large pour permettre le passage de l'ADNsb à travers l'anneau CTD et NTD pour qu'il s'insère à l'intérieur de l'hélicase.

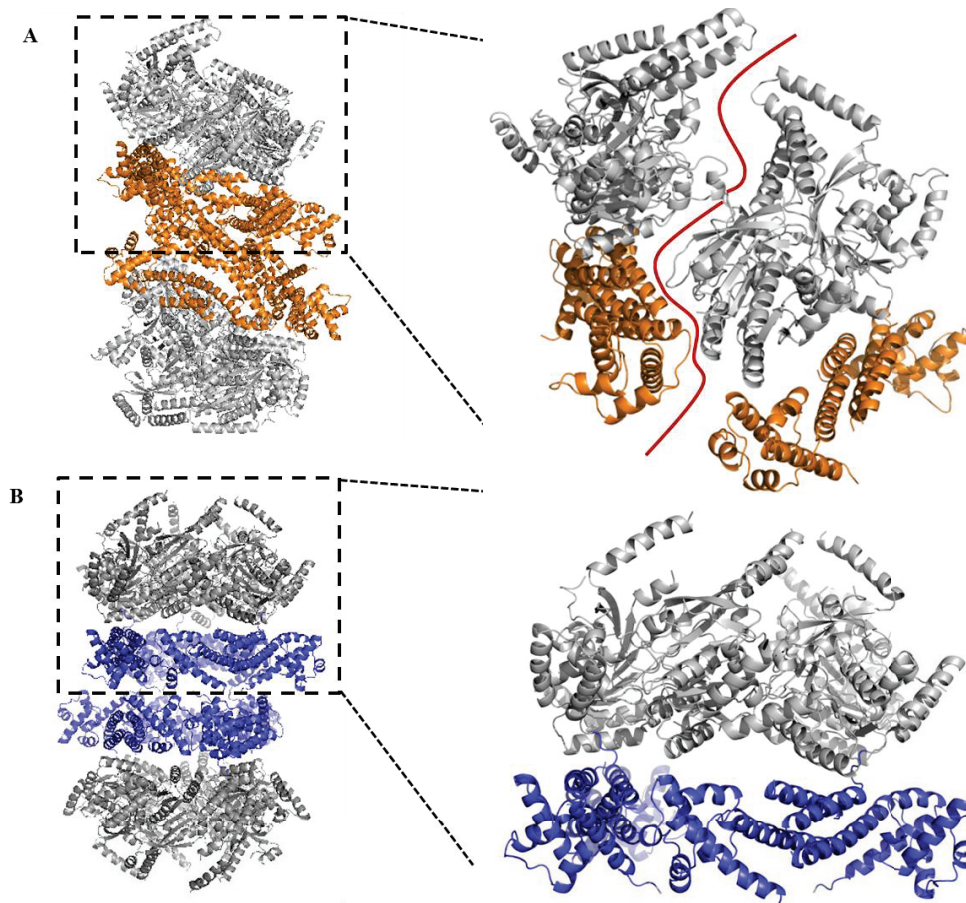


Figure 108 : Modèles de HpDnaB en conformation spirale-dilatée (A) et planaire-dilatée (B). Un hexamère a été agrandi avec seulement quatre sous-unités représentées pour plus de clarté. Le courbe rouge représente un ADNsb pouvant passer à travers l'ouverture latérale des anneaux NTD et CTD du dodécamère en conformation spirale-dilatée.

- **Formation du primosome de *H. pylori***

Après le chargement sur l'ADN, l'hélicase sert de plateforme pour l'arrimage de la primase DnaG. Chez *B. stearothermophilus*, DnaG interagit grâce à son domaine de fixation à l'hélicase (HBD) au niveau de deux sites de fixation situés sur deux domaines NTD (Figure 109). Chez *E. coli* et *B. stearothermophilus*, le complexe entre la primase et l'hélicase possède une stœchiométrie 1 :2 soit une molécule de DnaG pour deux molécules de DnaB (Mitkova *et al.*, 2003 ; Bird *et al.*, 2000 ; Bailey *et al.*, 2007). De plus, une étude plus récente a également montré que les protéines BstDnaB et BstDnaG formaient un complexe ternaire transitoire avec un facteur de chargement (Liu *et al.*, 2013).

Dans le dodécamère de HpDnaB, les hexamères interagissent tête-tête au niveau des anneaux NTD. Ainsi la zone d'interaction entre HpDnaG^{HBD} et le NTD 1 est stériquement encombrée par l'interaction avec le NTD 1 du deuxième hexamère (Figure 109).

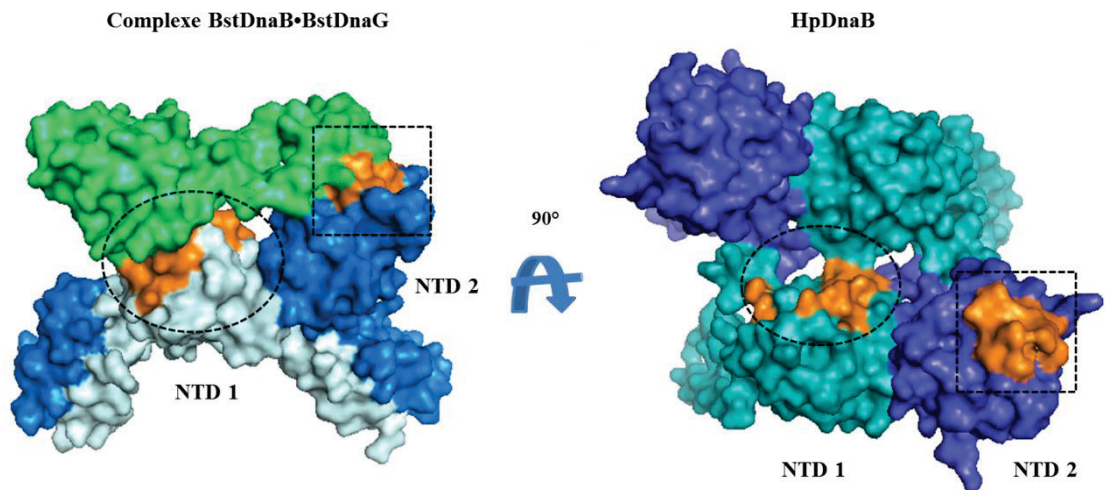


Figure 109 : Zone d'interaction (orange) d'une molécule de BstDnaG (vert) avec deux BstDnaBNTD (bleu clair et foncé) (à gauche). Représentation de cette même zone sur deux NTD de HpDnaB (à droite).

Nos études de titrations calorimétriques montrent que l'interaction entre HpDnaB et HpDnaG^{HBD} est faible, de l'ordre de 20 μM . Cette faible affinité est comparable à ce qui est observé avec les protéines d'*E. coli* (Oakley *et al.*, 2005 ; Mitkova *et al.*, 2003).

Cette faible affinité pourrait avoir une fonction durant la synthèse des amorces par la primase. Ainsi chez *E. coli*, du fait de la faible affinité entre DnaB et DnaG, la primase se détacherait du primosome et resterait accrochée aux amorces d'ARN synthétisées sur le brin retard (Soultanas, 2005 ; Yuzhakov *et al.*, 1999). Néanmoins même avec une affinité de l'ordre du micromolaire, HpDnaG^{HBD} a un effet important sur les activités enzymatiques de HpDnaB et sur son affinité pour l'ADNsb. En effet, le dodécamère est peu actif (activités ATPase et hélicase) alors qu'en présence d'HpDnaG^{HBD} ces deux activités augmentent. De plus, lorsque HpDnaB en dodécamère interagit avec HpDnaG^{HBD}, son affinité pour l'ADNsb augmente. Cela suggère que le dodécamère serait une forme inactive mais qu'après interaction avec HpDnaG^{HBD}, l'affinité de HpDnaB pour l'ADNsb augmente puis est activée. L'ensemble de nos résultats indiquent donc que l'interaction avec HpDnaG^{HBD} sépare le dodécamère. Cette interaction avec la primase représenterait alors un moyen de réguler l'action de l'hélicase en absence de facteur de chargement.

- **Modèle de l'initiation de la réplication de *H. pylori***

L'ensemble de nos résultats suggère que lors de l'initiation de la réplication, *H. pylori*, pourrait utiliser un mécanisme différent d'*E. coli* ou de *B. subtilis*. Chez *H. pylori* une protéine unique HobA nécessaire à l'initiation de la réplication pourrait modifier la topologie

de l'ADN en interagissant avec DnaA. L'origine de réplication bipartite de la bactérie, est composée de deux sous-régions *oriC1* et *oriC2* séparées par le gène *dnaA*. Cette organisation, unique chez les bactéries à Gram négatif mais existant chez les firmicutes, suggère qu'après la liaison de HpDnaA aux deux sous-régions, HobA en se liant au domaine N-terminal de HpDnaA (Natrajan *et al.*, 2009), pourrait participer à l'orisome qui prendrait la forme d'une boucle. En effet, l'affinité de DnaA pour *oriC1* est significativement augmentée en présence de HobA d'une manière similaire à ce qui est observé chez *E. coli* avec la protéine DiaA (Zawilak-Pawlik *et al.*, 2007 et 2001). De plus, des études de biochimie ont révélées que HpDnaA se liait sur la région *oriC2* seulement si cette région de l'ADN était superenroulée. Ceci contraste avec la région *oriC1* sur laquelle HpDnaA se lie quelque soit la topologie de l'ADN (Donczew *et al.*, 2011). Cependant le rôle de HobA sur HpDnaA liée à *oriC2* est encore inconnu à ce jour. Le génome de *H. pylori* ne code pas pour des protéines comme IHF ou Fis chez *E. coli* qui régulent la formation de l'orisome chez cet organisme. HobA pourrait donc servir à réguler la topologie de l'ADN pour ainsi contrôler la formation de l'orisome chez *H. pylori* (Donczew *et al.*, 2014).

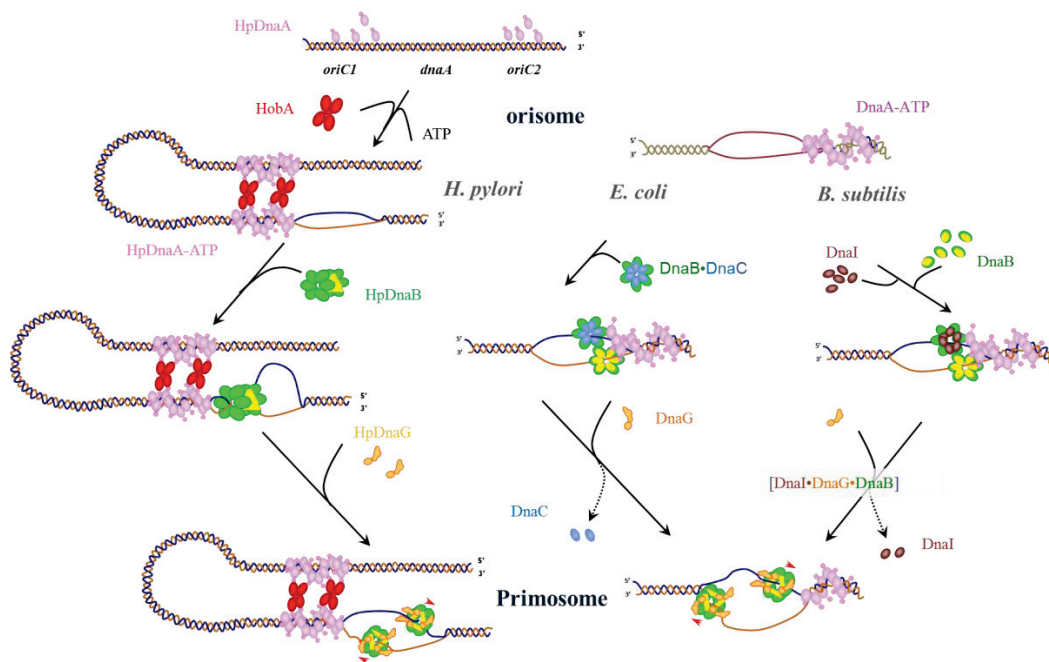


Figure 110 : Comparaison du modèle de l'initiation de la réplication chez *H. pylori* (à gauche) avec ceux d'*E. coli* et de *B. subtilis* (à droite).

L'ensemble de nos résultats nous permet de proposer un modèle original de l'assemblage du primosome de *H. pylori* différent des organismes modèles (Figure 110). Une fois l'orisome de *H. pylori* formé et après ouverture de la région riche en AT, le dodécamère de HpDnaB pourrait se positionner directement sur l'ADNsb par le passage de l'ADN à l'intérieur du

double anneau spiralé. La reconnaissance du site d'ouverture se ferait par un procédé différent d'*E. coli*. En effet, il n'existe pas de preuves d'interaction entre HpDnaA et HpDnaB. Des travaux non publiés réalisés par notre équipe montrent que les protéines HpDnaB et HpDnaA n'interagissent pas *in vitro* ou en système double hybride. Il est donc possible que HpDnaB se charge sur l'ADNsb sans interagir avec HpDnaA. La conformation en dodécamère permet de compenser l'absence de facteur de chargement puisque les deux hexamères sont orientés dans des directions opposées. Dans notre modèle nous proposons que les deux brins d'ADNsb se positionnent à l'intérieur d'un seul dodécamère comme cela est observé pour LTag. Cette étape n'est cependant pas démontrée et il est possible qu'un dodécamère soit positionné sur chaque ADNsb.

Le dodécamère de HpDnaB n'a qu'une faible activité hélicase suggérant qu'il est inactif pour l'élongation des fourches. L'ADNsb et l'hydrolyse de l'ATP ne permettant pas la dissociation du double anneau, un autre facteur est donc nécessaire. Nous proposons que la primase HpDnaG soit ce facteur. En interagissant avec HpDnaB, la primase provoquerait alors la séparation des deux hexamères permettant ainsi à la réplication de débuter. Cette étape ressemble d'un point de vue fonctionnel aux autres organismes car la primase DnaG active également DnaB d'*E. coli* ou de *B. subtilis*. Si ce modèle reste incomplet, il est compatible avec nos observations et expliquerait les résultats obtenus *in vivo* par d'autres équipes. Cependant, le chargement de HpDnaB sur l'ADNsb n'a pas encore été démontré et devra faire l'objet d'études futures. De plus, en absence d'interaction entre HpDnaA et HpDnaB, il est difficile dans l'état actuel de nos connaissances de comprendre comment HpDnaB peut reconnaître la région sur laquelle elle doit se charger. On peut supposer qu'un autre facteur pourrait intervenir dans ce processus.

Enfin, nos travaux démontrent, si cela était nécessaire, qu'il est important d'étudier des systèmes différents d'*E. coli* ou de *B. subtilis*. En effet, malgré le caractère essentiel de la réplication et la forte pression de sélection s'exerçant sur les protéines impliquées, les systèmes de réplifications peuvent varier considérablement d'une bactérie à une autre. Ainsi, le modèle d'initiation de *H. pylori* représente vraisemblablement une adaptation à une niche écologique particulière. Il sera intéressant de déterminer si ce modèle existe chez d'autres bactéries ou, plus vraisemblablement, si d'autres modèles d'initiation existent.

Références Bibliographiques

- Abdul Rehman, S. A., Verma, V., Mazumder, M., Dhar, S. K. and Gourinath, S. (2013). "Crystal structure and mode of helicase binding of the C-terminal domain of primase from *Helicobacter pylori*." *J Bacteriol* **195**(12): 2826-2838.
- Abe, Y., Jo, T., Matsuda, Y., Matsunaga, C., Katayama, T. and Ueda, T. (2007). "Structure and function of DnaA N-terminal domains: specific sites and mechanisms in inter-DnaA interaction and in DnaB helicase loading on *oriC*." *J Biol Chem* **282**(24): 17816-17827.
- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." *Acta Crystallogr D Biol Crystallogr* **66**(Pt 2): 213-221.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). "Basic local alignment search tool." *J Mol Biol* **215**(3): 403-410.
- Arias-Palomo, E., O'Shea, V. L., Hood, I. V. and Berger, J. M. (2013). "The bacterial DnaC helicase loader is a DnaB ring breaker." *Cell* **153**(2): 438-448.
- Arthur, A. K., Hoss, A. and Fanning, E. (1988). "Expression of simian virus 40 T antigen in *Escherichia coli*: localization of T-antigen origin DNA-binding domain to within 129 amino acids." *J Virol* **62**(6): 1999-2006.
- Atlung, T., Clausen, E. S. and Hansen, F. G. (1985). "Autoregulation of the *dnaA* gene of *Escherichia coli* K12." *Mol Gen Genet* **200**(3): 442-450.
- Backert, S. and Clyne, M. (2011). "Pathogenesis of *Helicobacter pylori* infection." *Helicobacter* **16** Suppl 1: 19-25.
- Bailey, S., Eliason, W. K. and Steitz, T. A. (2007a). "The crystal structure of the *Thermus aquaticus* DnaB helicase monomer." *Nucleic Acids Res* **35**(14): 4728-4736.
- Bailey, S., Eliason, W. K. and Steitz, T. A. (2007b). "Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase." *Science* **318**(5849): 459-463.

- Baker, T. A., Sekimizu, K., Funnell, B. E. and Kornberg, A. (1986). "Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the *Escherichia coli* chromosome." *Cell* **45**(1): 53-64.
- Bastos, J., Carreira, H., La Vecchia, C. and Lunet, N. (2013). "Childcare attendance and *Helicobacter pylori* infection: systematic review and meta-analysis." *Eur J Cancer Prev* **22**(4): 311-319.
- Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. (2011). "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM." *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4): 271-281.
- Bell, S. P. and Stillman, B. (1992). "ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex." *Nature* **357**(6374): 128-134.
- Bird, L. E., Pan, H., Soultanas, P. and Wigley, D. B. (2000). "Mapping protein-protein interactions within a stable complex of DNA primase and DnaB helicase from *Bacillus stearothermophilus*." *Biochemistry* **39**(1): 171-182.
- Biswas, E. E. and Biswas, S. B. (1999). "Mechanism of DnaB helicase of *Escherichia coli*: structural domains involved in ATP hydrolysis, DNA binding, and oligomerization." *Biochemistry* **38**(34): 10919-10928.
- Biswas, E. E., Chen, P. H. and Biswas, S. B. (2002). "Modulation of enzymatic activities of *Escherichia coli* DnaB helicase by single-stranded DNA-binding proteins." *Nucleic Acids Res* **30**(13): 2809-2816.
- Biswas, S. B., Chen, P. H. and Biswas, E. E. (1994). "Structure and function of *Escherichia coli* DnaB protein: role of the N-terminal domain in helicase activity." *Biochemistry* **33**(37): 11307-11314.
- Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. and Bricogne, G. (2004). "Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT." *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2210-2221.
- Blundell, T. L. and Johnson, L. N. (1976). *Protein Crystallography*. New-York, Academic Press Inc.

- Bochman, M. L., Bell, S. P. and Schwacha, A. (2008). "Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability." *Mol Cell Biol* **28**(19): 5865-5873.
- Bowman, G. D., O'Donnell, M. and Kuriyan, J. (2004). "Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex." *Nature* **429**(6993): 724-730.
- Braun, R. E., O'Day, K. and Wright, A. (1985). "Autoregulation of the DNA replication gene *dnaA* in *E. coli* K-12." *Cell* **40**(1): 159-169.
- Brewster, A. S., Wang, G., Yu, X., Greenleaf, W. B., Carazo, J. M., Tjajadi, M., Klein, M. G. and Chen, X. S. (2008). "Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase." *Proc Natl Acad Sci U S A* **105**(51): 20191-20196.
- Brown, L. M. (2000). "Helicobacter pylori: epidemiology and routes of transmission." *Epidemiol Rev* **22**(2): 283-297.
- Brunger, A. T. (1992a). "Free R value : a novel statistical quantity for assesseing the accuracy of crystal structure." *Nature* **355**: 475.
- Camara, J. E., Breier, A. M., Brendler, T., Austin, S., Cozzarelli, N. R. and Crooke, E. (2005). "Hda inactivation of DnaA is the predominant mechanism preventing hyperinitiation of *Escherichia coli* DNA replication." *EMBO Rep* **6**(8): 736-741.
- Caspi, R., Pacek, M., Consiglieri, G., Helinski, D. R., Toukdarian, A. and Konieczny, I. (2001). "A broad host range replicon with different requirements for replication initiation in three bacterial species." *EMBO J* **20**(12): 3262-3271.
- Chang, Y. P., Xu, M., Machado, A. C., Yu, X. J., Rohs, R. and Chen, X. S. (2013). "Mechanism of origin DNA recognition and assembly of an initiator-helicase complex by SV40 large tumor antigen." *Cell Rep* **3**(4): 1117-1127.
- Chodavarapu, S., Felczak, M. M., Yaniv, J. R. and Kaguni, J. M. (2008). "Escherichia coli DnaA interacts with HU in initiation at the *E. coli* replication origin." *Mol Microbiol* **67**(4): 781-792.
- Colliex, C. (1996). *La Microscopie électronique*.

- Costa, A., Hood, I. V. and Berger, J. M. (2013). "Mechanisms for initiating cellular DNA replication." *Annu Rev Biochem* **82**: 25-54.
- Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R. and Berger, J. M. (2011). "The structural basis for MCM2-7 helicase activation by GINS and Cdc45." *Nat Struct Mol Biol* **18**(4): 471-477.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006a). "Structural studies of the archaeal MCM complex in different functional states." *J Struct Biol* **156**(1): 210-219.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006b). "Structural basis of the *Methanothermobacter thermautotrophicus* MCM helicase activity." *Nucleic Acids Res* **34**(20): 5829-5838.
- Covacci, A., Telford, J. L., Del Giudice, G., Parsonnet, J. and Rappuoli, R. (1999). "*Helicobacter pylori* virulence and genetic geography." *Science* **284**(5418): 1328-1333.
- David, G. and Perez, J. (2009). "Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline." *Journal of Applied Crystallography* **42**(5): 892-900.
- Deb, S., DeLucia, A. L., Baur, C. P., Koff, A. and Tegtmeyer, P. (1986). "Domain structure of the simian virus 40 core origin of replication." *Mol Cell Biol* **6**(5): 1663-1670.
- Diederichs, K. and Karplus, P. A. (2013). "Better models by discarding data?" *Acta Crystallogr D Biol Crystallogr* **69**(Pt 7): 1215-1222.
- Diffley, J. F., Cocker, J. H., Dowell, S. J. and Rowley, A. (1994). "Two steps in the assembly of complexes at yeast replication origins in vivo." *Cell* **78**(2): 303-316.
- Dmitri, I. S. and Michel, H. J. K. (2003). "Small-angle scattering studies of biological macromolecules in solution." *Reports on Progress in Physics* **66**(10): 1735.
- Donczew, R., Mielke, T., Jaworski, P., Zakrzewska-Czerwinska, J. and Zawilak-Pawlik, A. (2014a). "Assembly of *Helicobacter pylori* initiation complex is determined by sequence-specific and topology-sensitive DnaA-oriC interactions." *J Mol Biol* **426**(15): 2769-2782.

- Donczew, R., Weigel, C., Lurz, R., Zakrzewska-Czerwinska, J. and Zawilak-Pawlik, A. (2012). "Helicobacter pylori oriC--the first bipartite origin of chromosome replication in Gram-negative bacteria." *Nucleic Acids Res* **40**(19): 9647-9660.
- Donczew, R., Zakrzewska-Czerwinska, J. and Zawilak-Pawlik, A. (2014b). "Beyond DnaA: the role of DNA topology and DNA methylation in bacterial replication initiation." *J Mol Biol* **426**(12): 2269-2282.
- Doublet, S. (1997). "Preparation of selenomethionyl proteins for phase determination." *Methods Enzymol* **276**: 523-530.
- Duderstadt, K. E., Chuang, K. and Berger, J. M. (2011). "DNA stretching by bacterial initiators promotes replication origin opening." *Nature* **478**(7368): 209-213.
- Emsley, P. and Cowtan, K. (2004). "Coot: model-building tools for molecular graphics." *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1): 2126-2132.
- Erzberger, J. P., Mott, M. L. and Berger, J. M. (2006). "Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling." *Nat Struct Mol Biol* **13**(8): 676-683.
- Erzberger, J. P., Pirruccello, M. M. and Berger, J. M. (2002). "The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation." *EMBO J* **21**(18): 4763-4773.
- ESPrript. (2014). "<http://esprript.ibcp.fr>." from <http://esprript.ibcp.fr>.
- Evans, P. (2006). "Scaling and assessment of data quality." *Acta Crystallogr D Biol Crystallogr* **62**(Pt 1): 72-82.
- Evans, P. R. (2011). "An introduction to data reduction: space-group determination, scaling and intensity statistics." *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4): 282-292.
- Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B. and Speck, C. (2009). "A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication." *Proc Natl Acad Sci U S A* **106**(48): 20240-20245.
- Farge, G., Holmlund, T., Khvorostova, J., Rofougaran, R., Hofer, A. and Falkenberg, M. (2008). "The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities." *Nucleic Acids Res* **36**(2): 393-403.

Feigin, L. A. and Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*, Plenum Press/Springer.

Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. and Craievich, A. F. (2010). "Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale." *Journal of Applied Crystallography* **43**(1): 101-109.

Flatten, I. and Skarstad, K. (2013). "The Fis protein has a stimulating role in initiation of replication in *Escherichia coli* in vivo." *PLoS One* **8**(12): e83562.

Fletcher, R. J., Bishop, B. E., Leon, R. P., Sclafani, R. A., Ogata, C. M. and Chen, X. S. (2003). "The structure and function of MCM from archaeal *M. Thermoautotrophicum*." *Nat Struct Biol* **10**(3): 160-167.

FoXS. "<http://modbase.compbio.ucsf.edu/foxs/>

FOXSTROT. "<http://wwwsynchrotron-soleil.fr/Recherche/LignesLumiere/SWING.>"

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. and Leith, A. (1996). "SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields." *J Struct Biol* **116**(1): 190-199.

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T. and Yokoyama, S. (2003). "Structural basis of replication origin recognition by the DnaA protein." *Nucleic Acids Res* **31**(8): 2077-2086.

Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T. and Yokoyama, S. (2003). "Structural basis of replication origin recognition by the DnaA protein." *Nucleic Acids Res* **31**(8): 2077-2086.

Fuller, R. S., Funnell, B. E. and Kornberg, A. (1984). "The dnaA protein complex with the *E. coli* chromosomal replication origin (*oriC*) and other DNA sites." *Cell* **38**(3): 889-900.

Gai, D., Zhao, R., Li, D., Finkielstein, C. V. and Chen, X. S. (2004). "Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen." *Cell* **119**(1): 47-60.

Garza-Gonzalez, E., Perez-Perez, G. I., Maldonado-Garza, H. J. and Bosques-Padilla, F. J. (2014). "A review of *Helicobacter pylori* diagnosis, treatment, and methods to detect eradication." *World J Gastroenterol* **20**(6): 1438-1449.

Gaudier, M., Schuwirth, B. S., Westcott, S. L. and Wigley, D. B. (2007). "Structural basis of DNA replication origin recognition by an ORC protein." *Science* **317**(5842): 1213-1216.

GEFH. "<http://www.helicobacter.fr/>

Gil, R., Silva, F. J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., Rausell, C., Kamerbeek, J., Gadau, J., Holldobler, B., van Ham, R. C., Gross, R. and Moya, A. (2003). "The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes." *Proc Natl Acad Sci U S A* **100**(16): 9388-9393.

Gomez-Llorente, Y., Fletcher, R. J., Chen, X. S., Carazo, J. M. and San Martin, C. (2005). "Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from *Methanobacterium thermoautotrophicum*." *J Biol Chem* **280**(49): 40909-40915.

Goodwin, C. S. and Armstrong, J. A. (1990). "Microbiological aspects of *Helicobacter pylori* (*Campylobacter pylori*)." *Eur J Clin Microbiol Infect Dis* **9**(1): 1-13.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. and Lopez, R. (2010). "A new bioinformatics analysis tools framework at EMBL-EBI." *Nucleic Acids Res* **38**(Web Server issue): W695-699.

Grimwade, J. E., Ryan, V. T. and Leonard, A. C. (2000). "IHF redistributes bound initiator protein, DnaA, on supercoiled oriC of *Escherichia coli*." *Mol Microbiol* **35**(4): 835-844.

Guinier, A. (1939). *Diffraction des rayons X aux très petits angles: application à l'étude des phénomènes ultramicroscopiques*, Masson.

Hayashi, M., Ogura, Y., Harry, E. J., Ogasawara, N. and Moriya, S. (2005). "Bacillus subtilis YabA is involved in determining the timing and synchrony of replication initiation." *FEMS Microbiol Lett* **247**(1): 73-79.

Hwang, D. S. and Kornberg, A. (1992). "Opening of the replication origin of *Escherichia coli* by DnaA protein with protein HU or IHF." *J Biol Chem* **267**(32): 23083-23086.

Ilyina, T. V., Gorbalenya, A. E. and Koonin, E. V. (1992). "Organization and evolution of bacterial and bacteriophage primase-helicase systems." *J Mol Evol* **34**(4): 351-357.

Ioannou, C., Schaeffer, P. M., Dixon, N. E. and Soultanas, P. (2006). "Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in *Bacillus subtilis*." *Nucleic Acids Res* **34**(18): 5247-5258.

Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K. and Katayama, T. (2004). "DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication." *J Biol Chem* **279**(44): 45546-45555.

Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J. and Steitz, T. A. (2012). "The hexameric helicase DnaB adopts a nonplanar conformation during translocation." *Cell* **151**(2): 267-277.

IUCr. "<http://it.iucr.org/>

Iyer, L. M., Leipe, D. D., Koonin, E. V. and Aravind, L. (2004). "Evolutionary history and higher order classification of AAA+ ATPases." *J Struct Biol* **146**(1-2): 11-31.

Jacob, F. and Brenner, S. (1963). "[On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon]." *C R Hebd Seances Acad Sci* **256**: 298-300.

Jacques, D. A. and Trewhella, J. (2010). "Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls." *Protein Science* **19**(4): 642-657.

Jameson, K. H., Rostami, N., Fogg, M. J., Turkenburg, J. P., Grahl, A., Murray, H. and Wilkinson, A. J. (2014). "Structure and interactions of the *Bacillus subtilis* sporulation inhibitor of DNA replication, SirA, with domain I of DnaA." *Mol Microbiol* **93**(5): 975-991.

Jemt, E., Farge, G., Backstrom, S., Holmlund, T., Gustafsson, C. M. and Falkenberg, M. (2011). "The mitochondrial DNA helicase TWINKLE can assemble on a closed circular template and support initiation of DNA synthesis." *Nucleic Acids Res* **39**(21): 9238-9249.

Jiang, Y., Yao, S., Helinski, D. and Toukdarian, A. (2006). "Functional analysis of two putative chromosomal replication origins from *Pseudomonas aeruginosa*." *Plasmid* **55**(3): 194-200.

- Johnson, S. K., Bhattacharyya, S. and Griep, M. A. (2000). "DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates." *Biochemistry* **39**(4): 736-744.
- Joo, W. S., Kim, H. Y., Purviance, J. D., Sreekumar, K. R. and Bullock, P. A. (1998). "Assembly of T-antigen double hexamers on the simian virus 40 core origin requires only a subset of the available binding sites." *Mol Cell Biol* **18**(5): 2677-2687.
- Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." *J. Appl. Cryst.* **26**: 795-800.
- Kaguni, J. M. (1997). "Escherichia coli DnaA protein: the replication initiator." *Mol Cells* **7**(2): 145-157.
- Kaguni, J. M. (2006). "DnaA: controlling the initiation of bacterial DNA replication and more." *Annu Rev Microbiol* **60**: 351-375.
- Kantardjieff, K. A. and Rupp, B. (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." *Protein Sci* **12**(9): 1865-1871.
- Kaplan-Turkoz, B., Jimenez-Soto, L. F., Dian, C., Ertl, C., Remaut, H., Louche, A., Tosi, T., Haas, R. and Terradot, L. (2012). "Structural insights into Helicobacter pylori oncoprotein CagA interaction with beta1 integrin." *Proc Natl Acad Sci U S A* **109**(36): 14640-14645.
- Kaplan-Turkoz, B. and Terradot, L. (2013). "[Structure and mode of injection of the oncoprotein CagA of Helicobacter pylori]." *Med Sci (Paris)* **29**(1): 33-36.
- Karplus, P. A. and Diederichs, K. (2012). "Linking crystallographic model and data quality." *Science* **336**(6084): 1030-1033.
- Kashav, T., Nitharwal, R., Abdulrehman, S. A., Gabdoulkhakov, A., Saenger, W., Dhar, S. K. and Gourinath, S. (2009). "Three-dimensional structure of N-terminal domain of DnaB helicase and helicase-primase interactions in Helicobacter pylori." *PLoS One* **4**(10): e7515.
- Katayama, T., Kubota, T., Kurokawa, K., Crooke, E. and Sekimizu, K. (1998). "The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase." *Cell* **94**(1): 61-71.

- Katayama, T., Ozaki, S., Keyamura, K. and Fujimitsu, K. (2010). "Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC." *Nat Rev Microbiol* **8**(3): 163-170.
- Kato, J. and Katayama, T. (2001). "Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*." *EMBO J* **20**(15): 4253-4262.
- Keck, J. L., Roche, D. D., Lynch, A. S. and Berger, J. M. (2000). "Structure of the RNA polymerase domain of *E. coli* primase." *Science* **287**(5462): 2482-2486.
- Kelch, B. A., Makino, D. L., O'Donnell, M. and Kuriyan, J. (2011). "How a DNA polymerase clamp loader opens a sliding clamp." *Science* **334**(6063): 1675-1680.
- Kelch, B. A., Makino, D. L., O'Donnell, M. and Kuriyan, J. (2012). "Clamp loader ATPases and the evolution of DNA replication machinery." *BMC Biol* **10**: 34.
- Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Su'etsugu, M., Fujimitsu, K., Kagawa, W., Yokoyama, S., Kurumizaka, H. and Katayama, T. (2007). "The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes." *Genes Dev* **21**(16): 2083-2099.
- Kitagawa, R., Mitsuki, H., Okazaki, T. and Ogawa, T. (1996). "A novel DnaA protein-binding site at 94.7 min on the *Escherichia coli* chromosome." *Mol Microbiol* **19**(5): 1137-1147.
- Kitagawa, R., Ozaki, T., Moriya, S. and Ogawa, T. (1998). "Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for *Escherichia coli* DnaA protein." *Genes Dev* **12**(19): 3032-3043.
- Kleywegt, G. J. and Brunger, A. T. (1996). "Checking your imagination: applications of the free R value." *Structure* **4**(8): 897-904.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. and Svergun, D. I. (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." *Journal of Applied Crystallography* **36**(5): 1277-1282.
- Konieczny, I. and Helinski, D. R. (1997). "Helicase delivery and activation by DnaA and TrfA proteins during the initiation of replication of the broad host range plasmid RK2." *J Biol Chem* **272**(52): 33312-33318.

Koonin, E. V. (1992). "DnaC protein contains a modified ATP-binding motif and belongs to a novel family of ATPases including also DnaA." *Nucleic Acids Res* **20**(8): 1997.

Kornberg, A. and Baker, T. A. (1992). University Science Books.

Kucherer, C., Lothar, H., Kolling, R., Schauzu, M. A. and Messer, W. (1986). "Regulation of transcription of the chromosomal dnaA gene of Escherichia coli." *Mol Gen Genet* **205**(1): 115-121.

Kusters, J. G., van Vliet, A. H. and Kuipers, E. J. (2006). "Pathogenesis of Helicobacter pylori infection." *Clin Microbiol Rev* **19**(3): 449-490.

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

Lee, P. S. and Grossman, A. D. (2006). "The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in Bacillus subtilis." *Mol Microbiol* **60**(4): 853-869.

Leonard, A. C. and Grimwade, J. E. (2011). "Regulation of DnaA assembly and activity: taking directions from the genome." *Annu Rev Microbiol* **65**: 19-35.

Leslie, A. G. (2006). "The integration of macromolecular diffraction data." *Acta Crystallogr D Biol Crystallogr* **62**(Pt 1): 48-57.

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J. A., Fanning, E., Jochimiak, A., Szakonyi, G. and Chen, X. S. (2003). "Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen." *Nature* **423**(6939): 512-518.

Liu, B., Eliason, W. K. and Steitz, T. A. (2013). "Structure of a helicase-helicase loader complex reveals insights into the mechanism of bacterial primosome assembly." *Nat Commun* **4**: 2495.

Lo, Y. H., Tsai, K. L., Sun, Y. J., Chen, W. T., Huang, C. Y. and Hsiao, C. D. (2009). "The crystal structure of a replicative hexameric helicase DnaC and its complex with single-stranded DNA." *Nucleic Acids Res* **37**(3): 804-814.

- Lopper, M., Boonsombat, R., Sandler, S. J. and Keck, J. L. (2007). "A hand-off mechanism for primosome assembly in replication restart." *Mol Cell* **26**(6): 781-793.
- Loscha, K. V., Jaudzems, K., Ioannou, C., Su, X. C., Hill, F. R., Otting, G., Dixon, N. E. and Liepinsh, E. (2009). "A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader." *Nucleic Acids Res* **37**(7): 2395-2404.
- Lu, M., Campbell, J. L., Boye, E. and Kleckner, N. (1994). "SeqA: a negative modulator of replication initiation in *E. coli*." *Cell* **77**(3): 413-426.
- Ludlam, A. V., McNatt, M. W., Carr, K. M. and Kaguni, J. M. (2001). "Essential amino acids of *Escherichia coli* DnaC protein in an N-terminal domain interact with DnaB helicase." *J Biol Chem* **276**(29): 27345-27353.
- Ludtke, S. J., Baldwin, P. R. and Chiu, W. (1999). "EMAN: semiautomated software for high-resolution single-particle reconstructions." *J Struct Biol* **128**(1): 82-97.
- Luo, X., Sanford, D. G., Bullock, P. A. and Bachovchin, W. W. (1996). "Solution structure of the origin DNA-binding domain of SV40 T-antigen." *Nat Struct Biol* **3**(12): 1034-1039.
- Lyubimov, A. Y., Costa, A., Bleichert, F., Botchan, M. R. and Berger, J. M. (2012). "ATP-dependent conformational dynamics underlie the functional asymmetry of the replicative helicase from a minimalist eukaryote." *Proc Natl Acad Sci U S A* **109**(30): 11999-12004.
- Makowska-Grzyska, M. and Kaguni, J. M. (2010). "Primase directs the release of DnaC from DnaB." *Mol Cell* **37**(1): 90-101.
- Marahrens, Y. and Stillman, B. (1992). "A yeast chromosomal origin of DNA replication defined by multiple functional elements." *Science* **255**(5046): 817-823.
- Marshall, B. J. and Warren, J. R. (1984). "Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration." *Lancet* **1**(8390): 1311-1315.
- Marszalek, J. and Kaguni, J. M. (1994). "DnaA protein directs the binding of DnaB protein in initiation of DNA replication in *Escherichia coli*." *J Biol Chem* **269**(7): 4883-4890.
- Matthews, B. W. (1968). "Solvent content of protein crystals." *J Mol Biol* **33**(2): 491-497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007). "Phaser crystallographic software." *J Appl Crystallogr* **40**(Pt 4): 658-674.

- McCoy, A. J., Storoni, L. C. and Read, R. J. (2004). "Simple algorithm for a maximum-likelihood SAD function." *Acta Crystallogr D Biol Crystallogr* **60**(Pt 7): 1220-1228.
- McGarry, K. C., Ryan, V. T., Grimwade, J. E. and Leonard, A. C. (2004). "Two discriminatory binding sites in the Escherichia coli replication origin are required for DNA strand opening by initiator DnaA-ATP." *Proc Natl Acad Sci U S A* **101**(9): 2811-2816.
- McPherson, A., Malkin, A. J. and Kuznetsov, Y. G. (1995). "The science of macromolecular crystallization." *Structure* **3**(8): 759-768.
- Megraud, F. (2003). "[When and how does Helicobacter pylori infection occur?]." *Gastroenterol Clin Biol* **27**(3 Pt 2): 374-379.
- Megraud, F. (2013). "Current recommendations for Helicobacter pylori therapies in a world of evolving resistance." *Gut Microbes* **4**(6): 541-548.
- Mesa, P., Alonso, J. C. and Ayora, S. (2006). "Bacillus subtilis bacteriophage SPP1 G40P helicase lacking the n-terminal domain unwinds DNA bidirectionally." *J Mol Biol* **357**(4): 1077-1088.
- Messer, W. (2002). "The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication." *FEMS Microbiol Rev* **26**(4): 355-374.
- Mitkova, A. V., Khopde, S. M. and Biswas, S. B. (2003). "Mechanism and stoichiometry of interaction of DnaG primase with DnaB helicase of Escherichia coli in RNA primer synthesis." *J Biol Chem* **278**(52): 52253-52261.
- Miyata, T., Suzuki, H., Oyama, T., Mayanagi, K., Ishino, Y. and Morikawa, K. (2005). "Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis." *Proc Natl Acad Sci U S A* **102**(39): 13795-13800.
- Molt, K. L., Sutura, V. A., Jr., Moore, K. K. and Lovett, S. T. (2009). "A role for nonessential domain II of initiator protein, DnaA, in replication control." *Genetics* **183**(1): 39-49.
- Mott, M. L., Erzberger, J. P., Coons, M. M. and Berger, J. M. (2008). "Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators." *Cell* **135**(4): 623-634.

Mulcair, M. D., Schaeffer, P. M., Oakley, A. J., Cross, H. F., Neylon, C., Hill, T. M. and Dixon, N. E. (2006). "A molecular mousetrap determines polarity of termination of DNA replication in *E. coli*." *Cell* **125**(7): 1309-1319.

Murray, H. and Errington, J. (2008). "Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA." *Cell* **135**(1): 74-84.

Natrajan, G., Noirot-Gros, M. F., Zawilak-Pawlik, A., Kapp, U. and Terradot, L. (2009). "The structure of a DnaA/HobA complex from *Helicobacter pylori* provides insight into regulation of DNA replication in bacteria." *Proc Natl Acad Sci U S A* **106**(50): 21115-21120.

Nitharwal, R. G., Paul, S., Dar, A., Choudhury, N. R., Soni, R. K., Prusty, D., Sinha, S., Kashav, T., Mukhopadhyay, G., Chaudhuri, T. K., Gourinath, S. and Dhar, S. K. (2007). "The domain structure of *Helicobacter pylori* DnaB helicase: the N-terminal domain can be dispensable for helicase activity whereas the extreme C-terminal region is essential for its function." *Nucleic Acids Res* **35**(9): 2861-2874.

Nitharwal, R. G., Verma, V., Dasgupta, S. and Dhar, S. K. (2011). "*Helicobacter pylori* chromosomal DNA replication: current status and future perspectives." *FEBS Lett* **585**(1): 7-17.

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**(12): 8342-8347.

Noirot-Gros, M. F., Velten, M., Yoshimura, M., McGovern, S., Morimoto, T., Ehrlich, S. D., Ogasawara, N., Polard, P. and Noirot, P. (2006). "Functional dissection of YabA, a negative regulator of DNA replication initiation in *Bacillus subtilis*." *Proc Natl Acad Sci U S A* **103**(7): 2368-2373.

Nunez-Ramirez, R., Robledo, Y., Mesa, P., Ayora, S., Alonso, J. C., Carazo, J. M. and Donate, L. E. (2006). "Quaternary polymorphism of replicative helicase G40P: structural mapping and domain rearrangement." *J Mol Biol* **357**(4): 1063-1076.

Oakley, A. J., Loscha, K. V., Schaeffer, P. M., Liepinsh, E., Pintacuda, G., Wilce, M. C., Otting, G. and Dixon, N. E. (2005). "Crystal and solution structures of the helicase-binding domain of *Escherichia coli* primase." *J Biol Chem* **280**(12): 11495-11504.

Ogura, Y., Ogasawara, N., Harry, E. J. and Moriya, S. (2003). "Increasing the ratio of Soj to Spo0J promotes replication initiation in *Bacillus subtilis*." *J Bacteriol* **185**(21): 6316-6324.

Ozaki, S. and Katayama, T. (2009). "DnaA structure, function, and dynamics in the initiation at the chromosomal origin." *Plasmid* **62**(2): 71-82.

Ozaki, S. and Katayama, T. (2012). "Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation." *Nucleic Acids Res* **40**(4): 1648-1665.

Pan, H. and Wigley, D. B. (2000). "Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase." *Structure* **8**(3): 231-239.

Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003). "Hexameric ring structure of the full-length archaeal MCM protein complex." *EMBO Rep* **4**(11): 1079-1083.

Pelikan, M., Hura, G. L. and Hammel, M. (2009). "Structure and flexibility within proteins as identified through small angle X-ray scattering." *Gen Physiol Biophys* **28**(2): 174-189.

Pomerantz, R. T. and O'Donnell, M. (2007). "Replisome mechanics: insights into a twin DNA polymerase machine." *Trends Microbiol* **15**(4): 156-164.

Reimer, L. (1997). *Transmission Electron Microscopy, Physics of Image Formation and Microanalysis*, Springer Series in Optical Sciences.

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P. and Diffley, J. F. (2009). "Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing." *Cell* **139**(4): 719-730.

Rice, P. A., Yang, S., Mizuuchi, K. and Nash, H. A. (1996). "Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn." *Cell* **87**(7): 1295-1306.

Robert, X. and Gouet, P. (2014). "Deciphering key features in protein structures with the new ENDscript server." *Nucleic Acids Res* **42**(Web Server issue): W320-324.

Robinson, A., Causer, R. J. and Dixon, N. E. (2012). "Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target." *Curr Drug Targets* **13**(3): 352-372.

- Robinson, A. and van Oijen, A. M. (2013). "Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies." *Nat Rev Microbiol* **11**(5): 303-315.
- 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**(1): 25-38.
- Roesler, B. M., Rabelo-Goncalves, E. M. and Zeitune, J. M. (2014). "Virulence Factors of *Helicobacter pylori*: A Review." *Clin Med Insights Gastroenterol* **7**: 9-17.
- Rossmann, M. G. and Blow, D. M. (1962). "The detection of sub-unit within the crystallographic asymmetric unit." *Acta. Cryst.* **15**: 24-31.
- Roth, A., Urmoneit, B. and Messer, W. (1994). "Functions of histone-like proteins in the initiation of DNA replication at *oriC* of *Escherichia coli*." *Biochimie* **76**(10-11): 917-923.
- Ryan, V. T., Grimwade, J. E., Nievera, C. J. and Leonard, A. C. (2002). "IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli oriC* by two different mechanisms." *Mol Microbiol* **46**(1): 113-124.
- Schaeffer, P. M., Headlam, M. J. and Dixon, N. E. (2005). "Protein--protein interactions in the eubacterial replisome." *IUBMB Life* **57**(1): 5-12.
- Schneidman-Duhovny, D., Hammel, M. and Sali, A. (2010). "FoXS: a web server for rapid computation and fitting of SAXS profiles." *Nucleic Acids Res* **38**(Web Server issue): W540-544.
- Schneidman-Duhovny, D., Hammel, M. and Sali, A. (2010). "FoXS: a web server for rapid computation and fitting of SAXS profiles." *Nucleic Acids Res* **38**(Web Server issue): W540-544.
- Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. and Sali, A. (2013). "Accurate SAXS profile computation and its assessment by contrast variation experiments." *Biophys J* **105**(4): 962-974.
- Seitz, H., Weigel, C. and Messer, W. (2000). "The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*." *Mol Microbiol* **37**(5): 1270-1279.

- Seybert, A., Singleton, M. R., Cook, N., Hall, D. R. and Wigley, D. B. (2006). "Communication between subunits within an archaeal clamp-loader complex." *EMBO J* **25**(10): 2209-2218.
- Shaikh, T. R., Gao, H., Baxter, W. T., Asturias, F. J., Boisset, N., Leith, A. and Frank, J. (2008). "SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs." *Nat Protoc* **3**(12): 1941-1974.
- Sharma, A., Nitharwal, R. G., Singh, B., Dar, A., Dasgupta, S. and Dhar, S. K. (2009). "Helicobacter pylori single-stranded DNA binding protein--functional characterization and modulation of H. pylori DnaB helicase activity." *FEBS J* **276**(2): 519-531.
- Simonetta, K. R., Kazmirski, S. L., Goedken, E. R., Cantor, A. J., Kelch, B. A., McNally, R., Seyedin, S. N., Makino, D. L., O'Donnell, M. and Kuriyan, J. (2009). "The mechanism of ATP-dependent primer-template recognition by a clamp loader complex." *Cell* **137**(4): 659-671.
- Singleton, M. R., Dillingham, M. S. and Wigley, D. B. (2007). "Structure and mechanism of helicases and nucleic acid translocases." *Annu Rev Biochem* **76**: 23-50.
- Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K. and Kleckner, N. (1995). "E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration." *Cell* **82**(6): 927-936.
- Soni, R. K., Mehra, P., Choudhury, N. R., Mukhopadhyay, G. and Dhar, S. K. (2003). "Functional characterization of Helicobacter pylori DnaB helicase." *Nucleic Acids Res* **31**(23): 6828-6840.
- Soni, R. K., Mehra, P., Mukhopadhyay, G. and Dhar, S. K. (2005). "Helicobacter pylori DnaB helicase can bypass Escherichia coli DnaC function in vivo." *Biochem J* **389**(Pt 2): 541-548.
- Soultanas, P. (2002). "A functional interaction between the putative primosomal protein DnaI and the main replicative DNA helicase DnaB in Bacillus." *Nucleic Acids Res* **30**(4): 966-974.
- Soultanas, P. (2005). "The bacterial helicase-primase interaction: a common structural/functional module." *Structure* **13**(6): 839-844.
- Soultanas, P. (2012). "Loading mechanisms of ring helicases at replication origins." *Mol Microbiol* **84**(1): 6-16.

- Sreekumar, K. R., Prack, A. E., Winters, D. R., Barbaro, B. A. and Bullock, P. A. (2000). "The simian virus 40 core origin contains two separate sequence modules that support T-antigen double-hexamer assembly." *J Virol* **74**(18): 8589-8600.
- Stelter, M., Gutsche, I., Kapp, U., Bazin, A., Bajic, G., Goret, G., Jamin, M., Timmins, J. and Terradot, L. (2012). "Architecture of a dodecameric bacterial replicative helicase." *Structure* **20**(3): 554-564.
- Stinchcomb, D. T., Struhl, K. and Davis, R. W. (1979). "Isolation and characterisation of a yeast chromosomal replicator." *Nature* **282**(5734): 39-43.
- Strycharska, M. S., Arias-Palomo, E., Lyubimov, A. Y., Erzberger, J. P., O'Shea, V. L., Bustamante, C. J. and Berger, J. M. (2013). "Nucleotide and partner-protein control of bacterial replicative helicase structure and function." *Mol Cell* **52**(6): 844-854.
- Suerbaum, S. and Josenhans, C. (2007). "Helicobacter pylori evolution and phenotypic diversification in a changing host." *Nat Rev Microbiol* **5**(6): 441-452.
- Suerbaum, S. and Michetti, P. (2002). "Helicobacter pylori infection." *N Engl J Med* **347**(15): 1175-1186.
- Su'etsugu, M., Harada, Y., Keyamura, K., Matsunaga, C., Kasho, K., Abe, Y., Ueda, T. and Katayama, T. (2013). "The DnaA N-terminal domain interacts with Hda to facilitate replicase clamp-mediated inactivation of DnaA." *Environ Microbiol* **15**(12): 3183-3195.
- Su'etsugu, M., Shimuta, T. R., Ishida, T., Kawakami, H. and Katayama, T. (2005). "Protein associations in DnaA-ATP hydrolysis mediated by the Hda-replicase clamp complex." *J Biol Chem* **280**(8): 6528-6536.
- Sutton, M. D., Carr, K. M., Vicente, M. and Kaguni, J. M. (1998). "Escherichia coli DnaA protein. The N-terminal domain and loading of DnaB helicase at the E. coli chromosomal origin." *J Biol Chem* **273**(51): 34255-34262.
- Sutton, M. D. and Kaguni, J. M. (1997). "The Escherichia coli dnaA gene: four functional domains." *J Mol Biol* **274**(4): 546-561.
- Svergun, D. I. (1992). "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." *Journal of Applied Crystallography* **25**(4): 495-503.

Svergun, D. I. and Koch, M. H. (2003). "Small-angle scattering studies of biological macromolecules in solution." *Reports on Progress in Physics* **66**(10): 1735.

Svergun, D. I. and Koch, M. H. (2013). *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*, Oxford University Press.

Terwilliger, T. C. (2000). "Maximum-likelihood density modification." *Acta Crystallogr D Biol Crystallogr* **56**(Pt 8): 965-972.

Thirlway, J., Turner, I. J., Gibson, C. T., Gardiner, L., Brady, K., Allen, S., Roberts, C. J. and Soutanas, P. (2004). "DnaG interacts with a linker region that joins the N- and C-domains of DnaB and induces the formation of 3-fold symmetric rings." *Nucleic Acids Res* **32**(10): 2977-2986.

Toth, E. A., Li, Y., Sawaya, M. R., Cheng, Y. and Ellenberger, T. (2003). "The crystal structure of the bifunctional primase-helicase of bacteriophage T7." *Mol Cell* **12**(5): 1113-1123.

Tougu, K., Peng, H. and Marians, K. J. (1994). "Identification of a domain of *Escherichia coli* primase required for functional interaction with the DnaB helicase at the replication fork." *J Biol Chem* **269**(6): 4675-4682.

Tsai, K. L., Lo, Y. H., Sun, Y. J. and Hsiao, C. D. (2009). "Molecular interplay between the replicative helicase DnaC and its loader protein DnaI from *Geobacillus kaustophilus*." *J Mol Biol* **393**(5): 1056-1069.

Valle, M., Chen, X. S., Donate, L. E., Fanning, E. and Carazo, J. M. (2006). "Structural basis for the cooperative assembly of large T antigen on the origin of replication." *J Mol Biol* **357**(4): 1295-1305.

Valle, M., Gruss, C., Halmer, L., Carazo, J. M. and Donate, L. E. (2000). "Large T-antigen double hexamers imaged at the simian virus 40 origin of replication." *Mol Cell Biol* **20**(1): 34-41.

van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. and Schatz, M. (1996). "A new generation of the IMAGIC image processing system." *J Struct Biol* **116**(1): 17-24.

Veening, J. W., Murray, H. and Errington, J. (2009). "A mechanism for cell cycle regulation of sporulation initiation in *Bacillus subtilis*." *Genes Dev* **23**(16): 1959-1970.

Velten, M., McGovern, S., Marsin, S., Ehrlich, S. D., Noirot, P. and Polard, P. (2003). "A two-protein strategy for the functional loading of a cellular replicative DNA helicase." *Mol Cell* **11**(4): 1009-1020.

Voet, D. and Voet, J. G. (2005). *Biochimie, De Boeck Supérieur*.

Wagner, J. K., Marquis, K. A. and Rudner, D. Z. (2009). "SirA enforces diploidy by inhibiting the replication initiator DnaA during spore formation in *Bacillus subtilis*." *Mol Microbiol* **73**(5): 963-974.

Wang, Q. P. and Kaguni, J. M. (1987). "Transcriptional repression of the *dnaA* gene of *Escherichia coli* by *dnaA* protein." *Mol Gen Genet* **209**(3): 518-525.

Warren, J. R. and Marshall, B. (1983). "Unidentified curved bacilli on gastric epithelium in active chronic gastritis." *Lancet* **1**(8336): 1273-1275.

Watari, J., Chen, N., Amenta, P. S., Fukui, H., Oshima, T., Tomita, T., Miwa, H., Lim, K. J. and Das, K. M. (2014). "*Helicobacter pylori* associated chronic gastritis, clinical syndromes, precancerous lesions, and pathogenesis of gastric cancer development." *World J Gastroenterol* **20**(18): 5461-5473.

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

Weigelt, J., Brown, S. E., Miles, C. S., Dixon, N. E. and Otting, G. (1999). "NMR structure of the N-terminal domain of *E. coli* DnaB helicase: implications for structure rearrangements in the helicase hexamer." *Structure* **7**(6): 681-690.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. and Wilson, K. S. (2011). "Overview of the CCP4 suite and current developments." *Acta Crystallogr D Biol Crystallogr* **67**(Pt 4): 235-242.

Wiseman, T., Williston, S., Brandts, J. F. and Lin, L.-N. (1989). "Rapid measurement of binding constants and heats of binding using a new titration calorimeter." *Analytical Biochemistry* **179**(1): 131-137.

Wlodawer, A. and Hodgson, K. O. (1975). "Crystallization and drystal data for monellin." *Proc. Nat. Acad. Sc.* **72**: 398-399.

Wyatt, P. J. (1993). "Light scattering and the absolute characterization of macromolecules." *Analytica Chimica Acta* **272**(1): 1-40.

Xu, Q., McMullan, D., Abdubek, P., Astakhova, T., Carlton, D., Chen, C., Chiu, H. J., Clayton, T., Das, D., Deller, M. C., Duan, L., Elsliger, M. A., Feuerhelm, J., Hale, J., Han, G. W., Jaroszewski, L., Jin, K. K., Johnson, H. A., Klock, H. E., Knuth, M. W., Kozbial, P., Sri Krishna, S., Kumar, A., Marciano, D., Miller, M. D., Morse, A. T., Nigoghossian, E., Nopakun, A., Okach, L., Oommachen, S., Paulsen, J., Puckett, C., Reyes, R., Rife, C. L., Sefcovic, N., Trame, C., van den Bedem, H., Weekes, D., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. and Wilson, I. A. (2009). "A structural basis for the regulatory inactivation of DnaA." *J Mol Biol* **385**(2): 368-380.

Yang, S., Yu, X., VanLoock, M. S., Jezewska, M. J., Bujalowski, W. and Egelman, E. H. (2002). "Flexibility of the rings: structural asymmetry in the DnaB hexameric helicase." *J Mol Biol* **321**(5): 839-849.

Yuzhakov, A., Kelman, Z. and O'Donnell, M. (1999). "Trading places on DNA--a three-point switch underlies primer handoff from primase to the replicative DNA polymerase." *Cell* **96**(1): 153-163.

Zawilak-Pawlik, A., Donczew, R., Szafranski, S., Mackiewicz, P., Terradot, L. and Zakrzewska-Czerwinska, J. (2011). "DiaA/HobA and DnaA: a pair of proteins co-evolved to cooperate during bacterial orisome assembly." *J Mol Biol* **408**(2): 238-251.

Zawilak-Pawlik, A., Kois, A., Majka, J., Jakimowicz, D., Smulczyk-Krawczynszyn, A., Messer, W. and Zakrzewska-Czerwinska, J. (2005). "Architecture of bacterial replication initiation complexes: orisomes from four unrelated bacteria." *Biochem J* **389**(Pt 2): 471-481.

Zawilak-Pawlik, A., Kois, A., Stingl, K., Boneca, I. G., Skrobuk, P., Piotr, J., Lurz, R., Zakrzewska-Czerwinska, J. and Labigne, A. (2007). "HobA--a novel protein involved in initiation of chromosomal replication in *Helicobacter pylori*." *Mol Microbiol* **65**(4): 979-994.

Article 1

Architecture of a Dodecameric Bacterial Replicative Helicase

Stelter, M., Gutsche, I., Kapp, U., **Bazin, A.**, Bajic, G., Goret, G., Jamin, M., Timmins, J. and Terradot, L. (2012). "Architecture of a dodecameric bacterial replicative helicase." *Structure* **20**(3): 554-564.

Architecture of a Dodecameric Bacterial Replicative Helicase

Meike Stelter,^{1,5,6} Irina Gutsche,^{2,5} Ulrike Kapp,¹ Alexandre Bazin,³ Goran Bajic,³ Gaël Goret,² Marc Jamin,² Joanna Timmins,^{1,4,6} and Laurent Terradot^{1,3,*}

¹Structural Biology Group, European Synchrotron Radiation Facility, BP 220 38043 Grenoble Cedex 9, France

²Unit of Virus Host-Cell Interactions, UJF-EMBL-CNRS, UMI3265, 38042 Grenoble Cedex 9, France

³Institut de Biologie et Chimie des Protéines, Université Lyon 1, Université Lyon, France, CNRS, UMR 5086, IBCP 7 Passage du Vercors, 69367 Lyon, France

⁴Institut de Biologie Structurale J.P. Ebel, 41, Rue Jules Horowitz, 38027 Grenoble Cedex, France

⁵These authors contributed equally to this work

⁶Present address: Institut de Biologie Structurale J.P. Ebel, 41, Rue Jules Horowitz, 38027 Grenoble Cedex, France

*Correspondence: laurent.terradot@ibcp.fr

DOI 10.1016/j.str.2012.01.020

SUMMARY

Hexameric DnaB helicases are often loaded at DNA replication forks by interacting with the initiator protein DnaA and/or a helicase loader (DnaC in *Escherichia coli*). These loaders are not universally required, and DnaB from *Helicobacter pylori* was found to bypass DnaC when expressed in *E. coli* cells. The crystal structure of *Helicobacter pylori* DnaB C-terminal domain (HpDnaB-CTD) reveals a large two-helix insertion (named HPI) in the ATPase domain that protrudes away from the RecA fold. Biophysical characterization and electron microscopy (EM) analysis of the full-length protein show that HpDnaB forms head-to-head double hexamers remarkably similar to helicases found in some eukaryotes, archaea, and viruses. The docking of the HpDnaB-CTD structure into EM reconstruction of HpDnaB provides a model that shows how hexamerization of the CTD is facilitated by HPI-HPI interactions. The HpDnaB double-hexamer architecture supports an alternative strategy to load bacterial helicases onto forks in the absence of helicase loaders.

INTRODUCTION

Replicative helicases are essential enzymes that participate in genome duplication in all organisms. Homo- or hetero-oligomers of these helicases are recruited to active replication forks by interacting with the replication initiators and/or loaders. Once loaded, DNA helicases unwind dsDNA by using ATP hydrolysis and serve as a platform to target and load other replication enzymes at the open DNA forks. In bacteria, DnaB (named DnaC in Gram-positive bacteria, not to be confused with *Escherichia coli* helicase loader DnaC) is the main replicative helicase and is arranged as ring-shaped hexamers. These proteins contain two domains: an N-terminal domain (NTD), and a C-terminal domain (CTD). The NTD is divided into a globular

six-helix “head” (N-glob) followed by two helices forming an α -hairpin (Bailey et al., 2007a). The NTD is structurally and functionally homologous to the helicase-interacting CTD of bacterial primases (Chintakayala et al., 2007; Oakley et al., 2005; Syson et al., 2005). In *E. coli* the NTD of DnaB is dispensable for DNA binding and ATPase activity but is important for helicase activity (Biswas and Biswas, 1999a, 1999b). The CTD has a RecA-like fold typical of the RecA/F subgroup of the additional strand catalytic glutamate (ASCE) superfamily of P loop ATPases (Bailey et al., 2007a). As all members of the RecA/F group, DnaB CTDs exhibit five characteristic motifs: H1 (corresponds to Walker A), H1a and H2 (corresponds to Walker B), and H3 and H4 (Ilyina et al., 1992). The NTD and CTD are connected by a flexible linker containing a helix that mediates hexamerization by binding to a conserved pocket in the CTD of an adjacent subunit (Bailey et al., 2007a, 2007b).

The ring-shaped hexamers of DnaB are formed by two rings, one made by the NTDs (NTD-ring) and one by the CTDs (CTD-ring). Different structural arrangements of these rings have been observed, illustrating the dynamics of the assembly. Electron microscopy (EM) studies of *E. coli* DnaB (EcDnaB) and of the DnaC-like protein G40P from *Bacillus subtilis* SSP1 bacteriophage indicated that hexamers can have either a 6-fold rotational symmetry (C_6), 3-fold rotational symmetry (C_3), or an intermediate state (C_3C_6) (Núñez-Ramírez et al., 2006; Yang et al., 2002). Crystal structures of hexameric DnaBs have been described for G40P, *Geobacillus stearothermophilus* (GsDnaB), and *G. kaustophilus* (GkDnaB) (Bailey et al., 2007b; Lo et al., 2009; Wang et al., 2008). In these structures, DnaB adopts a similar organization where DnaB subunits display two different conformations, which differ in the position of the NTD relative to the CTD. *cis* and *trans* subunits assemble in alternating arrangement where the NTD-ring can be described as a trimer of dimers (C_3), whereas the CTD-ring has a 6- or near 6-fold symmetry (Bailey et al., 2007a; Lo et al., 2009; Wang et al., 2008). Atomic force microscopy and crystallographic studies have shown that a DnaB hexamer can accommodate up to three DnaG moieties that bind at the NTD dimer interface (Bailey et al., 2007b; Thirlway et al., 2004). DnaBs are thought to bind DNA via both the CTD and the NTD with one strand threaded within the ring and the other passing outside. Based on the structure of the

Structure

Structure of *Helicobacter pylori* DnaB

T7 GP4 helicase, four DNA binding motifs were identified in the CTD: H4, loop I, II and III with loops I and II, and H4 structurally conserved in the RecA/F family (Singleton et al., 2000). However, the structure of GkDnaB in complex with a single-stranded (dT15) oligonucleotide showed that ssDNA bound primarily to the NTD and to two arginines from the loop I of the CTD (Lo et al., 2009).

To accomplish their functions during replication, DnaB hexamers must be loaded in the correct orientation at each replication fork, i.e., with the CTD-ring facing the dsDNA, whereas ssDNA is passing through the hexameric channel. In some (but not all) Gram-negative bacteria such as *E. coli*, DnaB recruitment requires helicase loaders such as DnaC (Davey and O'Donnell, 2003). In these bacteria two DnaB hexamers are loaded as DnaB₆DnaC₆ complexes at active replication forks formed by the origin of replication *oriC* and multimers of the DNA initiator protein DnaA (reviewed in Katayama et al., 2010). One DnaB hexamer can be loaded on ssDNA via interaction with the NTD of DnaA (Seitz et al., 2000; Sutton et al., 1998). The other hexamer might be loaded in the opposite direction by interaction with DnaC, acting as a molecular adaptor by simultaneously interacting with the ATP-bound DnaA, as suggested for *Aquifex aeolicus* (Mott et al., 2008). However, no interaction has been detected between *E. coli* DnaA and DnaC (Keyamura et al., 2009; Schaeffer et al., 2005). Instead, DnaC was proposed to bind and to open up the helicase ring, which is then loaded onto the DNA via an interaction with DnaA (Ozaki and Katayama, 2011). In the Gram-positive *B. subtilis*, three different proteins, DnaD, DnaB, and DnaI, are required to load the replicative helicase at *oriC*. DnaI is the helicase loader and is a sequence homolog of DnaC (Ioannou et al., 2006). DnaD (not to be confused with the EcDnaB helicase) and DnaB assemble onto the replication origin for DnaI to load the helicase DnaC (Smits et al., 2010; Velten et al., 2003).

Although helicase loaders have been identified in *E. coli* and in *B. subtilis*, they are absent in most bacterial species. For instance unlike the EcDnaB helicase, the closely related DnaB from *Pseudomonas aeruginosa* could be delivered and activated at an origin of replication in the absence of DnaC (Caspi et al., 2001). As those of most bacteria, the genome of the human pathogen *Helicobacter pylori* does not encode for a DnaC homolog (reviewed in Nitharwal et al., 2011). Genetic complementation experiments have shown that *H. pylori* DnaB (HpDnaB) can complement *E. coli* strain J58, a conditional lethal mutant for the EcDnaB allele. More intriguingly, it is also able to complement two different *E. coli* dnaCts mutants when expressed in *trans*, suggesting that the protein is able to bypass DnaC in these cells (Soni et al., 2005). In contrast to other DnaBs, the N-glob of HpDnaB is dispensable for helicase activity in vitro, and the corresponding mutated *hpdnaB* is able to complement an *E. coli* dnaBts strain in vivo (Nitharwal et al., 2007). To better understand the properties of HpDnaB, we have solved the structure of its CTD (HpDnaB-CTD) at 2.5 Å resolution. The structure exhibits a conserved fold but also a unique insertion that adopts a helix-turn-helix motif that protrudes away from the core domain. Biophysical characterization of HpDnaB in solution unexpectedly revealed that the protein forms assemblies larger than hexamers. EM analysis and single-particle reconstruction demonstrated that HpDnaB forms head-to-head double-hex-

amer structures similar to MCM and Mcm2-7 helicases. The structure of HpDnaB-CTD was fitted into the 23 Å resolution EM density map, which revealed an organization different from those observed previously in DnaB crystal structures (Kashav et al., 2009). To our knowledge, our work describes the first example of a dodecameric bacterial helicase, which may explain the peculiar properties of the DnaB helicase of *H. pylori* and also suggests an alternative mechanism for loading replicative helicases in the absence of helicase loaders.

RESULTS

Crystal Structure of HpDnaB-CTD

The crystal structure of a C-terminal fragment of HpDnaB, corresponding to residues 151–488, was solved by the single-wavelength anomalous dispersion method using Se-Met substituted protein. HpDnaB-CTD crystallized in the space group P2₁2₁2 with two molecules per asymmetric unit (AU) (Table 1). Chain A consists of residues 151–473 and chain B of residues 151–317 and 330–468. Refinement statistics are shown in Table 1. As expected, HpDnaB-CTD adopts a RecA-like fold (Story et al., 1992), consisting of a bent, predominantly parallel β sheet flanked by α helices similar to other DnaB helicases (Figure 1B). Part of the linker region (residues 151–175) is attached to the CTD via a flexible loop and contains the so-called linker helix (Figures 1A and 1B; see Figure S1 available online).

The five motifs defining the DnaB family of helicases (Ilyina et al., 1992) (H1, H1a, H2, H3, and H4) are present in the HpDnaB-CTD crystal structure (Figures 1B and S1). The Walker A and Walker B motifs (H1 and H2) are conserved, whereas the DNA binding elements of HpDnaB differ significantly compared to other DnaB crystal structures. Loop I (residues 314–330) that binds DNA in the GkDnaB-ssDNA complex (Lo et al., 2009) adopts a very different conformation than in other DnaBs, i.e., a short β strand followed by a short helix (Figure 1C). The neighboring loop (residues 371–382, Figure 1C) designated as the H4 motif shows a β-hairpin conformation similar to the T7 GP4 helicase (Singleton et al., 2000) but different from other DnaBs (Figure 1C). The most remarkable feature of HpDnaB is that the unique insertion of 34 residues in the ATPase domain forms a helix-turn-helix motif (named hereafter HPI for *Helicobacter pylori* insertion) that protrudes away from the core ATPase domain (Figure 1B). This feature replaces loop III, which was proposed to bind DNA in the T7 GP4 helicase (Singleton et al., 2000). The two helices of the HPI are heavily charged with ten positively charged and nine negatively charged residues (Figures 1D and S1). These residues are evenly distributed in the protein sequence, but not in the helices. One side of the HPI is clearly negatively charged, whereas the other one is positively charged, in particular near the core of the CTD, and forms a large positive patch together with the nucleotide binding pocket and loop I (Figure 1D). A search for homologous structures in the protein database (DALI) retrieved a wealth of protein structures, many of which are involved in DNA replication. A particularly significant hit was the C-terminal helical hairpin of *B. subtilis* UvrB, which is believed to be responsible for protein-protein interactions and dimerization (Waters et al., 2006).

The AU contains a HpDnaB-CTD dimer formed by interactions between the β sheet, the HPI and the linker helix, with a buried

surface of 1,820 Å² (Figure S2). No hexameric arrangement could be distinguished. This dimer does not seem to be biologically relevant because the two molecules interact predominantly via the HPI in a head-to-head manner (Figure S2). In solution, HpDnaB-CTD does not form hexamers but is predominantly a tetramer with lower molecular weight (MW) complexes also observed (see below). However, the linker helix inserts into the hexamerization pocket of a symmetry-related CTD (Figure S1), as previously found in the crystal structure of *Thermophilus aquaticus* DnaB (Bailey et al., 2007a). The orientation of the HpDnaB-CTD is completely different from the ones described in hexameric DnaB structures, but the binding mechanism is very similar (Figure 1E). This suggests that the helix-mediated hexamerization found in DnaBs and other helicases is most likely conserved in HpDnaB.

Oligomeric State of HpDnaB and 3D Reconstruction by EM

To gain insight into the structure of full-length HpDnaB, we first examined the oligomeric state of the protein by size exclusion chromatography (SEC) coupled to multi-angle laser light scattering (MALLS; Figure 2A). HpDnaB elutes as a single peak as described previously (Soni et al., 2003), but unexpectedly, SEC-MALLS revealed that HpDnaB has a molecular mass of 680 kDa, which is much larger than that expected for the hexamer (~340 kDa). The size of the assembly corresponds to a dodecamer (theoretical mass of 684 kDa), and no significant variation was observed when nucleotides (ADP, ATP, or the nonhydrolysable ATP analog AMP-PNP) were added to the protein and when the SEC buffers were changed (Figure 2A). In contrast, HpDnaB-CTD forms various oligomers predominantly tetramers but also lower MW assemblies (Figure 2B). In solution the protein showed an ATPase activity of 0.069 and 0.279 μM ATP/s in the presence of 50-mer ssDNA. Although lower, these rates are in agreement with rates observed in an independent study (Nitharwal et al., 2007) (Table S1). In the presence of the long ssM13mp18 DNA, HpDnaB activity did not increase, suggesting that long ssDNA had little effect on its activity. HpDnaB-CTD has no ATPase activity (Table S1).

HpDnaB was imaged by negative-stain EM (Figure 3A), and approximately 35,000 individual particle frames were selected for image analysis (see Supplemental Experimental Procedures for details). The eigenimages of the centered rotationally unaligned particles were used as an indicator of the data set symmetry, and strongly suggested the presence of a C₆ on the top views and a 2-fold rotational symmetry on the side views (Figure 3B). No evidence for a C₃ could be found from the eigenimages, even if a careful visual inspection showed a minor amount of apparently 3-fold symmetric rings (Figure 3A, upper row, compare the left 6-fold symmetric and second from the left 3-fold symmetric class averages). Remarkably, in addition to top views of a single hexameric ring, many class averages could be unambiguously identified as side views and intermediate views of a double ring (Figure 3A, lower row), strikingly similar to those observed in EM studies of double-ring replicative helicases such as *Methanobacterium thermoautotrophicum* MCM (MtMCM) (Costa et al., 2006b) and the yeast Mcm2-7 helicases (Evrin et al., 2009; Remus et al., 2009). These large

Table 1. Data Collection Statistics

	Native Crystal	Selenomethionine Derivative
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Cell parameters, a, b, c (Å)	100.9, 102.2, 85.1	102.1, 102.1, 85.9
Resolution range (Å)	45.6–2.5 (2.64–2.50)	43.9–2.8 (2.95–2.80)
R _{merge}	0.059 (0.586) ^a	0.064 (0.340)
R _{ano}		0.078 (0.219)
Number of observed reflections	150,742 (22,036)	116,868 (15,063)
Number of unique reflections	30,696 (4,447)	22,621 (3,134)
Mean (I/σ(I))	16.2 (2.5)	15.5 (3.7)
Completeness (%)	99.0 (99.8)	99.4 (96.2)
Multiplicity	4.9 (5.0)	5.2 (4.8)
Phasing Statistics		
Number of sites	20	
Phasing power	1.8 (2.6 at 4 Å)	
FOM	0.37 (0.5 at 4 Å)	
FOMdm (53% solvent)	0.85	
Refinement Statistics		
Resolution range for refinement (Å)	40.2–2.5	
Number of unique reflections in refinement	30,672	
Working set	29,125	
Test set	1,547	
Completeness (%)	98.6	
R factor (%)	21.7	
Free R factor (%)	25.1	
Weighted rmsds from ideal values		
Bond length (Å)	0.005	
Angles (°)	0.875	
B wilson (Å ²)	57.4	
Average B factor (Å ²)	83.3	
Main chain/side chain plus waters	82.3/84.4	

^aValues in parentheses refer to the indicated resolution shell.

particles correspond in fact to two hexamers stacked together (Figures 3A and 3B).

Despite quantitative variations between purification fractions, both hexameric and double-hexameric views were systematically observed in all preparations, suggesting that HpDnaB dodecamers can dissociate during sample preparation for EM analysis. Thus, we considered a mixed population of single rings (as previously observed by Nitharwal et al., 2007) and double rings of DnaB, and processed the data up to 3D reconstructions of these two assemblies (Figure S3; Supplemental Experimental Procedures). The 21 Å resolution reconstruction of the single hexamer was practically indistinguishable from the individual

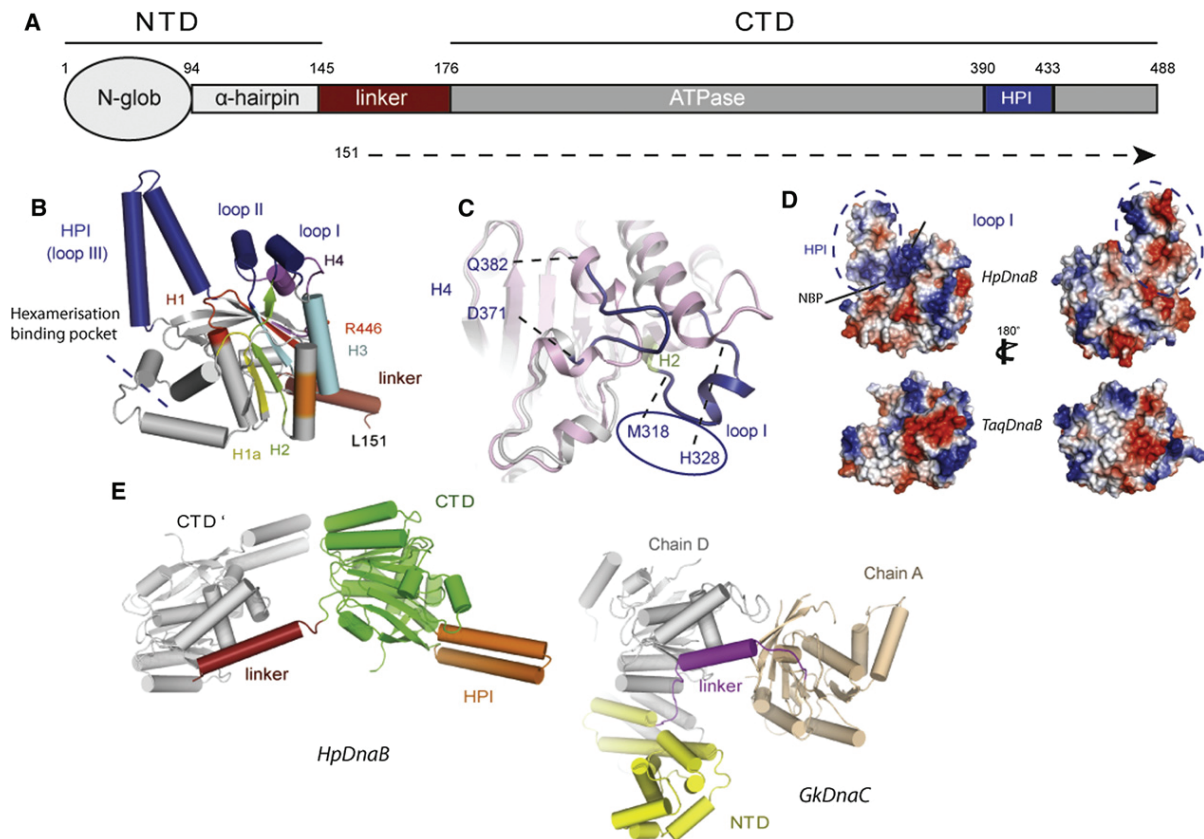


Figure 1. Crystal Structure of HpDnaB-CTD

(A) Schematic representation of the domain organization of HpDnaB. The NTD is composed of a globular domain (N-glob) followed by a helical hairpin. The CTD contains the ATPase domain in which the HPI is inserted. The linker region contains the helix that mediates the hexamerization of the helicase. HpDnaB amino acid numbers are indicated at the domain boundaries. An arrow with a dashed line indicates the region of HpDnaB that was crystallized in our study.

(B) Crystal structure of HpDnaB-CTD displayed as ribbon and cylinders. The HPI replaces loop III in other DnaBs and is colored blue. Motifs characteristic of the RecA/F family are indicated: H1 or Walker A is colored in red, H1a in yellow, H2 or Walker B in green, and H4 in magenta. The DNA binding loops I and II are also colored in blue. The motif RARR found in TaqDnaB (Bailey et al., 2007a) is colored in orange.

(C) Close-up view of a superimposition of HpDnaB-CTD (gray) with the GP40 crystal structure (pink). Two major differences (colored in blue in HpDnaB-CTD) can be noted in the P loop containing the Walker B motif and in the H4 motif.

(D) Surface representation of HpDnaB-CTD and comparison with the *Thermophilus aquaticus* DnaB-CTD (TaqDnaB) colored according to electrostatic potential. Positively charged regions are colored blue, whereas negatively charged regions are colored red.

(E) Left view is a cartoon representation of two symmetry-related HpDnaB-CTD monomers showing the insertion of the linker helix (red) into the hexamerization binding pocket of the neighboring subunit. Right view is a similar representation of a *Geobacillus kaustophilus* DnaC (GkDnaC) dimer from the crystal structure of the GkDnaC hexamer.

See also Figure S1.

hexamers in the dodecameric state, which in turn was reconstructed to 23 Å resolution. The HpDnaB hexamer is composed of two layers that can be readily interpreted as the NTD- and the CTD-rings (Figures 3C and 3D). As in other double-hexameric helicases, such as MtMCM, Mcm2-7, or the viral E1 and large T antigen (LTAg) helicases, the two hexamers of HpDnaB interact via their NTD-rings. Compared to these helicases, the HpDnaB double ring is shorter, with a height of about 150 Å (Figure 3C).

The hexameric ring with a diameter of about 130 Å and a height of 70 Å can at this resolution be considered to possess 6-fold symmetry (Figure 3D). The structure is organized around a central cavity, which is about 55 Å wide in the middle of the CTD-ring and the NTD-ring, but tightens up to approximately 35 Å at the

distal end of the CTD-ring partly occluded by a collar. This hexameric organization is reminiscent of the EM structure of EcDnaB but different from the crystal structures of DnaBs (Figure 3D). Both the CTD- and NTD-rings exhibit a rotational 6-fold symmetry. In DnaB crystal structures the CTD-ring showed a C_6 or slightly modified symmetry, whereas the NTD-ring had a clear C_3 symmetry (Figure 3D). Although a C_6 hexamer appears to best fit the HpDnaB single-particle data for both single and double hexamers (see Supplemental Experimental Procedures), we cannot rule out that a mixed C_3C_6 state (as seen in the crystal structures of DnaBs) is present. This uncertainty is due to both the resolution of the data and to the fact that the NTDs, which would be the region likely to adopt a trimer-of-dimers state, have less scattering mass than the CTDs, which are almost

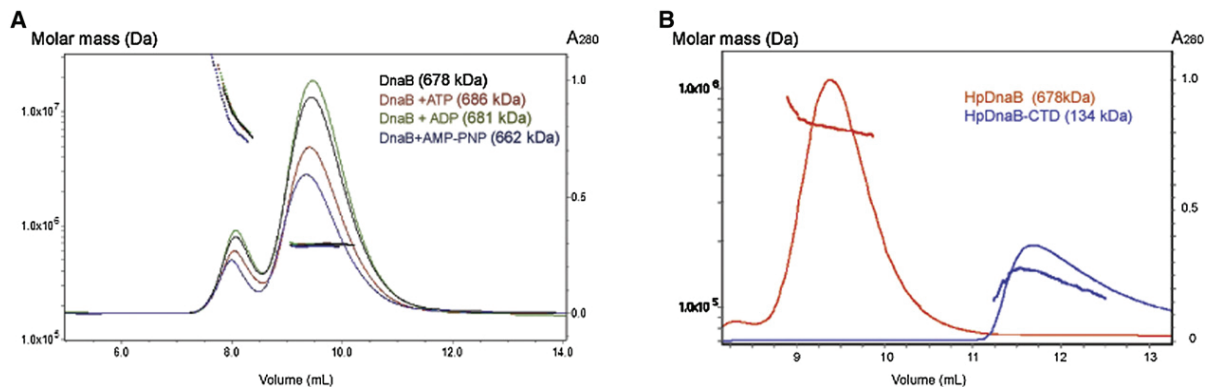


Figure 2. HpDnaB Oligomerization

(A) Chromatograms obtained from SEC-MALLS measurements with HpDnaB (black), HpDnaB with ATP (red), ADP (green), or AMP-PNP (blue) in SEC buffer for HpDnaB (20 mM Tris [pH 8], 200 mM sodium chloride, and 5% glycerol) and supplemented with 5 mM $MgCl_2$ and the corresponding nucleotide. Molar mass (Da) points obtained are indicated above the curves for the first peak and below for the second. The first peak corresponds to HpDnaB aggregation (heterogeneous peak containing very high MW assemblies), whereas the second peak has a mass of around 680 kDa, corresponding to HpDnaB dodecamers. (B) Comparison of the chromatograms obtained with HpDnaB (red curve) and HpDnaB-CTD (blue curve) with the calculated molar masses from MALLS. HpDnaB-CTD is around 38 kDa. The molar mass calculated by MALLS is 134 kDa. Thus, the calculated mass is closer to a tetramer of HpDnaB-CTD (152 kDa). See also Figure S2 and Table S1.

certainly in a C_6 -like state. Given that all NTD structures, which comprise both the helical hairpin and globular domain, form dimers through the α -helical hairpins, these contacts may well be preserved in HpDnaB.

In our EM reconstructions the CTD-ring of HpDnaB has a 6-fold symmetry with a propeller shape in which each ATPase domain is clearly distinguishable, separated by a crevasse of around 10 Å. These large side openings have been previously observed in structures of EcDnaB (Yang et al., 2002), LTag (Li et al., 2003), and MtMCM (Gómez-Llorente et al., 2005) (Figure 3D). The CTD-ring does not make extensive contacts with the NTD-ring, suggesting that the CTDs could undergo concerted conformational changes during DnaB activity. On the top of the CTD-ring, a triangular-shaped density protrudes from each subunit and forms a collar that closes the hexamer into a 35 Å wide hole, which is smaller than in DnaB crystal structures and EcDnaB EM studies (Figure 3D). This unusual feature is reminiscent of the cap found in the MtMCM EM reconstruction (Bae et al., 2009).

A Model of the HpDnaB Double Hexamer

A model for the hexameric HpDnaB-CTD was then constructed by docking the crystal structure into the EM map using Chimera (Pettersen et al., 2004) and VEDA (<http://mem.ibs.fr/VEDA>), the new graphic version of URO (Navaza et al., 2002) (Figure 4). The HpDnaB-CTDs fit remarkably well within the density (Figure 4A). Parts of the crystal structure penetrate the density, suggesting that structural rearrangements occur upon hexamerization, in particular for the DNA binding loop I and the HPI.

In this model the HPIs of adjacent subunits are located at the collar, at the top of the CTD-ring (Figure 4A). Although the two helices of the HPI are positioned to interact with each other, they are too distant to do so in the present model (Figure 4A). Because part of the HPI does not fit completely into the EM

density, the conformation of HPIs is likely different in the context of the hexamer than in the crystal structure. Indeed, the conformation of the HPI in the HpDnaB-CTD crystal structure is considerably constrained by the crystal packing because the HPI mediates the dimer formation in the AU (Figure S2). This is also supported by normal mode analysis of HpDnaB-CTD that indicated that the HPI is highly flexible and could undergo large conformational changes (Figure S5; Supplemental Experimental Procedures). These observations strongly suggest that the collar is formed by HPI-HPI interactions, which largely contribute to the stability of the CTD-ring assembly. Accordingly, mutagenesis and biochemical studies showed that when the HPI is deleted, HpDnaB does not form oligomers and displays no ATPase or helicase activity (Nitharwal et al., 2007).

Compared with hexameric DnaB crystal structures, each HpDnaB-CTD undergoes a rotation of around 30° (Figures 4B and 4C). The HPIs form a continuum of DNA binding motifs with loops I and II in the inner channel of the hexamer. The charge is more pronounced on one side of each subunit and provides a large positive area all along the central channel, starting from the collar to the base of the RecA fold. Interestingly, the H4 motif lies at the interface between two subunits and is solvent accessible. This contrasts with the previous structures of DnaBs where H4 is mostly buried between two subunits and is only accessible from inside the ring (Figure 4D). It should be noted that a separation between DnaB CTDs is seen between some subunits in the crystal structures of GsDnaB (Bailey et al., 2007b). In HpDnaB the side channels are large and provide access from the outer side of the ring to the H4 motif and the DNA binding loops I and II (Figure 4D).

We then attempted to dock the NTD of HpDnaB (HpDnaB-NTD) into the EM map to obtain a complete model. The crystal structure of HpDnaB-NTD (residues 1–122) consists of a N-terminal globular domain (N-glob) followed by a long α helix forming

Structure

Structure of *Helicobacter pylori* DnaB

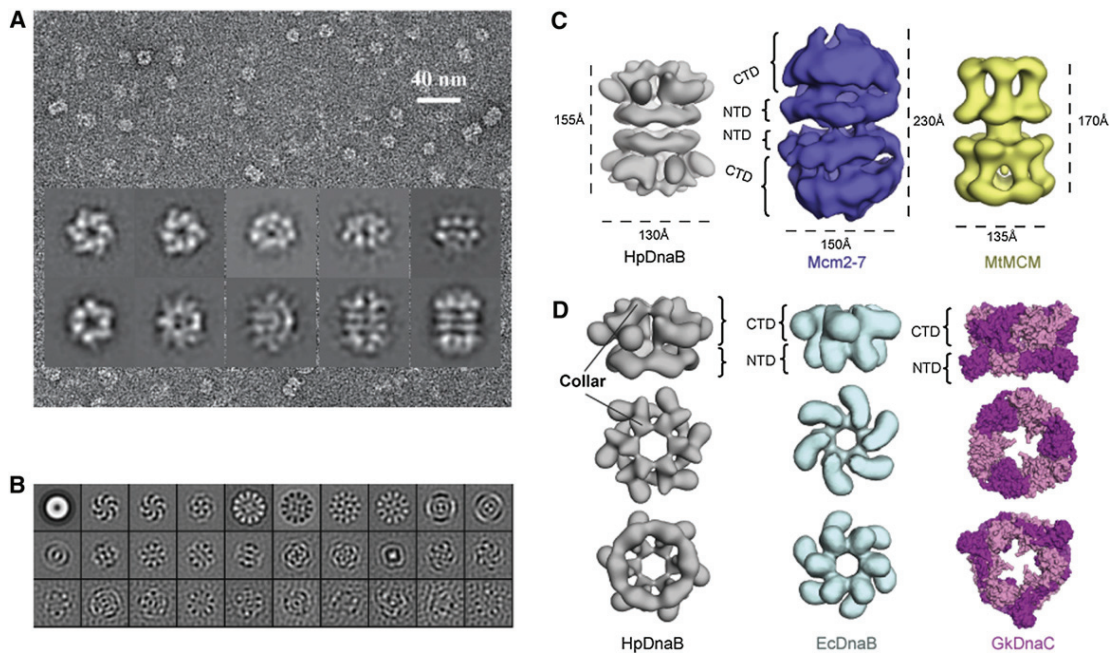


Figure 3. EM Reconstruction of HpDnaB Double Hexamers

(A) EM image of HpDnaB and five representative classes of particles corresponding to hexamers (top panels) and dodecamers (bottom panels). (B) First eigenimages of the rotationally unaligned HpDnaB data set. (C) EM density envelope from the reconstruction of HpDnaB double hexamer (gray) and comparison with Mcm2-7 (blue) and MtMCM (yellow) reconstructions. (D) Comparison of HpDnaB EM hexamer (gray) with EcDnaB EM reconstruction (blue) in the C_6 conformation and the crystal structure of GkDnaC (adjacent subunits are colored in pink and magenta). Hexamers are displayed in the same orientation: viewed from the side (top), CTD (middle), and NTD (bottom). See also Figures S3 and S4.

the first half of the α -hairpin (Sharma et al., 2009) (Figures 1 and S1). In the crystal structures of hexameric DnaBs, the NTD can adopt two conformations, a *cis* or a *trans*, relative to the CTD (Figure S6). These two conformations result in a NTD-ring with a C_3 symmetry, a conformation suitable to bind DnaG-CTDs (Bailey et al., 2007b; Biswas and Tsodikov, 2008; Wang et al., 2008). However, in our EM reconstructions of HpDnaB dodecamer and hexamer, the NTD-ring had a C_6 . We generated a monomer of HpDnaB (HpDnaB-NTD and HpDnaB-CTD, total of 421 residues out of 488) in *cis* and *trans* conformations and fitted them into the EM map. The best fit was obtained with the HpDnaB in the *cis* conformation. The docking of HpDnaB-NTD was improved using Chimera (Pettersen et al., 2004), and the positioning of the six subunits was optimized using VEDA (Figure 4E). The resulting model of the dodecamer consists of two identical hexamers of HpDnaB with six NTDs in a conformation similar to the *cis* conformation (Figures 4E and S6). It should be noted, however, that when HpDnaB-NTD was fitted independently of HpDnaB-CTD, several orientations were obtained, in particular, the NTD could be rotated 180° about an axis perpendicular to the 6-fold symmetry axis. The resulting position of the NTD is in fact very similar to the one obtained in the EM model of *E. coli* DnaB (Yang et al., 2002). Thus, although the density indicates where the globular domain and the helical hairpin are likely to reside such as in Figure 4E, a more precise determination of the NTD position would require a higher-resolution EM reconstruction.

DISCUSSION

DnaBs are essential helicases that are loaded onto active replication forks. Until our study, to our knowledge, DnaBs were found exclusively as hexamers in solution. Here, we presented the crystal structure of the CTD of DnaB from *H. pylori* and the architecture of the full-length protein based on EM reconstructions. Our work clearly establishes that HpDnaB forms stable dodecamers in solution that are organized as head-to-head double hexamers. Interestingly, it was proposed earlier that EcDnaB could form head-to-head double hexamers, a model described as the DNA twin pump system (Kaplan and O'Donnell, 2004). Dodecamers of an EcDnaB mutant (D82N) have also been identified by electrospray ionization mass spectrometry (Watt et al., 2007), but to our knowledge, no structural study has been reported to support a dodecameric model for DnaB.

The dodecameric architecture we describe is strikingly similar to those of some replicative helicases found in viruses (LTag, E1), archaea (MtMCM), and in yeast (Mcm2-7). Sv40 LTag and E1 assemble from monomers on replication origins to adopt a double-hexamer state (Schuck and Stenlund, 2005; Valle et al., 2000). Mcm2-7 and MtMCM have been shown to form a double hexamer on origin DNA or on short DNA duplexes, respectively (Costa et al., 2006a; Evrin et al., 2009; Remus et al., 2009). HpDnaB and these proteins share a similar organization consisting of a head-to-head double hexamer mediated by interactions between NTDs. The NTDs of these proteins,

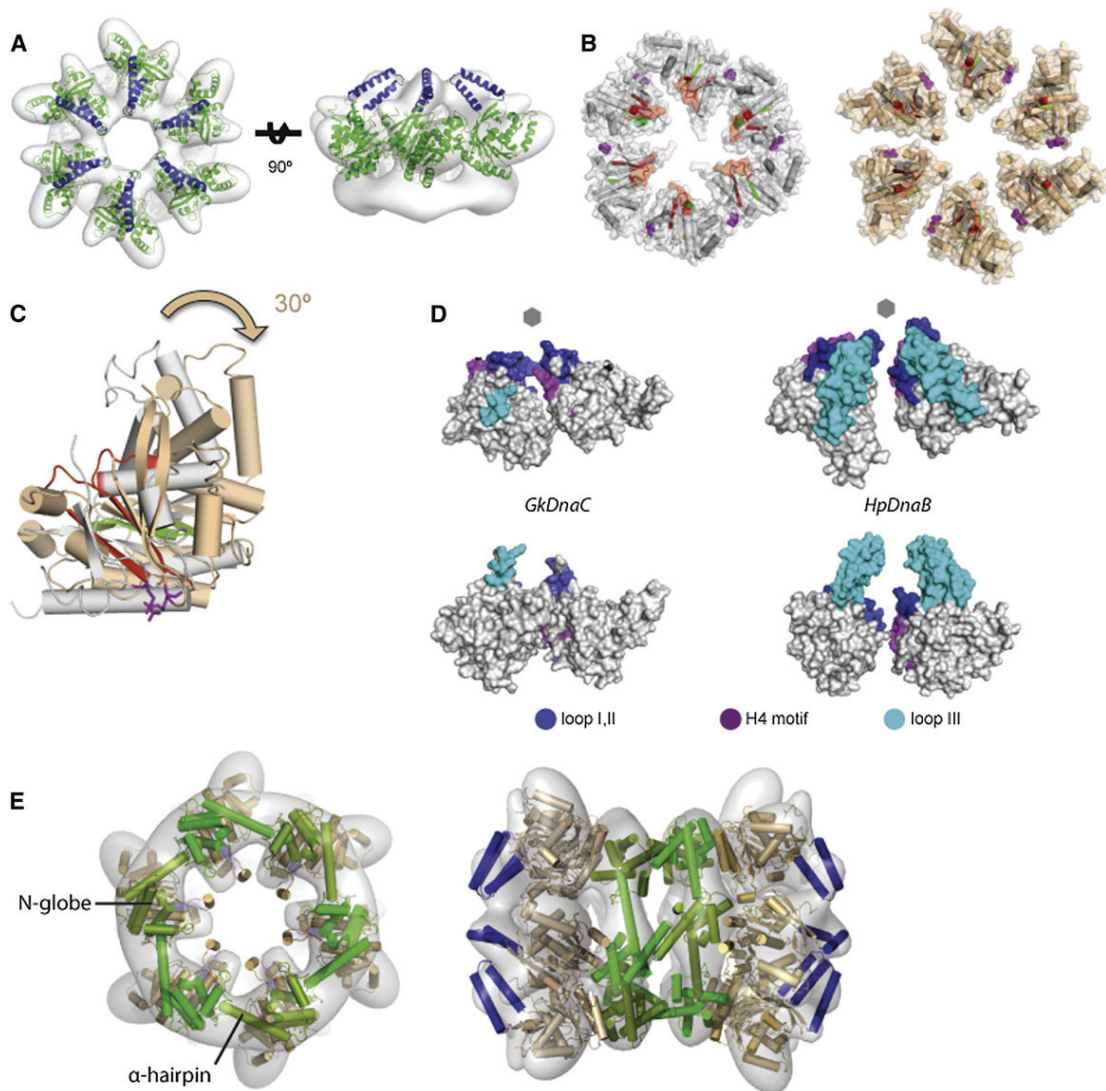


Figure 4. Hexamer of HpDnaB-CTD

(A) Top (left panel) and side view (right panel) of the model of HpDnaB CTD-ring with each subunit depicted as a green ribbon with the HPI colored in blue fitted into the EM density map (gray).

(B) Comparison of the GkDnaC (left) and HpDnaB (right) CTD-rings represented as ribbon with surface depicted in transparency. The nucleotide binding motifs are colored in red (H1) and green (H2), and the arginine finger is shown as a magenta sphere. For clarity, HPIs have been removed from HpDnaB subunits.

(C) Structural superimposition of one GkDnaC-CTD and one HpDnaB-CTD in the hexameric rings colored as in (B).

(D) Comparison of HpDnaB and GkDnaC DNA binding motifs from two adjacent subunits viewed from the top (upper panel) and the side of the ring (bottom panel). Gray hexagons indicate the position of the 6-fold symmetry axis.

(E) Left view is the hexameric assembly of the HpDnaB-NTD subunits (green), shown as ribbon, docked into the NTD-ring of the EM map (gray). Right view is a model of the near full-length dodecameric HpDnaB fit into the EM map with the CTDs colored in grey with the HPIs in blue and the NTD colored in dark and light green in alternating subunits.

See also Figures S5 and S6.

however, share little sequence/structure similarity, suggesting that the specific protein-protein contacts responsible for such assemblies are most likely not conserved. Interestingly, HpDnaB double hexamers present large side channels in the CTD-ring that resemble those found in MtMCM and LTag helicases (Bae et al., 2009; Costa et al., 2006a, 2006b; Gómez-Llorente et al., 2005; Li et al., 2003; Pape et al., 2003) and in the EM reconstructions of EcDnaB and GP40 hexamers in the C_6 symmetry

(Núñez-Ramírez et al., 2006; Yang et al., 2002). Thus, this architecture is likely to represent a common functional state of a subset of replicative helicases. In the case of LTag, it was proposed that ssDNA could thread through these positively charged channels (Li et al., 2003).

A particular feature of the HpDnaB structure that we propose is the mode of oligomerization, which is mediated in part by a unique insertion (HPI) that forms a helix-turn-helix structure

Structure

Structure of *Helicobacter pylori* DnaB

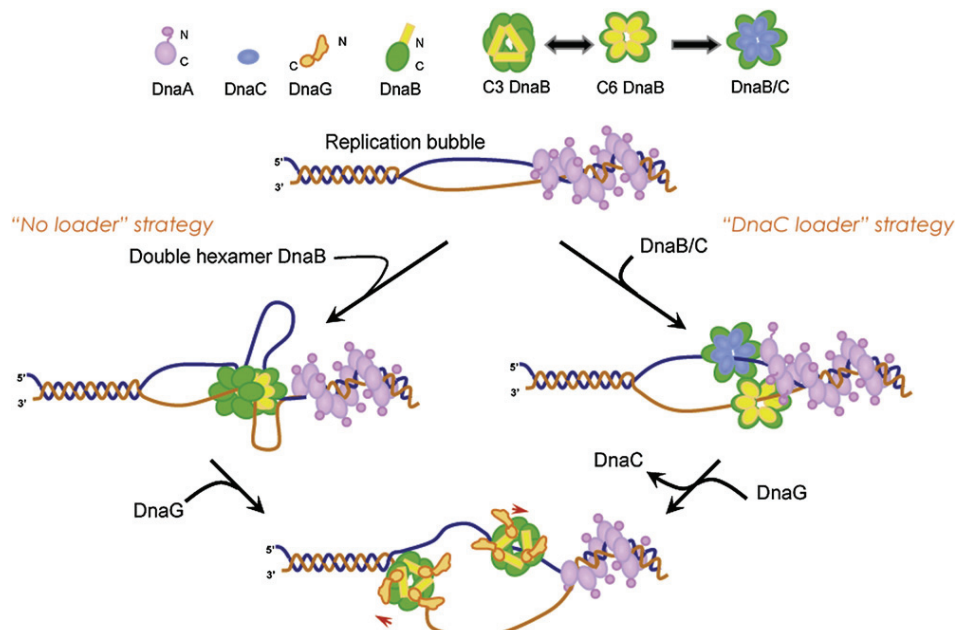


Figure 5. Strategies for DNA Replication Initiation in Different Gram-Negative Bacteria

ATP-bound DnaA oligomerization triggers the formation of the replication bubble (Erzberger et al., 2006). In the helicase loader strategy two DnaB hexamers are loaded by DnaC-DnaA in *A. aeolicus* (Mott et al., 2008) and DnaB-DnaA interactions using DnaC as a ring opener (*E. coli*; Keyamura et al., 2009). In *H. pylori* (and possibly other bacteria lacking helicase loaders), twin DnaB hexamers with NTD-rings in C₆ conformation are loaded onto the replication fork leading to the opening of the bubble. Interaction between DnaA and the double hexamer might direct the loading of the helicase, but additional factors could also be involved such as HobA (Natrajan et al., 2009). The double hexamer might also bind simultaneously to the two ssDNAs that could penetrate the ring via the side opening of the CTDs. Subsequent activation of DnaB's helicase activity may be achieved by binding of three DnaG molecules to the NTD-ring of a DnaB hexamer, which in the loader strategy is known to release DnaC (Makowska-Grzyska and Kaguni, 2010) and trigger the change of conformation of the NTD-ring of DnaB from C₆ to C₃ (Thirlway et al., 2004), thus opening the NTD-ring and allowing the passage of ssDNA.

instead of the DNA binding loop III found in RecA/F family of ATPases. From the reconstruction the HPis assemble as a collar and appear to contribute to the stability of the hexamer, although they are not sufficient for HpDnaB-CTD hexamerization. Together with DNA binding loops I and II, the HPI forms a continuous positively charged patch that significantly modifies the charge of the distal part of the CTD-ring and provides an obvious nucleic acid binding surface. It is noteworthy that the MtMCM double-hexamer reconstruction showed a similar "cap" in the presence of DNA that was proposed to consist of the C-terminal winged helix-turn-helix (wHTH) motif (Costa et al., 2006b). Further work should aim at identifying the role of the HPI and loop III in other DnaBs and comparing it to the wHTH domain of MCMs.

Although replicative helicases share common functions and structural characteristics, many aspects of their mode of action vary such as their polarity, their active oligomeric state, or how they are recruited to DNA replication initiation sites. At these sites, two replicative helicase hexamers must be loaded in opposite directions to achieve bidirectional unwinding of dsDNA and replication fork progression. The viral LTag and E1 helicases recognize the viral origin of replication, and the active helicase assembly consists of a double hexamer that initially pumps DNA at replication forks (Schuck and Stenlund, 2005; Valle et al., 2000). In archaea and eukaryotes, MCM and Mcm2-7 helicases form stable dodecamers at the origin and are recruited by

interaction with initiator proteins Orc1/Cdc6 and ORC, respectively (Costa et al., 2006b; Evrin et al., 2009; Gambus et al., 2011; Remus et al., 2009). Some MCM proteins might use helicase loaders, but not all (Barry and Bell, 2006). These helicase assemblies are subsequently activated by various regulatory kinases in order to initiate DNA unwinding, and at present, it is still unclear whether the double hexamers represent the active form of such helicases (Katayama et al., 2010; Konieczny, 2003). In the most-studied bacterial system, *E. coli*, DNA replication starts with the recognition of origin sequence on the chromosome by the initiator protein DnaA (Figure 5). ATP-bound DnaA oligomerization triggers the formation of the replication bubble (Erzberger et al., 2006). In *E. coli* and *B. subtilis*, DnaBs are loaded via a "loader" strategy, but this might not be the case in most bacteria where helicase loaders have yet to be identified (Caspi et al., 2001; Ioannou et al., 2006; Konieczny, 2003; Velten et al., 2003). Moreover, the "loader strategy" involves different sets of proteins in different bacteria and may also involve several mechanisms for a given set of proteins (Keyamura et al., 2009; Mott et al., 2008; Ozaki and Katayama, 2011; Velten et al., 2003). How two DnaB hexamers can be loaded onto the replication forks in the absence of a helicase loader is still an unanswered question. *H. pylori*'s genome does not contain a sequence homolog of DnaC. Intriguingly, HpDnaB could complement both *dnaB* or *dnaC* temperature-sensitive mutants of *E. coli* (Soni et al., 2003, 2005), suggesting that HpDnaB is able

to load onto *E. coli* replication forks in the absence of DnaC and function as an active helicase.

The double-hexamer structure of HpDnaB provides a simple, yet attractive solution for loading simultaneously and in opposite directions two DnaB hexamers on a replication fork in the absence of a loader (Figure 5). We speculate that in *H. pylori* (and possibly other bacteria lacking helicase loaders), twin DnaB hexamers with NTD-rings in C₆ conformation could be loaded onto the unwound DNA by integration of ssDNA using the side openings of the CTD-ring as proposed for LTag. This step might also be accompanied by the opening of the bubble (Fang et al., 1999). The loading of the helicase may involve an interaction between HpDnaB and the HpDnaA/oriC complex, but additional factors could also play a role. In *E. coli*, DiaA was found to prevent the loading of DnaB by interacting with the N terminus of DnaA (Keyamura et al., 2009). In *H. pylori*, HobA is a structural homolog of DiaA (Keyamura et al., 2009; Natrajan et al., 2007) and forms a HobA₄/DnaA₄ complex absolutely required for bacterial survival (Natrajan et al., 2009). It is, therefore, possible that HobA and/or other factors participate in the targeting of HpDnaB to oriC in *H. pylori*.

The fate of the double ring is at present unclear, but the separation of the two rings could be triggered by the interaction with the DnaG primase. In both the loader and the no loader strategy that we propose here, three DnaG subunits would then interact with the NTD-ring of a DnaB hexamer. In the loader strategy this would lead to the release of DnaC (Makowska-Grzyska and Kaguni, 2010) possibly by triggering the change of conformation of DnaB from C₆ to C₃, the latter being required for DnaB to form a stable complex with DnaG (Bailey et al., 2007b; Wang et al., 2008). Such a change in symmetry of DnaB upon interaction with DnaG has already been observed by atomic force microscopy (Thirlway et al., 2004). In the absence of a loader, interactions between DnaB NTDs and DnaG may disrupt the double hexamer, using a similar C₆-C₃ transition and allow each of the hexameric rings to move along the DNA in opposite directions. Although this speculative model will need further experimentations, the structure of the HpDnaB dodecamer will provide a novel framework to investigate the molecular interplay between replisome proteins and help to unravel the molecular basis of helicase loading onto replication forks.

EXPERIMENTAL PROCEDURES

Protein Purification and Crystallization

HpDnaB and HpDnaB-CTD (residues 152–488) were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen). HpDnaB and HpDnaB-CTD were purified on a His-Trap column (GE Healthcare) and concentrated to 5–10 mg/ml. Proteins were submitted to SEC (Superdex 200; GE Healthcare) in buffer GF (10 mM Tris-HCl [pH 7.5] and 200 mM NaCl). Selenomethionine-substituted HpDnaB-CTD was purified as for the HpDnaB-CTD. Native and selenomethionine-substituted HpDnaB-CTDs were concentrated to 5 mg/ml in buffer GF for crystallization. Crystals of HpDnaB-CTD and selenomethionine-substituted HpDnaB-CTD were grown by vapor diffusion at 20°C using reservoir solution of 0.1 M citric acid (pH 5.0) and 2%–10% w/v Polyethylene Glycol 6000.

Data Collection, Structure Determination, and Refinement

Crystals were flash frozen in liquid nitrogen using the mother liquor supplemented with 25% glycerol for cryoprotection. Crystals of HpDnaB-CTD belonged to the space group P2₁2₁2 with cell dimensions of a = 100.9,

b = 102.2, and c = 85.1 Å, and diffracted to a resolution of 2.5 Å. The structure of HpDnaB-CTD was solved by the SAD method using selenomethionine derivative data. The position of 20 selenium atoms was determined with SHELXD (Schneider and Sheldrick, 2002), refined with the program SHARP (de La Fortelle and Bricogne, 1997), and density modification was performed using DM (Collaborative Computational Project, Number 4, 1994). The model was built using Coot (Emsley and Cowtan, 2004), and the refinement using PHENIX (Adams et al., 2002) converged to yield a model with R and R_{free} factors of 21.7% and 25.1%, respectively, with good geometry (Table 1). The structure has been deposited in the protein data bank with coordinates 4A1F. Figures were generated using PyMOL (DeLano, 2002). The structure-based sequence alignment was generated with ESPript (Gouet et al., 2003).

SEC-MALLS

SEC combined with MALLS and refractometry (RI) experiments were performed with an analytical Superdex 200 column (GE Healthcare) equilibrated with a buffer containing 20 mM Tris (pH 8), 200 mM sodium chloride, and 5% glycerol. For experiments performed with nucleotide-bound HpDnaB, the protein was incubated 30 min on ice in the same buffer containing 5 mM MgCl₂ with 1 mM ATP, ADP, or AMP-PNP, and the SEC experiments were performed in the corresponding buffer. Online MALLS detection was performed with a miniDAWN-TREOS detector (Wyatt Technology, Santa Barbara, CA, USA) using a laser emitting at 690 nm and by refractive index measurement using an Optilab T-REX system (Wyatt Technology). Weight-averaged molar masses (MW) were calculated using the ASTRA software (Wyatt Technology).

ATPase Activity

The rate of ATP hydrolysis by 80 nM HpDnaB in reaction buffer (20 mM Tris [pH 8], 100 mM KCl, 8 mM DTT, 4% sucrose, 80 µg/ml BSA, and 5 mM MgCl₂) was measured at a concentration of 1 mM ATP and in the presence and absence of 1 µM single-stranded polydT 18-mer, 50-mer, or the M13mp18 DNA using pyruvate kinase and lactic dehydrogenase to link the reaction to the oxidation of NADH, using the previously described spectrophotometric method (Panuska and Goldthwait, 1980). The hydrolysis rates presented in Table S1 correspond to the mean of three independent experiments performed with each sample.

Negative-Staining EM and Image Analysis

HpDnaB sample from SEC-MALLS (0.1 mg/ml) was applied to the clear side of carbon on a carbon-mica interface and stained with 2% (w/v) uranyl acetate. Images were recorded with a JEOL 1200 EX II microscope at 100kV and at a nominal 40,000× magnification. Negatives were digitized on a Zeiss scanner (Photoscan TD) to a pixel size of 3.5 Å at the specimen level. A total of 35,389 HpDnaB individual particle subframes were iteratively selected, CTF corrected, aligned, and classified. An elaborate sorting protocol enabled division of the data set into 7,716 frames corresponding to different views of the double-ring form and 26,859 frames containing views of the single-ring form that were used for reconstruction. The resolutions of the hexamer and dodecamer reconstructions were assessed to be around 21 and 23 Å, respectively. Image-processing details are provided in Supplemental Experimental Procedures. Details of experimental procedures used are given in Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, one table, and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.str.2012.01.020.

ACKNOWLEDGMENTS

We are indebted to Professors Edward H. Egelman and Edward P. Morris for providing us with the EM density maps of EcDnaB and Mcm2-7 helicases, respectively. Thanks are due to Samira Acajjaoui for help in setting up the ATPase assay and to Wyatt staff members for help with MALLS data analysis. I.G. thanks the EM platform of the Partnership for Structural Biology (Grenoble) for access to EM equipment. This work was supported by the ESRF in-house

research program for M.S., L.T., and J.T. and also by the ATIP-Avenir CNRS-Ligue contre le Cancer program for G.B. and L.T. A.B. is supported by a region Rhône-Alpes Cible 2011 fellowship.

Received: October 23, 2011

Revised: January 17, 2012

Accepted: January 22, 2012

Published: March 6, 2012

REFERENCES

- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954.
- Bae, B., Chen, Y.H., Costa, A., Onesti, S., Brunzelle, J.S., Lin, Y., Cann, I.K., and Nair, S.K. (2009). Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. *Structure* **17**, 211–222.
- Bailey, S., Eliason, W.K., and Steitz, T.A. (2007a). The crystal structure of the *Thermus aquaticus* DnaB helicase monomer. *Nucleic Acids Res.* **35**, 4728–4736.
- Bailey, S., Eliason, W.K., and Steitz, T.A. (2007b). Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science* **318**, 459–463.
- Barry, E.R., and Bell, S.D. (2006). DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.* **70**, 876–887.
- Biswas, E.E., and Biswas, S.B. (1999a). Mechanism of DNA binding by the DnaB helicase of *Escherichia coli*: analysis of the roles of domain gamma in DNA binding. *Biochemistry* **38**, 10929–10939.
- Biswas, E.E., and Biswas, S.B. (1999b). Mechanism of DnaB helicase of *Escherichia coli*: structural domains involved in ATP hydrolysis, DNA binding, and oligomerization. *Biochemistry* **38**, 10919–10928.
- Biswas, T., and Tsodikov, O.V. (2008). Hexameric ring structure of the N-terminal domain of *Mycobacterium tuberculosis* DnaB helicase. *FEBS J.* **275**, 3064–3071.
- Caspi, R., Pacek, M., Consiglieri, G., Helinski, D.R., Toukdarian, A., and Konieczny, I. (2001). A broad host range replicon with different requirements for replication initiation in three bacterial species. *EMBO J.* **20**, 3262–3271.
- Chintakayala, K., Larson, M.A., Grainger, W.H., Scott, D.J., Griep, M.A., Hinrichs, S.H., and Soultanas, P. (2007). Domain swapping reveals that the C- and N-terminal domains of DnaG and DnaB, respectively, are functional homologues. *Mol. Microbiol.* **63**, 1629–1639.
- Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A., and Onesti, S. (2006a). Structural basis of the *Methanothermobacter thermoautotrophicus* MCM helicase activity. *Nucleic Acids Res.* **34**, 5829–5838.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A., and Onesti, S. (2006b). Structural studies of the archaeal MCM complex in different functional states. *J. Struct. Biol.* **156**, 210–219.
- Davey, M.J., and O'Donnell, M. (2003). Replicative helicase loaders: ring breakers and ring makers. *Curr. Biol.* **13**, R594–R596.
- de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Meth. Enzymol.* **276**, 472–494.
- DeLano, W.L. (2002). <http://pymol.sourceforge.net>.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
- Erzberger, J.P., Mott, M.L., and Berger, J.M. (2006). Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nat. Struct. Mol. Biol.* **13**, 676–683.
- Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad. Sci. USA* **106**, 20240–20245.
- Fang, L., Davey, M.J., and O'Donnell, M. (1999). Replisome assembly at oriC, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. *Mol. Cell* **4**, 541–553.
- Gambus, A., Khoudoli, G.A., Jones, R.C., and Blow, J.J. (2011). MCM2-7 form double hexamers at licensed origins in *Xenopus* egg extract. *J. Biol. Chem.* **286**, 11855–11864.
- Gómez-Llorente, Y., Fletcher, R.J., Chen, X.S., Carazo, J.M., and San Martín, C. (2005). Polymorphism and double hexamer structure in the archaeal mini-chromosome maintenance (MCM) helicase from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* **280**, 40909–40915.
- Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323.
- Ilyina, T.V., Gorbalenya, A.E., and Koonin, E.V. (1992). Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J. Mol. Evol.* **34**, 351–357.
- Ioannou, C., Schaeffer, P.M., Dixon, N.E., and Soultanas, P. (2006). Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in *Bacillus subtilis*. *Nucleic Acids Res.* **34**, 5247–5258.
- Kaplan, D.L., and O'Donnell, M. (2004). Twin DNA pumps of a hexameric helicase provide power to simultaneously melt two duplexes. *Mol. Cell* **15**, 453–465.
- Kashav, T., Nitharwal, R., Abdulrehman, S.A., Gabdoulkhalov, A., Saenger, W., Dhar, S.K., and Gourinath, S. (2009). Three-dimensional structure of N-terminal domain of DnaB helicase and helicase-primase interactions in *Helicobacter pylori*. *PLoS One* **4**, e7515.
- Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010). Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and oriC. *Nat. Rev. Microbiol.* **8**, 163–170.
- Keyamura, K., Abe, Y., Higashi, M., Ueda, T., and Katayama, T. (2009). DiaA dynamics are coupled with changes in initial origin complexes leading to helicase loading. *J. Biol. Chem.* **284**, 25038–25050.
- Konieczny, I. (2003). Strategies for helicase recruitment and loading in bacteria. *EMBO Rep.* **4**, 37–41.
- Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakonyi, G., and Chen, X.S. (2003). Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* **423**, 512–518.
- Lo, Y.H., Tsai, K.L., Sun, Y.J., Chen, W.T., Huang, C.Y., and Hsiao, C.D. (2009). The crystal structure of a replicative hexameric helicase DnaC and its complex with single-stranded DNA. *Nucleic Acids Res.* **37**, 804–814.
- Makowska-Grzyska, M., and Kaguni, J.M. (2010). Primase directs the release of DnaC from DnaB. *Mol. Cell* **37**, 90–101.
- Mott, M.L., Erzberger, J.P., Coons, M.M., and Berger, J.M. (2008). Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. *Cell* **135**, 623–634.
- Natrajan, G., Hall, D.R., Thompson, A.C., Gutsche, I., and Terradot, L. (2007). Structural similarity between the DnaA-binding proteins HobA (HP1230) from *Helicobacter pylori* and DiaA from *Escherichia coli*. *Mol. Microbiol.* **65**, 995–1005.
- Natrajan, G., Noirot-Gros, M.F., Zawilak-Pawlik, A., Kapp, U., and Terradot, L. (2009). The structure of a DnaA/HobA complex from *Helicobacter pylori* provides insight into regulation of DNA replication in bacteria. *Proc. Natl. Acad. Sci. USA* **106**, 21115–21120.
- Navaza, J., Lepault, J., Rey, F.A., Alvarez-Rúa, C., and Borge, J. (2002). On the fitting of model electron densities into EM reconstructions: a reciprocal-space formulation. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1820–1825.
- Nitharwal, R.G., Paul, S., Dar, A., Choudhury, N.R., Soni, R.K., Prusty, D., Sinha, S., Kashav, T., Mukhopadhyay, G., Chaudhuri, T.K., et al. (2007). The domain structure of *Helicobacter pylori* DnaB helicase: the N-terminal domain can be dispensable for helicase activity whereas the extreme C-terminal region is essential for its function. *Nucleic Acids Res.* **35**, 2861–2874.

- Nitharwal, R.G., Verma, V., Dasgupta, S., and Dhar, S.K. (2011). *Helicobacter pylori* chromosomal DNA replication: current status and future perspectives. *FEBS Lett.* **585**, 7–17.
- Núñez-Ramírez, R., Robledo, Y., Mesa, P., Ayora, S., Alonso, J.C., Carazo, J.M., and Donate, L.E. (2006). Quaternary polymorphism of replicative helicase G40P: structural mapping and domain rearrangement. *J. Mol. Biol.* **357**, 1063–1076.
- Oakley, A.J., Loscha, K.V., Schaeffer, P.M., Liepinsh, E., Pintacuda, G., Wilce, M.C., Otting, G., and Dixon, N.E. (2005). Crystal and solution structures of the helicase-binding domain of *Escherichia coli* primase. *J. Biol. Chem.* **280**, 11495–11504.
- Ozaki, S., and Katayama, T. (2011). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic Acids Res.*, in press. Published online November 3, 2011.
- Panuska, J.R., and Goldthwait, D.A. (1980). A DNA-dependent ATPase from T4-infected *Escherichia coli*. Purification and properties of a 63,000-dalton enzyme and its conversion to a 22,000-dalton form. *J. Biol. Chem.* **255**, 5208–5214.
- Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M., and Onesti, S. (2003). Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep.* **4**, 1079–1083.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
- Remus, D., Beuron, F., Tolun, G., Griffith, J.D., Morris, E.P., and Diffley, J.F. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730.
- Schaeffer, P.M., Headlam, M.J., and Dixon, N.E. (2005). Protein–protein interactions in the eubacterial replisome. *IUBMB Life* **57**, 5–12.
- Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1772–1779.
- Schuck, S., and Stenlund, A. (2005). Assembly of a double hexameric helicase. *Mol. Cell* **20**, 377–389.
- Seitz, H., Weigel, C., and Messer, W. (2000). The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. *Mol. Microbiol.* **37**, 1270–1279.
- Sharma, A., Nitharwal, R.G., Singh, B., Dar, A., Dasgupta, S., and Dhar, S.K. (2009). *Helicobacter pylori* single-stranded DNA binding protein—functional characterization and modulation of *H. pylori* DnaB helicase activity. *FEBS J.* **276**, 519–531.
- Singleton, M.R., Sawaya, M.R., Ellenberger, T., and Wigley, D.B. (2000). Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101**, 589–600.
- Smits, W.K., Goranov, A.I., and Grossman, A.D. (2010). Ordered association of helicase loader proteins with the *Bacillus subtilis* origin of replication in vivo. *Mol. Microbiol.* **75**, 452–461.
- Soni, R.K., Mehra, P., Choudhury, N.R., Mukhopadhyay, G., and Dhar, S.K. (2003). Functional characterization of *Helicobacter pylori* DnaB helicase. *Nucleic Acids Res.* **31**, 6828–6840.
- Soni, R.K., Mehra, P., Mukhopadhyay, G., and Dhar, S.K. (2005). *Helicobacter pylori* DnaB helicase can bypass *Escherichia coli* DnaC function in vivo. *Biochem. J.* **389**, 541–548.
- Story, R.M., Weber, I.T., and Steitz, T.A. (1992). The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**, 318–325.
- Sutton, M.D., Carr, K.M., Vicente, M., and Kaguni, J.M. (1998). *Escherichia coli* DnaA protein. The N-terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin. *J. Biol. Chem.* **273**, 34255–34262.
- Syson, K., Thirlway, J., Hounslow, A.M., Soultanas, P., and Waltho, J.P. (2005). Solution structure of the helicase-interaction domain of the primase DnaG: a model for helicase activation. *Structure* **13**, 609–616.
- Thirlway, J., Turner, I.J., Gibson, C.T., Gardiner, L., Brady, K., Allen, S., Roberts, C.J., and Soultanas, P. (2004). DnaG interacts with a linker region that joins the N- and C-domains of DnaB and induces the formation of 3-fold symmetric rings. *Nucleic Acids Res.* **32**, 2977–2986.
- Valle, M., Gruss, C., Halmer, L., Carazo, J.M., and Donate, L.E. (2000). Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. *Mol. Cell. Biol.* **20**, 34–41.
- Velten, M., McGovern, S., Marsin, S., Ehrlich, S.D., Noirot, P., and Polard, P. (2003). A two-protein strategy for the functional loading of a cellular replicative DNA helicase. *Mol. Cell* **11**, 1009–1020.
- Wang, G., Klein, M.G., Tokonzaba, E., Zhang, Y., Holden, L.G., and Chen, X.S. (2008). The structure of a DnaB-family replicative helicase and its interactions with primase. *Nat. Struct. Mol. Biol.* **15**, 94–100.
- Waters, T.R., Eryilmaz, J., Geddes, S., and Barrett, T.E. (2006). Damage detection by the UvrABC pathway: crystal structure of UvrB bound to fluorescein-adducted DNA. *FEBS Lett.* **580**, 6423–6427.
- Watt, S.J., Urathamakul, T., Schaeffer, P.M., Williams, N.K., Sheil, M.M., Dixon, N.E., and Beck, J.L. (2007). Multiple oligomeric forms of *Escherichia coli* DnaB helicase revealed by electrospray ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 132–140.
- Yang, S., Yu, X., VanLoock, M.S., Jezewska, M.J., Bujalowski, W., and Egelman, E.H. (2002). Flexibility of the rings: structural asymmetry in the DnaB hexameric helicase. *J. Mol. Biol.* **321**, 839–849.

Article 2

Structure and primase-mediated activation of a bacterial dodecameric replicative helicase

Bazin, A., Cherrier, M.V., Gutsche, I., Timmins, J., and Terradot, L. (2015) “Structure and primase-mediated activation of a bacterial dodecameric replicative helicase” *Nucleic Acid Res* **43**(17):8564-76.

Structure and primase-mediated activation of a bacterial dodecameric replicative helicase

Alexandre Bazin¹, Mickaël V. Cherrier¹, Irina Gutsche², Joanna Timmins^{3,4,5} and Laurent Terradot^{1,*}

¹CNRS, UMR 5086 Bases Moléculaires et Structurales de Systèmes Infectieux, Institut de Biologie et Chimie des Protéines, 7 Passage du Vercors, F-69367, Lyon, France. Université de Lyon, F-69622, Lyon, France; Université Claude Bernard Lyon 1, F-69622, Villeurbanne, France, ²Unit of Virus Host-Cell Interactions, UJF-EMBL-CNRS, UMI3265, F-38044 Grenoble Cedex 9, France, ³Univ. Grenoble Alpes, Institut de Biologie Structurale, F-38044 Grenoble, France, ⁴CNRS, Institut de Biologie Structurale, F-38044 Grenoble, France and ⁵CEA, Institut de Biologie Structurale, F-38044 Grenoble, France

Received May 18, 2015; Revised July 20, 2015; Accepted July 22, 2015

ABSTRACT

Replicative helicases are essential ATPases that unwind DNA to initiate chromosomal replication. While bacterial replicative DnaB helicases are hexameric, *Helicobacter pylori* DnaB (*HpDnaB*) was found to form double hexamers, similar to some archaeal and eukaryotic replicative helicases. Here we present a structural and functional analysis of *HpDnaB* protein during primosome formation. The crystal structure of the *HpDnaB* at 6.7 Å resolution reveals a dodecameric organization consisting of two hexamers assembled via their N-terminal rings in a stack-twisted mode. Using fluorescence anisotropy we show that *HpDnaB* dodecamer interacts with single-stranded DNA in the presence of ATP but has a low DNA unwinding activity. Multi-angle light scattering and small angle X-ray scattering demonstrate that interaction with the DnaG primase helicase-binding domain dissociates the helicase dodecamer into single ringed primosomes. Functional assays on the proteins and associated complexes indicate that these single ringed primosomes are the most active form of the helicase for ATP hydrolysis, DNA binding and unwinding. These findings shed light onto an activation mechanism of *HpDnaB* by the primase that might be relevant in other bacteria and possibly other organisms exploiting dodecameric helicases for DNA replication.

INTRODUCTION

DNA replication is an essential process in all organisms and consists of the faithful duplication of the genetic material to

be transmitted to daughter cells. Initiation of DNA replication relies on helicases, key motor proteins that unwind DNA so that it can be used as a template for DNA polymerases (1). Replicative helicases are hexameric ring-shaped proteins that hydrolyse ATP and unwind double-stranded DNA (dsDNA), with one strand running through the central cavity of the hexamer and the other being excluded from the ring. In bacteria, the replicative helicase DnaB functions as a hexamer to unwind dsDNA in the 5' to 3' direction (1).

The DnaB monomer consists of an N-terminal domain (NTD, comprising a head and a helical hairpin) and a C-terminal RecA type ATPase domain (CTD) associated through a linker (Figure 1A) (2). The NTDs and CTDs assemble into separate rings (namely NTD- and CTD-rings) that have different symmetries and form a two-tiered hexamer (3,4). In the absence of a ligand, the NTD-ring adopts a trimer of dimers organization (4) (described as 'dilated' (5)), which opens a large central channel in the helicase. In the presence of nucleotides, the NTD can also adopt a triskel like structure (described as 'constricted'), a conformation compatible with interaction with the τ subunit of the polymerase (5).

Because DnaB forms closed rings, threading of single-stranded DNA (ssDNA) through the central channel of DnaB requires ring-breaking or ring-making mechanisms during initiation of replication (6). In addition, because of their polarity, two hexamers of DnaB must be positioned in opposite directions at the origin on each of the DNA forks for replication to proceed bi-directionally (7). In some bacteria, loading of the helicase is assisted by AAA+ helicase loaders such as DnaC in *Escherichia coli* or DnaI in *Bacillus subtilis* (*BsuDnaI*) (7). *E. coli* DnaC forms a 6:6 complex with *EcDnaB* (8,9). The interaction with the loader opens the *EcDnaB* rings into a lock-washer conformation and triggers a helical conformation of the overall complex, thereby providing an entry point for ss-

*To whom correspondence should be addressed. Tel: +33 472722652; Fax: +33 472722604; Email: laurent.terrardot@ibcp.fr

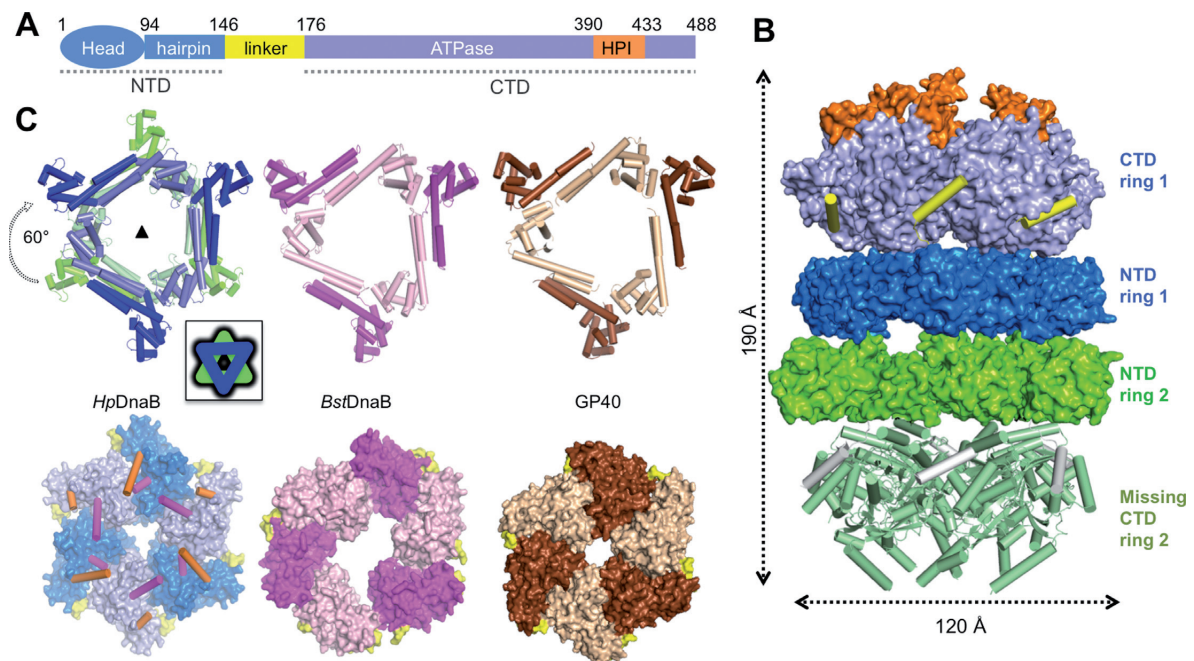


Figure 1. Crystal structure of *HpDnaB* dodecamer. (A) Schematic representation of *HpDnaB* domain organization. (B) Surface representation of His-*HpDnaB* crystal structure with one hexamer colored according to the domain organization described in (A) and the second one in green. The second CTD-ring, absent in the crystal structure has been modeled and is shown as a cartoon. (C) Top panels, structures of NTD-rings of His-*HpDnaB*, apo *BstDnaB* and GP40. The inset depicts a schematic representation of the apparent six-fold symmetry of His-*HpDnaB* NTD-rings. Bottom panels, surface representation of *HpDnaB*, apo *BstDnaB* and GP40 CTD-ring structures. Subunits are colored in alternating shades for clarity, linker helices are colored in yellow, HPI helices $\alpha 15$ (orange) and $\alpha 16$ (magenta) are shown as cylinders.

DNA (10). A non-planar architecture has also been observed in the pre-priming complex composed of *BsuDnaI* and *Bacillus stearothermophilus* helicase (*BstDnaB*), as well as the helicase-binding domain (HBD) of the primase (*BstDnaG*^{HBD}) (11), suggesting that helicase loading might depend on a transition of the helicase from planar to non-planar conformation induced or stabilized by the AAA+ helicase loader (10,11).

During replication, the active form of DnaB is a complex of the helicase with up to three DnaG primases, named the primosome. Transition from a loaded helicase•helicase-loaders•primase complex to an active unwinding primosome requires additional conformational changes of the helicase. Closing of the helicase rings is likely facilitated by nucleotide and ssDNA binding since the translocation conformation of DnaB on ssDNA is also spiral (12). *BstDnaG*^{HBD} interaction with the pre-primed *BstDnaB*•*BsuDnaI* complex resulted in a ternary complex that might be loaded onto the replication fork (11). In *E. coli*, primer synthesis by DnaG in complex with DnaB releases DnaC (13) and stimulates DnaB helicase and ATPase activities (14). Therefore, bacteria appear to have developed different strategies for the formation of primosomes. Moreover, AAA+ loader-dependent primosomes are unlikely to represent a universal strategy since loader homologues have been identified in only a subset of bacterial species (~20%) (15). Concomitantly, the human pathogen *Helicobacter pylori* does not encode for a DnaC/I homologue. Moreover, *H. pylori* DnaB (*HpDnaB*) could complement two temperature-sensitive mutants of *E. coli dnaBts* and *dnaCts*, suggesting

that *HpDnaB* was able to bypass DnaC in these cells (16). *HpDnaB* formed a double hexamer (17) with an architecture remarkably similar to head-to-head double hexamers of some replicative helicases of archaea, yeast and some viruses (7,17). However, the absence of higher resolution structures of dodecameric helicases and the limited resolution of our *HpDnaB* reconstruction (23 Å) prevented insightful structural comparison. It was also unclear whether such a double hexameric assembly could bind and unwind DNA and what its fate was during primosome formation.

Here, we have solved the crystal structure of the *HpDnaB* dodecamer at a resolution of 6.7 Å, which reveals that the helicase can assemble as a planar, stack-twisted double hexamer. Using a complementary set of methods, we found that *HpDnaB* interacts preferentially with ssDNA as a dodecamer in the presence of ATP but is poorly active for DNA unwinding. Interaction with *HpDnaG*^{HBD} dissociates the helicase dodecamer into single-ringed primosomes that represent a more active form of the helicase for DNA unwinding. We propose that in the absence of a control mechanism by AAA+ loaders, dissociation of the double hexamer of the helicase by the primase represents a key switch to launch DNA replication in *H. pylori* and possibly other replication systems utilizing double hexamer replicative helicases.

MATERIALS AND METHODS

Gene cloning, protein expression and purification

His-*HpDnaB* was purified as previously described (17) with an N-terminal histidine-tag consisting of 28

amino-acids (MHHHHHHGKPIPPLLGLDSTENLYFQG). Selenomethionine-substituted His-*HpDnaB* (His-*HpDnaB*^{Se}) was produced using the protocol described in (18) and purified as His-*HpDnaB* (17). For wild-type *HpDnaB* expression, the DNA sequence of *hpdnab* gene was amplified by polymerase chain reaction (PCR) using the forward primer 5'-agtcataatggatcattaaagcatttgcag-3' containing a *NdeI* restriction site (underlined) and reverse primer 5'-atactcgagttcaagttgtaactatataatcc-3' containing a *XhoI* site (underlined). The fragment was inserted into the pACYCDuet-1 vector to produce the pACYC*hpdnab* expression vector. Single point mutations were introduced into pACYC*hpdnab* using the Quick change site directed mutagenesis kit (Stratagene) using manufacturer's protocol. An N-terminal deletion mutant starting at residue 10 (*HpDnaB*^{Δ9}) was generated by PCR (forward primer 5'-caccatgttgcaaacattgaaggatcgtgc-3' and same reverse primer than for wild-type) and inserted into pET101/D-TOPO vector (Invitrogen). The sequence encoding for the helicase binding domain of *HpDnaG* (*HpDnaG*^{HBD}, residues 415–559) was amplified from *H. pylori* genomic DNA (strain 26695) using the forward 5'-caccgtctctttcagccttttacc-3' and reverse 5'-tcataatggcgactaattctcttg-3' primers and inserted into pET151/D-TOPO (Invitrogen) to generate the plasmid pET*hpdnag*^{HBD}.

HpDnaB and mutants were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen) grown in LB medium with chloramphenicol at 50 μg.l⁻¹ (or ampicillin at 100 μg.l⁻¹ for *HpDnaB*^{Δ9}) at 37°C until OD₆₀₀ = 0.6. Protein expression was induced with 1mM isopropyl β-D-thiogalactopyranoside (IPTG) during 16 h at 20°C. Harvested cells were resuspended in lysis buffer (10 mM phosphate pH 7.5, 200 mM NaCl) with protease inhibitor tablet (complete EDTA-free, Roche), lysozyme (Roche) and DNase I (Sigma-Aldrich). The cells were lysed by sonication and centrifuged at 16 000 g for 20 min. The soluble fraction was diluted four-fold with buffer A and applied onto a HiTrapTM Heparin 5 ml column equilibrated with buffer A (10 mM phosphate pH 7.5, 2 mM βME). The protein was eluted using a 0–100% gradient of buffer B (10 mM phosphate pH 7.5, 1 M NaCl, 2 mM βME). Fractions containing the protein were pooled and applied on a HiTrap Q HP (GE Healthcare) column as a second purification step using the same elution protocol. The protein was concentrated and loaded onto a Superdex 200 10/300 GL gel filtration column (GE Healthcare) equilibrated in 10 mM phosphate pH 7.5, 200 mM NaCl, 2 mM βME.

For *HpDnaG*^{HBD} expression, *E. coli* BL21 Star (DE3) cells (Invitrogen) carrying the pET*hpdnag*^{HBD} plasmid were grown in LB medium with ampicillin (100 μg.l⁻¹) at 37°C until an OD₆₀₀ of 0.6 and protein expression was induced by adding 1mM IPTG at 20°C overnight. The cells were harvested and resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl) with protease inhibitor (Roche), lysozyme (Roche) and DNase I (Sigma). The cells were sonicated and centrifuged at 16 000 g for 20 min at 4°C. The supernatant was applied on a HiTrapTM HP 5 ml (GE Healthcare) pre-equilibrated with buffer C (50 mM Tris pH 8.0, 200 mM NaCl, 1 mM DTT and 5% (v/v) glycerol). The protein was eluted by applying a linear gradient (10–100%) of

buffer D (buffer C with 500mM imidazole). Fractions containing *HpDnaG*^{HBD} were pooled and incubated with TEV protease with 1 mM DTT and 0.5 mM EDTA, and dialyzed overnight against buffer D at 4°C. After TEV cleavage the protein fraction was loaded on the HisTrapTM column and *HpDnaG*^{HBD} eluted with 10% of buffer B. The protein was injected on a Superdex 200 10/300 GL (GE Healthcare) gel filtration column and eluted with buffer C.

Crystallization, X-ray diffraction data collection and structure determination

Crystals of His-*HpDnaB* and His-*HpDnaB*^{Se} were grown at 292 K in polyethylene glycol 3350 26%, 100 mM Tris pH 8.6 and 200 mM lithium sulphate with protein incubated with 1 mM ADP and 1 mM MgCl₂. Drops were set-up by mixing 1.5 μl of protein solution (10 mg.ml⁻¹) with 1.5 μl of reservoir solution. Hexagonal crystal forms were obtained after one to two months. Tantalum Bromide derivative crystals were obtained by incubating the crystals for 2 h in the presence of 0.01 mg of Ta₆Br₁₂²⁺ powder (Jena Biosciences GmbH). Single crystals were harvested and flash frozen in liquid nitrogen (100 K) using the mother liquor supplemented with 10% sucrose for cryoprotection. Two single-wavelength X-ray diffraction data sets (Tant 1 and Tant 2) were collected from a single Tantalum Bromide crystal on beamline Proxima 2a at the synchrotron SOLEIL. His-*HpDnaB*^{Se} X-ray data were collected on the beamline ID14EH4 at the European Synchrotron Radiation Facility. The diffraction data were indexed and integrated using XDS (19) and scaled with SCALA from the CCP4 program suite (20). To verify that the crystals contained the full-length protein *HpDnaB*, crystals were dissolved and analyzed on SDS-PAGE.

Structure determination, model building and refinement

The structure of *HpDnaB* was solved by the single anomalous dispersion method with Tant 1 data collected at the inflection wavelength from a single crystal. The positions of three Tantalum clusters were found using HYSS (21) and were then submitted to Phaser (22), which is integrated in Autosol in PHENIX (23). The experimental map at 7.5 Å resolution was of excellent quality and revealed that the asymmetric unit contained four molecules of *HpDnaB*. Four molecules of the NTD high-resolution (2.2 Å) crystal structure (residues 1 to 120, (24)) of *HpDnaB* (*HpDnaB*^{NTD}) were placed into the electron density map and the second helices of the four NTD α-hairpins were built using COOT (25). These domains were used for NCS averaging using PARROT (26) which improved the electron density map. Two high-resolution (2.5 Å) crystal structures of the CTD (residues 176 to 473 (17)) of *HpDnaB* could then be placed and fitted into this averaged map. Placing these domains resulted in $R_{\text{factor}}/R_{\text{free}}$ of 0.30/0.32, respectively, after a first round of refinement with BUSTER TNT (27). The linker helices were placed and the *Helicobacter pylori* insertion (HPI) helices (Figure 1A) were adjusted by rigid body modeling. Density was observed for two helices of the second CTD-ring but the two subunits could not be placed reliably. Soaking the crystals in a solution containing

ssDNA or co-crystallization experiments in various conditions (ssDNA lengths, nucleotides and concentration) did not improve either the resolution or the electron density map corresponding to this region. No ADP molecule could be placed in the nucleotide binding site since no density was observed in the Fo-Fc map with or without NCS averaging. The resulting model was subsequently refined against Tant 2 data set to extend the resolution to 6.7 Å using BUSTER TNT (27) and PHENIX (23). In the final rounds of refinement, a cluster of Ta₆Br₁₂²⁺ could be placed in the density. During refinement, tight geometry constraints were maintained, group B-factors and NCS restraints were used. The final His-*HpDnaB* model was refined to a resolution of 6.7 Å with $R_{\text{factor}}/R_{\text{free}}$ of 0.25/0.29 and very good geometry. Data collection, phasing and refinement statistics are shown in Table 1. The model coordinates were deposited in the protein data bank (pdb code 4ZC0).

Size exclusion chromatography (SEC)-Multi-angle light scattering (MALS)

Size exclusion chromatography (SEC) combined with multi-angle light scattering (MALS) and refractometry (RI) experiments were performed with a Shodex KW405-4F size exclusion column equilibrated with 50 mM Tris pH 8.0, 200 mM NaCl, 1 mM DTT and 5% (v/v) glycerol. 12.5 μl of protein samples of *HpDnaB* or mutants (15 mg.ml⁻¹) and *HpDnaG*^{HBD} (20 mg.ml⁻¹) were injected onto the column. For samples containing nucleotides the proteins were first incubated with 5 mM ATP or AMPPNP and 5 mM MgCl₂ and the buffer was supplemented with 0.5 mM of the corresponding nucleotides and 5 mM MgCl₂. For samples of the *HpDnaB*•*HpDnaG*^{HBD} complex, the separated proteins were incubated at equal molar amounts at a final concentration of 15 mg.ml⁻¹. For ssDNA complex measurements the protein or complexes were first mixed with a 20mer polydT oligonucleotide at either 45 μM or 75 μM (final concentration). On-line MALS detection was performed with a miniDAWN-TREOS detector (Wyatt Technology Corp., Santa Barbara, CA) using a laser emitting at 690 nm and refractive index measurements were performed using an Optilab T-rex system (Wyatt Technology Corp., Santa Barbara, CA). Weight averaged molar masses were calculated using the ASTRA software (Wyatt Technology Corp., Santa Barbara, CA).

Negative staining electron microscopy and image analysis

HpDnaB sample from SEC-MALS (0.1 mg.ml⁻¹) was applied to the clear side of carbon on a carbon-mica interface and stained with 2% (w/v) uranyl acetate. Images were recorded with a JEOL 1200 EX II microscope at 100 kV and at a nominal 40 000x magnification. Negatives were digitized on a Zeiss scanner (Photoscan TD) to a pixel size of 3.5 Å at the specimen level. 9135 individual particles of *HpDnaB* were semi-automatically selected with the EMAN boxer routine (28), CTF-corrected with CTFIND3 (29) and bsoft (30), and low-path-filtered at 15 Å with Spider (31). This data set was subjected to multivariate statistical analysis and classification with Imagic-5 (32).

Small-angle X-ray scattering experiments

SAXS data were recorded on beamline Swing at SOLEIL Synchrotron (Gif sur Yvette, France) at a wavelength of 1.003 Å on a 17 cm × 17 cm low-noise Avix charge-coupled device detector positioned at a distance of 1800 mm from the sample, with the direct beam off-centered. The useful Q-range was 0.004–0.61 Å⁻¹, where $Q = 4\pi\sin\theta/\lambda$ is the scattering vector, and 2θ is the scattering angle. 50 μl of sample in buffer S (50 mM Tris pH 8.0, 200 mM NaCl, 5% (v/v) glycerol, 1 mM DTT) were injected into a size-exclusion column (Shodex KW405-4F) using an Agilent® HPLC system cooled at 288 K and eluted directly into the SAXS flow through capillary cell at a flow rate of 200 μl.min⁻¹, as previously described (33). Protein and buffer samples were prepared exactly as described for the MALS experiments, i.e. for samples containing nucleotides, the proteins were first incubated with 5 mM ATP or AMPPNP and 5 mM MgCl₂ and the buffer was supplemented with 0.5 mM of the corresponding nucleotides and 5 mM MgCl₂ to minimize the effects due to ATP hydrolysis. SAXS data were collected on-line throughout the whole elution time, with a frame duration of 2 s and a dead time between frames of 1 s. A first data set of 90 frames, collected before the void volume, was averaged to account for buffer scattering. A second data set was collected for the sample, from which the 10 frames corresponding to the top of the elution peak were averaged and used for data processing after baseline subtraction (Supplementary Figure S1). Data were processed using the local application FOXTROT (<http://www.synchrotron-soleil.fr/Recherche/LignesLumiere/SWING>) and analyzed using PRIMUS (34). After processing, Kratky plots were generated to verify if proteins were correctly folded and P(r) functions were calculated (Supplementary Figure S2). Theoretical curves from the models were generated by FoXS (35). Minimal Ensemble Search was performed using MES program (36).

Chemical cross-linking

Cross-linking of *HpDnaB* and mutants L4A, Q8A, E80A and Δ9 was performed in 20 mM HEPES pH 7.6, 150 mM NaCl, 0.5 mM βME and 5% (v/v) glycerol. For *HpDnaB* with AMPPNP and ssDNA, the protein was incubated 5 min with AMPPNP (5 mM), MgCl₂ (5 mM) and a 20mer polydT ssDNA oligonucleotide (75 μM final concentration) at room temperature before cross-linking. Crosslinking agent ethylene glycol bis(succinimidyl succinate) (EGS) (5 mM final concentration) was added to 20 μg of protein in 20 μl final volume. The mixture was then incubated at room temperature during 10 min and the reaction was quenched with 2 μl of 1M Tris pH 7.5. The samples were run on gradient SDS-PAGE (4–12%) and stained with Coomassie blue for analysis.

ATPase activity

The rate of ATP hydrolysis by *HpDnaB* (500 nM of monomer) in reaction buffer 20 mM Tris pH 8.0, 100 mM KCl, and 5 mM MgCl₂ was measured at a concentration of 5 mM ATP and in the presence and absence of 500 nM single stranded 20mer polydT DNA and/or 500 nM

Table 1. Data collection, phasing and refinement statistics

	Tant 1 (Ta ₆ Br ₁₂ ²⁺) SOLEIL-Proxima2	Tant 2 (Ta ₆ Br ₁₂ ²⁺) SOLEIL-Proxima2	SeMet ID14EH4-ESRF
Wavelength (Å)	1.2546	1.2546	0.97239
Space group	I2 ₁ 3	I2 ₁ 3	I2 ₁ 3
Unit-cell parameters (Å)	a = b = c = 284.03 β = 90	a = b = c = 283.47 β = 90	a = b = c = 283.11 β = 90
Resolution limits (Å)	48.8–7.5 (7.91–7.50)	47.2–6.7 (6.93–6.70)	48.6–8.0 (8.43–8.0)
No. of observations	171283 (25533)	308609 (45815)	29801 (4471)
No. of unique observations	5016 (725)	6954 (1005)	4111 (602)
Completeness (%)	99.6 (100)	99.7 (100)	99.5 (100)
Anomalous completeness (%)	99.7(100)	99.8 (100)	99.7(100)
Multiplicity	34.1 (35.2)	44.3 (45.6)	7.2 (7.4)
Anomalous multiplicity	18.1 (18.2)	23.3 (23.6)	3.8 (3.8)
R _{meas}	0.13 (> 1)	0.18 (> 1)	0.10 (0.64)
R _{pim}	0.03 (0.27)	0.03 (0.99)	0.05 (0.23)
R _{merge}	0.11 (> 1)	0.17 (> 1)	0.09 (0.55)
CC _{1/2}	99.9 (78.9)	99.9 (41.6)	99.8 (81.9)
I/σ(I)	27.6 (2.8)	18.3 (1.2)	17.4 (3.6)
Wilson B-factor (Å ²)	334	502	247
<i>Refinement</i>			
Reflections used all (free)		13211 (688)	
R-work		0.258 (0.332)	
R-free		0.2988 (0.347)	
Ligand atoms		18	
Protein residues		1173	
RMS (bonds)		0.004	
RMS (angles)		0.838	
Ramachandran favored (%)		96.9	
Ramachandran allowed (%)		2.9	
Ramachandran outliers (%)		0.2	
Clashscore		12.58	
Overall score		1.92	
Average B-factor (Å ²)		315.5	

Values in parentheses refer to the indicated resolution shell.

HpDnaG^{HBD}. ATP hydrolysis was measured using the previously described spectrophotometric method (37) based on the coupling of ATP hydrolysis to the oxidation of NADH with pyruvate kinase and lactate dehydrogenase. The hydrolysis rates presented correspond to the mean of three independent experiments performed with each sample.

DNA binding

Equilibrium DNA binding assays were performed on a Clariostar (BMG Labtech) microplate reader, fitted with polarization filters to measure fluorescence anisotropy. The binding assays were conducted in 384-well plates at room temperature in 40 μl reaction volumes in 20 mM Hepes pH 7.5, 5 mM MgCl₂, 50 mM potassium acetate and 5% (v/v) glycerol supplemented with 0.2 mg.ml⁻¹ BSA and 0.5 mM AMPPNP or ATP. 0 to 100 nM *HpDnaB*, *HpDnaB*•*HpDnaG*^{HBD} complex, or 0 to 15 μM *HpDnaG*^{HBD} were titrated into 1 nM 5'-FAM-labeled 20mer dsDNA composed of oligonucleotides 5'- GACTACGTACTGTTACGTCT-3' and 5'-AGACGTAACAGTACGTAGTC-3' or 1 nM 5'-FAM-labeled 20mer polydT ssDNA oligonucleotide. After subtracting the polarization values obtained for DNA alone, the mean data from three independent experiments were fitted to a standard binding equation ($y = B_{max} * x^h / (Kd^h + x^h)$) assuming a single binding site with Hill slope (h) using GraphPad Prism6. The fits were very

good, with R² values all above 0.98 (Supplementary Table S1).

DNA unwinding assay

Unwinding of forked DNA was carried out by first annealing a 5'-FAM-labeled oligonucleotide (5'-FAM-TACGTAACGAGCCTGC(dT)₂₅-3') to a 1.1 molar excess of a 3' dabcyI quencher-labeled strand (5'-(dT)₂₅-GCAGGCTCGTTACGTA-DabcyI-3'). A capture oligo (5'-GCAGGCTCGTTACGTA-3') complementary to the base-paired region of the FAM-labeled strand was added to all reactions to prevent reannealing of the unwound substrate. 80 μl reactions containing 50 nM *HpDnaB* dodecamer (with or without 7.3 μM *HpDnaG*^{HBD}) and 100 nM fork substrate in helicase buffer (10 mM Tris pH 8.0, 50 mM NaCl, 1% glycerol and 5 mM MgCl₂) were started by addition of 100 nM capture DNA and 2 mM ATP and were monitored at 37°C for 3 h using a Clarisotar (BMG Labtech) fluorescence plate reader. Reactions were performed in triplicate and presented data correspond to the average of at least three independent measurements.

RESULTS

Crystal structure of *HpDnaB*

Crystals of N-terminally 6-histidine-tagged *HpDnaB* (His-*HpDnaB*) diffracted to 6.7 Å resolution and belonged to

the high symmetry space group $I2_13$. The structure was solved by the single anomalous dispersion method and the final model was refined to an $R_{\text{factor}}/R_{\text{free}}$ of 0.25/0.29 with very good geometry (Supplementary Figure S3A, B and Table 1). Given the low resolution, SeMet-substituted protein crystals were grown and used to collect a single wavelength data set at the Se edge to ascertain the position of the methionines in the structure. An anomalous difference Fourier map confirmed the correct location of modeled methionines leaving no ambiguity regarding the position of the subdomains of His-*HpDnaB* (Supplementary Table S2 and Figure S3C).

The crystal structure consists of two His-*HpDnaB* homodimers (chains AB and CD) per asymmetric unit. While chains A and B comprise both the NTD and CTD, chains C and D display only the NTD. Two hexamers, each of which is composed of three homodimers, are generated with the operation of a crystallographic three-fold symmetry (Supplementary Figure S3A). As a result, hexamer 1 consisting of three copies of chains AB is complete with both NTD- and CTD-rings, while hexamer 2 consisting of three copies of chains CD has only the NTD-ring. The absence of the CTD-ring in hexamer 2 cannot be accounted for by protein degradation, since *HpDnaB* was intact in the crystals (Supplementary Figure S3D). Instead, analysis of the crystal packing indicates that the entire CTD-ring 2 is in a large solvent channel. A model of the complete dodecamer was generated by superimposing a copy of hexamer 1 onto the NTD-ring 2 (Figure 1B). The complete dodecamer measures 190 Å in length and 120 Å in diameter (Figure 1B). When this model is used to generate the crystal lattice, no clashes were observed between symmetry-related molecules and the subunits closest to the CTD-ring 2 are at 25 Å (Sup. Movie M1). This demonstrates that the solvent channel is sufficiently large to accommodate the CTD-ring 2. Given that no particular flexibility of the CTD was detected in EM studies (17), we concluded that the absence of density for the CTD-ring 2 is due to rigid-body movements of individual CTDs permitted by the high crystal solvent content (70%) and the lack of crystal contact in this area.

HpDnaB forms a stack-twisted double hexamer

The structure of His-*HpDnaB* reveals that the last two helices of the NTD fold into a helical hairpin that mediates NTD dimer formation (Figure 1C and Supplementary Figure S3B) as in all DnaB structures (3,4,38). The two NTD-rings of His-*HpDnaB* are identical and composed of a planar trimer of dimers delimiting a 50 Å wide channel, a state recently described as ‘dilated’, found in apo DnaBs (3,4,38,39) (Figure 1C). The NTD-rings stack on each other, and if the two hexamers are projected onto the same plane, the vectors through the center-of-mass are offset by 60 degrees (Figure 1C). These two hexamers can be described as a ‘stack-twisted’ assembly. Strikingly, the arrangement adopts an apparent six-fold symmetry when observed from the top thus explaining the six-fold symmetry detected in our previous negative staining electron microscopy (nsEM) reconstruction (17). The interactions of the NTD-rings are mediated by head domains associations, burying a total surface of $\sim 1400 \text{ \AA}^2$. Helices $\alpha 1$ and $\alpha 5$ of

chain A engage $\alpha 1$ of chain D and $\alpha 5$ of chain C (NTD-ring 2), respectively, and chain B (NTD-ring 1) interacts with chain C via $\alpha 1$ – $\alpha 1$ interactions (Figure 2A). Although the low resolution of the structure does not allow for a precise description of the dodecamer interface, $\alpha 1$ – $\alpha 1$ interactions seem predominantly hydrophobic, involving the first residues of the helix.

The structural basis for the connection between the CTD subunits is the same as in other DnaB structures, with each linker helix inserted in the hexamerisation pocket of the adjacent ATPase domain (Figure 1B,C) (3–5). The CTD-ring adopts a pseudo six-fold symmetry with a closed-ring similar to GP40 (Figure 1C) (3,4). One feature that we identified previously in the structure of the isolated *HpDnaB*^{CTD} is a large insertion (named HPI for *Helicobacter pylori* insertion) within the ATPase domain forming a helix-turn-helix composed of $\alpha 15$ and $\alpha 16$ (Figure 1B,C). In the dodecameric crystal structure, the HPI helices from chain A and chain B undergo important conformational changes resulting in different structures compared to the HPI of the *HpDnaB*^{CTD} structure (17). In chain A, $\alpha 16$ and the N-terminal portion of $\alpha 15$ are ordered while the situation is inverted in chain B with $\alpha 15$ and the C-terminal part of $\alpha 16$ ordered (Figure 1C and Supplementary Figure S4). As anticipated from the nsEM structure, the HPIs form a collar extending the CTD-ring (Figure 1B,C). Interestingly, $\alpha 15$ of chain A (but not of chain B) interacts with the adjacent ATPase domain. This feature suggests a structural explanation for some of the unusual properties of *HpDnaB*. In contrast with other DnaBs, *HpDnaB* maintains both ATPase and helicase activities in the absence of the NTD head domain (40). The structure of the His-*HpDnaB* dodecamer suggests that interaction of the HPI collar with the ATPase domain could stabilize the CTD-ring in the absence of NTD during DNA unwinding and/or translocation experiments. This is also supported by the fact that a HPI-deleted *HpDnaB* mutant protein does not form hexamers and is inactive for ATP hydrolysis and DNA unwinding (40).

Dynamics of *HpDnaB* in solution

We observed that in the crystal structure, some residues of the N-terminal His-tag are located at the NTD–NTD interface (Figure 2A). Concerned by the fact that the His-tag might influence the oligomeric state of *HpDnaB* as seen in other hexameric AAA+ ATPases (41), we purified the protein without the His-tag (named *HpDnaB* hereafter). Size-Exclusion Chromatography coupled with Multi-Angle Static Light Scattering (SEC-MALS) experiments showed that *HpDnaB* elutes as a single peak and has a mass of around 585 kDa. This molecular mass is lower than the expected *HpDnaB* dodecamer (669 kDa) suggesting that the protein assembly might be in equilibrium between hexamer and dodecamer. Online size exclusion chromatography coupled to small angle X-ray scattering (SAXS) was used to determine the radius of gyration (R_g) of the proteins. R_g s obtained for His-*HpDnaB* and *HpDnaB* are comparable (67.4 Å and 66.7 Å, respectively; Supplementary Table S3). R_g s calculated from the model of the *HpDnaB* hexamer or dodecamer based on our crystal structure are 46.8 Å and 65.5 Å, respectively (Table 2, see also Figures S1 and S2). Thus,

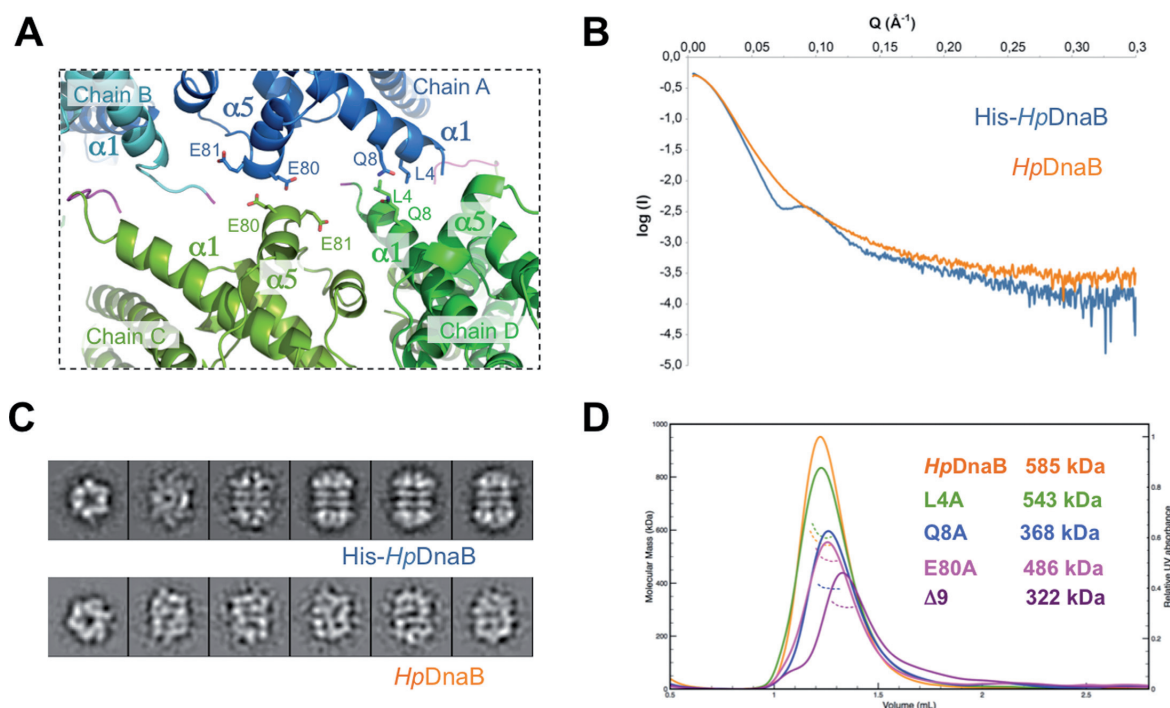


Figure 2. Conformation of the dodecamer and site-directed mutagenesis study of the hexamer-hexamer interface. (A) Close-up view of the hexamer-hexamer interface with side chains of participating residues shown as sticks. Helices are labeled and the histidine-tag is colored in magenta. (B) Experimental SAXS curves of His-*HpDnaB* (blue) and *HpDnaB* (orange). (C) Representative reference-free 2D class averages of His-*HpDnaB* and *HpDnaB*. (D) Size exclusion chromatograms (280 nm) of *HpDnaB* and mutants L4A, Q8A, E80A, and $\Delta 9$ performed on a Superdex 200 increase (GE, 0.45 ml.min⁻¹). MALS weight-averaged molar masses are indicated as dotted lines.

our SAXS data clearly demonstrate that His-*HpDnaB* and *HpDnaB* are both predominantly dodecameric in solution.

We noticed that the scattering curves of the two proteins were nonetheless different, with the His-*HpDnaB* curve exhibiting a dip and hump between Q values of 0.07 to 0.09 (Figure 2B). This feature has previously been found in ring shape helices in solution, including *EcDnaB* (5,42). This feature is absent in the *HpDnaB* SAXS curve, suggesting that *HpDnaB* is more dynamic and that the shapes of the two dodecameric particles are different. By comparing the SAXS profiles with a theoretical SAXS curve of the dodecamer model, we observed that the stack-twisted dodecamer fits better the His-*HpDnaB* SAXS curve ($\chi^2 = 11.8$) than the *HpDnaB* curve ($\chi^2 = 23.6$). Both fits were improved by using multi-ensemble searches (MES, i.e. a combination of different structures (36)). A better fit was obtained for His-*HpDnaB* data ($\chi^2 = 8.8$) using a combination of 91% of dodecamer and 9% of hexamer (Supplementary Figure S5A) and for *HpDnaB* data ($\chi^2 = 14.1$) using a combination of 75% of dodecamer and 25% of hexamer (Supplementary Figure S5B).

To gain further insights into the organization of *HpDnaB* dodecamers, nsEM images of a *HpDnaB* sample were compared with the previously obtained His-*HpDnaB* images (17). As in the case of His-*HpDnaB*, class averages of *HpDnaB* confirm the presence of four-layered particles, but the symmetry is less obvious than for His-*HpDnaB* (Figure 2C). No 3D reconstruction was attempted at this stage because, in contrast to His-*HpDnaB*, multivariate statis-

tical analysis did not reveal any clear cyclic or dihedral symmetry. Nevertheless, the obtained classes suggest that the *HpDnaB* dodecamer also relies on NTD-NTD interactions.

To determine if the interface between the NTD-rings seen in the crystal structure of His-*HpDnaB* is also relevant in *HpDnaB*, we generated a series of structure-guided mutants. An N-terminal deletion mutant starting at residue 10 (*HpDnaB* ^{$\Delta 9$}) and two point mutants L4A (*HpDnaB*^{L4A}) and Q8A (*HpDnaB*^{Q8A}) were constructed to disrupt the contacts between $\alpha 1$ helices. Another mutant, E80A (*HpDnaB*^{E80A}) was designed to target the $\alpha 5$ - $\alpha 5$ interaction. SEC-MALS experiments performed on the purified *HpDnaB* mutants showed that the mass of the *HpDnaB* ^{$\Delta 9$} was 322 kDa, so closer to that of a hexamer (334 kDa) (Figure 2D). In comparison, single mutants *HpDnaB*^{L4A}, *HpDnaB*^{E80A} and *HpDnaB*^{Q8A} had a mass of 543, 486 and 368 kDa, respectively. So, the dodecamer-hexamer equilibrium previously observed in *HpDnaB* MALS measurement was clearly affected in mutants *HpDnaB* ^{$\Delta 9$} , *HpDnaB*^{Q8A}, *HpDnaB*^{E80A} but not in *HpDnaB*^{L4A}. The oligomeric states of wild-type and mutants *HpDnaBs* were then compared using chemical cross-linking (Supplementary Figure S6). *HpDnaB* in complex with AMPPNP and ssDNA, which forms hexamers (see after), was used as a control. In the presence of cross-linking agent, *HpDnaB* and *HpDnaB*^{L4A} formed assemblies of the same size and larger than a hexamer, while the mutants *HpDnaB* ^{$\Delta 9$} , *HpDnaB*^{Q8A} and to a less extent *HpDnaB*^{E80A}

Table 2. Summary of the oligomeric states of *HpDnaB* and associated complexes studied

	MALS		SAXS		
	MW (kDa)	<i>HpDnaB</i> oligomer	R _g (Å)	D _{max} (Å)	<i>HpDnaB</i> oligomer
His- <i>HpDnaB</i>	687	12	67.4	234	12
<i>HpDnaB</i>	585	6–12	66.7	228	12
<i>HpDnaB</i> + ATP	546	6–12	67.9	234	12
<i>HpDnaB</i> + ATP + ssDNA	553	6–12	68.0	235	12
<i>HpDnaB</i> + AMPPNP	500	6–12	67.6	234	12
<i>HpDnaB</i> + AMPPNP + ssDNA	310	6	47.0	146	6
<i>HpDnaB</i> + <i>HpDnaG</i> ^{HBD}	345	6 (+ x <i>HpDnaG</i> ^{HBD})	55.0	193	6 (+ 3 <i>HpDnaG</i> ^{HBD})
Theoretical values[#]	MW (kDa)		R_g (Å)	Diameter (Å)	
<i>HpDnaB</i> hexamer	334	6	46.8	147	6
<i>HpDnaB</i> dodecamer	669	12	65.5	207	12
<i>HpDnaB</i> ₆ • <i>HpDnaG</i> ^{HBD} ₃ [‡]	385		50.7	152	

MW: Molecular weight; x = 1, 2 or 3; [‡] Model obtained using *BstDnaB*₆•*BstDnaG*^{HBD}₃ structure (pdb code 2R6A) as template, [#] Theoretical R_g and diameter values were obtained using CRYSOLOG (49).

formed predominantly hexamers (Supplementary Figure S6). These results are in agreement with MALS measurements. Collectively these experiments demonstrate that mutations targeting the interface seen in the crystal structure affect dodecamer formation of *HpDnaB*. We concluded that the structure of *HpDnaB* in solution is similar to the one seen in the crystal structure and that α1 is essential for hexamer–hexamer assembly. However in the absence of the His-tag, *HpDnaB* probably adopts multiple conformations, some of which might trigger dissociation of the two hexamers. Our data suggest that the N-terminal His-tag reduces the dynamics of the protein, possibly explaining that only the His-tagged protein crystallized. This is also supported by the observation that the values of R_g derived from the SAXS are more stable across the peak of His-*HpDnaB* than across the peak of *HpDnaB* (Supplementary Figure S1).

HpDnaB binds ssDNA but not dsDNA

Since DnaBs are loaded onto ssDNA at replication forks and *HpDnaB* is able to translocate in the 5′ to 3′ direction (40), we investigated the fate of the double ring on DNA. Using fluorescence anisotropy, we found that *HpDnaB* interacts with a 20mer polydT (20dT) ssDNA oligonucleotide in the presence of ATP or the non-hydrolysable ATP analogue, AMPPNP, with dissociation constants (K_D) of 15 ± 1 nM and 3.1 ± 0.1 nM, respectively, but not in the absence of nucleotide (Figure 3A). In contrast, no or very weak binding was observed with dsDNA even in the presence of nucleotides (Figure 3A). Increasing the length of the ssDNA oligonucleotide (Figure 3B) barely increased the binding affinity in the presence of AMPPNP (K_D = 1.8 ± 0.2 nM) and instead reduced the binding affinity of *HpDnaB* in the presence of ATP (K_D = 70 ± 6 nM). The binding of *HpDnaB* to ssDNA was found to be cooperative in the presence of AMPPNP, but not in the presence of ATP (Supplementary Table S1), indicating that the DNA binding mode may be affected by the nucleotide-bound state of *HpDnaB*.

Next, the complexes formed by *HpDnaB* and ssDNA were analyzed by SEC-MALS in the presence of either ATP or AMPPNP. The elution profile and the estimated molecular weight of *HpDnaB* in the presence and absence of

75 μM ssDNA were very similar when 0.5 mM ATP was added to the buffer (Figure 3C). In both cases, the mass of *HpDnaB* derived from the MALS was around 550 kDa thus corresponding to the mostly dodecameric *HpDnaB* (Figure 3B). When SAXS data were collected on the same sample and in the same size exclusion chromatography conditions, the R_gs of *HpDnaB*•ATP and *HpDnaB*•ATP•20dT samples were 67.9 Å and 68.0 Å, respectively (Supplementary Table S2). Together these data suggest that the complex *HpDnaB*•ATP•20dT is a dodecamer. In contrast, in the presence of AMPPNP, the addition of 45 μM of 20dT to the sample resulted in the reduction of the peak corresponding to the *HpDnaB* dodecamer and the appearance of second and third peaks at later elution volumes. The second peak contained *HpDnaB* and had a molecular weight of 310 kDa (Figure 3D). The addition of 75 μM of 20dT to the *HpDnaB*•AMPPNP sample resulted in the complete disappearance of the dodecamer peak. Interestingly, the R_g derived from the SAXS analysis of the sample corresponding to the second peak containing the *HpDnaB*•AMPPNP•20dT was 47.0 Å, corresponding to a hexamer while in the presence of AMPPNP alone, the R_g of *HpDnaB* was 67.6 Å and corresponds to a dodecamer (Table 2).

Altogether, our results show that *HpDnaB* binds ssDNA with high affinity in the presence of nucleotides both as a hexamer and a dodecamer. *HpDnaB* in its ATP-bound state remains a dodecamer when binding to ssDNA, while when ATP hydrolysis is prevented by incubating the complex with AMPPNP, *HpDnaB* binds ssDNA cooperatively as a hexamer (Table 2 and Supplementary Table S1) (5,12).

HpDnaB interaction with *HpDnaG*^{HBD} dissociates double rings into hexamers

We investigated the oligomeric state of *HpDnaB* during formation of the *HpDnaB*•*HpDnaG* complex. The crystal structure of the *BstDnaB*•*BstDnaG*^{HBD} revealed that each of the three DnaG^{HBD} engages the NTD dimer via interactions with the head domains of *BstDnaB* (4). To gain insight into the assembly of the primosome of *H. pylori* we analyzed the *HpDnaB*•*HpDnaG*^{HBD} complex obtained by mixing the two proteins together using SEC-MALS and

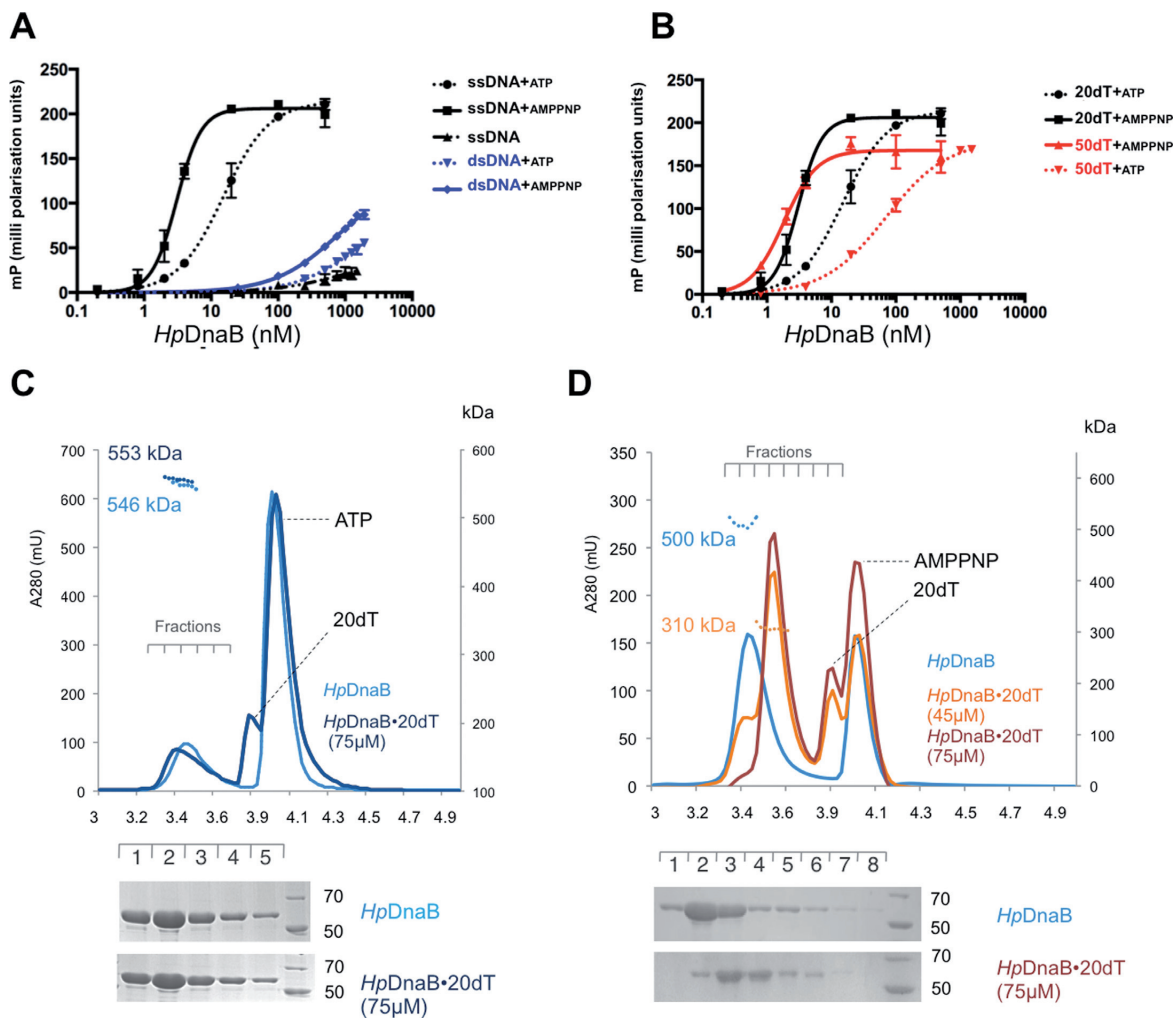


Figure 3. Effects of ATP and AMPPNP on the *HpDnaB* interaction with DNA. (A) Fluorescence anisotropy measurements of *HpDnaB* binding to either 5'-FAM labeled 20mer ssDNA (20dT) or dsDNA in the presence and absence of nucleotide (ATP and AMPPNP). The curves represent the mean of three independent experiments. (B) Similar experiments performed with 5'-FAM labeled 50mer ssDNA (50dT). (C) Size exclusion chromatograms (280 nm) of *HpDnaB* alone (light blue) or mixed with a 20dT oligonucleotide (75 μ M) and 5 mM ATP (dark blue). MALS weight-averaged molar masses are indicated as dotted lines. The SDS-PAGE analysis shows the *HpDnaB* protein elution fractions. (D) SEC-MALS experiment performed as in (C) except that *HpDnaB* was incubated with 5 mM AMPPNP and the running buffer contained 0.5 mM AMPPNP. Three samples were analyzed: *HpDnaB* (blue) and *HpDnaB* with 45 μ M (orange) or 75 μ M (red) 20dT. The SDS-PAGE analysis shows that *HpDnaB* elutes slightly later in the presence of ssDNA and AMPPNP.

SAXS. SEC experiments showed that this complex eluted as two peaks and SDS-PAGE showed that the first peak contained both *HpDnaB* and *HpDnaG*^{HBD}, while the second contained exclusively *HpDnaG*^{HBD} (Figure 4A). Measurements of the molecular weight by MALS indicated that the peak corresponding to the *HpDnaB*•*HpDnaG*^{HBD} complex had a molecular mass of approximately 345 kDa, suggesting that this peak corresponds to hexameric rather than dodecameric *HpDnaB* (the theoretical mass of the hexamer is 334 kDa) with one or perhaps more *HpDnaG*^{HBD} (17 kDa) molecules bound. *HpDnaG*^{HBD} alone had a mass of around 35 kDa corresponding to a dimer as previously observed (43).

SAXS data corroborated these observations with the complex *HpDnaB*•*HpDnaG*^{HBD} having an Rg of 55.0 Å, much lower than a dodecamer (66.7 Å) but greater than a theoretical *HpDnaB* hexamer (46.8 Å) (Supplementary

Table S3). We then modeled the *HpDnaB*₆•*HpDnaG*^{HBD}₃ complex using the crystal structures of His-*HpDnaB* and *HpDnaG*^{HBD} and the *BstDnaB*₆•*BstDnaG*^{HBD}₃ crystal structure as a template (43). A very good fit of the experimental SAXS profile was obtained with the theoretical curve derived from the *HpDnaB*₆•*HpDnaG*^{HBD}₃ model ($\chi^2 = 5.8$) compared with the curve derived from the *HpDnaB* dodecamer ($\chi^2 = 36.8$) (Figure 4B). When using MES, the fit was significantly improved with a mixture of 89% of the *HpDnaB*₆•*HpDnaG*^{HBD}₃ complex and of 11% of *HpDnaB* dodecamer ($\chi^2 = 3.5$; Figure 4B), suggesting that the peak observed in SEC-MALS contains both uncomplexed dodecameric *HpDnaB* and hexameric *HpDnaG*^{HBD}-bound *HpDnaB*. Together, these data reveal that formation of the *H. pylori* primosome results in the dissociation of the helicase dodecamer into hexamers that can interact with up

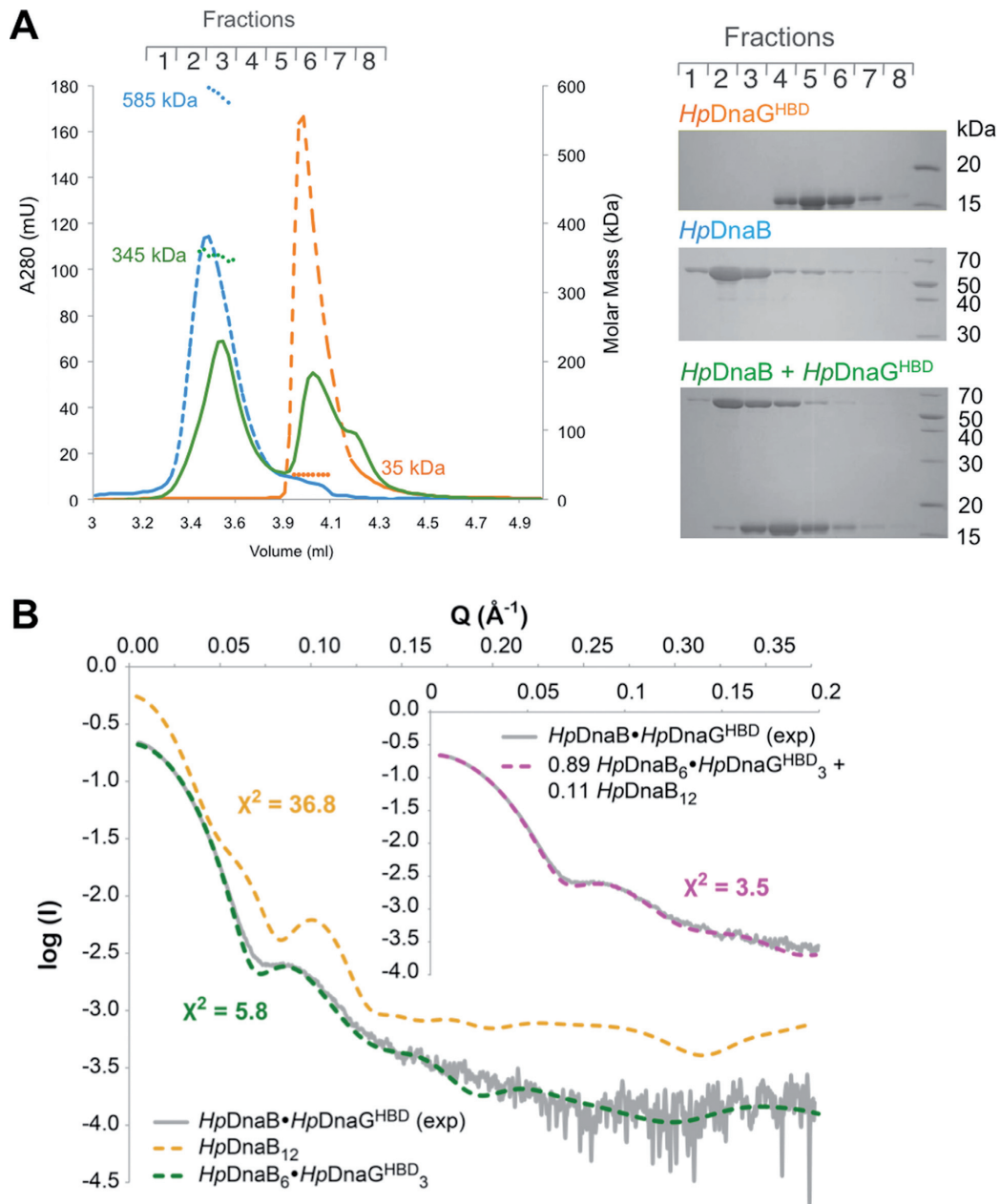


Figure 4. Interaction of *HpDnaG^{HBD}* dissociates *HpDnaB* dodecamer. (A) Size-exclusion chromatograms (280 nm) of *HpDnaB* (light blue), *HpDnaG^{HBD}* (orange) or the reconstituted primosome *HpDnaB + HpDnaG^{HBD}* (green). MALS weight-averaged molar masses measurements are indicated as dotted lines. The SDS-PAGE analysis shows the proteins contained in the elution fractions. (B) Experimental scattering curve of the *HpDnaB*•*HpDnaG^{HBD}* peak (grey) compared to theoretical curves of *HpDnaB* dodecamer (orange), and model of the *HpDnaB₆*•*HpDnaG^{HBD}₃* (green). The inset shows the improved fit obtained from a MES containing a mixture of *HpDnaB₆*•*HpDnaG^{HBD}₃* and *HpDnaB* dodecamer models.

to three *HpDnaG^{HBD}* in a conformation similar to the crystal structure of *BstDnaB₆*•*BstDnaG^{HBD}₃* (Table 2).

HpDnaG^{HBD} interacts with *HpDnaB* and ssDNA to prime helicase and ATPase activities

A major role of primosome formation is the modulation of primase and helicase activities (14,44–46). As for other DnaBs, the ATPase activity of *HpDnaB* has previously been shown to be stimulated by *HpDnaG^{HBD}* or by ssDNA

(17,40). We have thus investigated the effect of *HpDnaG^{HBD}* on the ability of *HpDnaB* to hydrolyse ATP, bind DNA and unwind forked DNA duplexes. *HpDnaB* had an ATPase activity of $0.04 \pm 0.03 \mu\text{M ATP}\cdot\text{s}^{-1}$ at a concentration of 500 nM (Supplementary Table S4). As expected, addition of *HpDnaG^{HBD}* or ssDNA increased the ATPase activity of *HpDnaB* approximately three-fold (Supplementary Table S4). The activity was increased almost eight-fold by adding both ssDNA and *HpDnaG^{HBD}*, suggesting that the

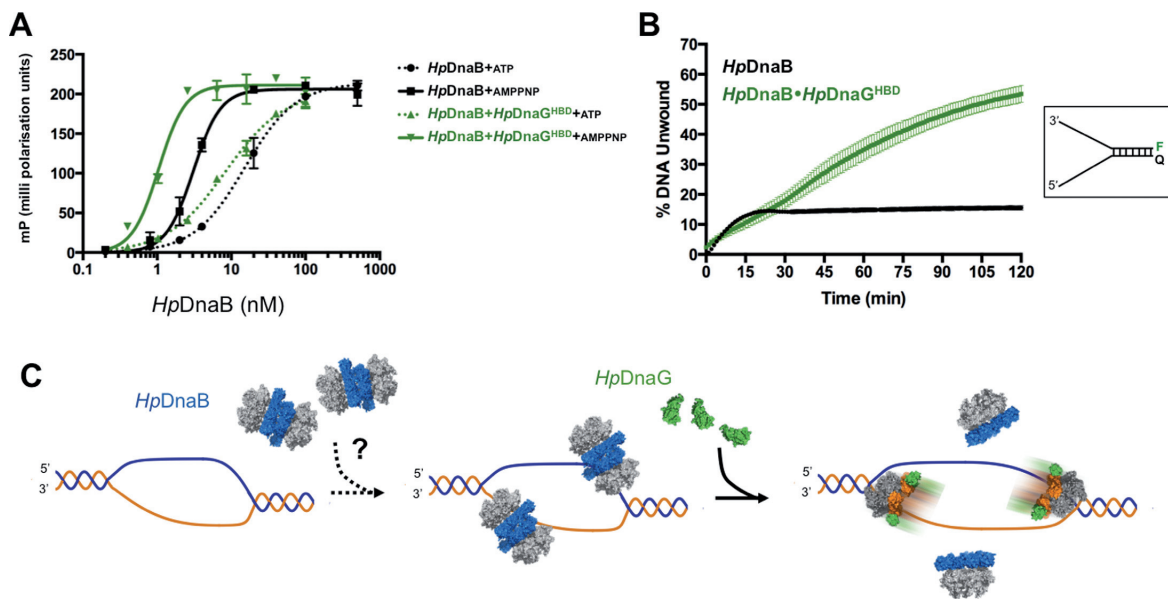


Figure 5. *HpDnaG^{HBD}* primes *HpDnaB* activities. (A) Single-stranded DNA binding activity of *HpDnaB* and *HpDnaB*•*HpDnaG^{HBD}* complex in the presence of ATP (0.5 mM) or AMPPNP (0.5 mM) as measured by fluorescence anisotropy. (B) Unwinding of a fluorophore (F)/quencher (Q)-labeled forked DNA substrate (shown schematically in inset) by *HpDnaB* (black) or *HpDnaB*•*HpDnaG^{HBD}* (green). The curve represents the mean of three independent experiments and the standard deviations are indicated for each measurement as error bars. (C) Schematic illustration of the proposed model of *HpDnaG^{HBD}* activation of the double hexamer on ssDNA during replication initiation in *Helicobacter pylori*.

ternary complex is the most active for ATP hydrolysis with a hydrolysis rate of $0.35 \pm 0.08 \mu\text{M ATP}\cdot\text{s}^{-1}$ (Supplementary Table S4).

The binding affinity of these proteins to ssDNA was then measured by fluorescence anisotropy. The *HpDnaB*•*HpDnaG^{HBD}* complex exhibited a higher affinity for ssDNA than *HpDnaB* alone in the presence of ATP ($K_D = 8.3 \pm 0.6 \text{ nM}$ versus $15 \pm 1 \text{ nM}$) or AMPPNP ($K_D = 1.0 \pm 0.1$ versus $3.1 \pm 0.1 \text{ nM}$) (Figure 5A). *HpDnaG^{HBD}* thus increases the affinity of *HpDnaB* for ssDNA in the presence of nucleotides. Interestingly, *HpDnaG^{HBD}* alone interacts with ssDNA, but with an affinity in the micromolar range (Supplementary Figure S7). Translocation of the double hexamer on ssDNA poses several mechanistic problems. (i) *HpDnaB* unwinds dsDNA in the 5' to 3' direction, so only one hexamer can be active; (ii) the second (inactive) hexamer would considerably hinder the first one's activity by binding non-specifically to ssDNA via its DNA binding loops; (iii) an inactive hexamer would represent a significant inert mass to be translocated by the active one. This suggests that for efficient helicase activity, a double hexamer would have to be separated to release the correctly positioned *HpDnaB* hexamer. To test this hypothesis we analyzed the helicase activity of *HpDnaB* in the presence and absence of *HpDnaG^{HBD}*. We used a fluorescence-based assay similar to the one used previously to measure *EcDnaB*•*EcDnaC* helicase activity (10). A forked DNA substrate was formed by annealing an oligonucleotide bearing a 3' ssDNA tail and a fluorescein label at its 5' end with a complementary oligonucleotide bearing a 5' ssDNA tail and a quenching dye (Dabcyl) at its 3' end. As seen in Figure 5B, double hexameric *HpDnaB* displayed weak DNA unwinding activity, rapidly reaching a maximum of only 10% DNA unwound. In con-

trast, in complex with *HpDnaG^{HBD}*, *HpDnaB* was able to unwind dsDNA, reaching 60% DNA unwound. The initial rates of DNA unwinding were very similar in both cases, confirming that *HpDnaB* is indeed an active DNA helicase, but the activity of the enzyme was drastically affected by the presence of the primase *HpDnaG^{HBD}*. No more activity was detected for *HpDnaB* alone after approximately 20 min, whereas in the presence of *HpDnaG^{HBD}*, helicase activity could still be detected 2 h after the start of the reaction.

DISCUSSION

DNA replication relies on replicative helicases to initiate progression of replication forks and to anchor the primase. While most DnaBs characterized to date have been reported to be hexameric, we previously found that *HpDnaB* from *H. pylori* was able to form double hexamers, thus questioning how this structure would assemble and function on replication forks. In this study, we have addressed the structural mechanisms of *HpDnaB* activation during primosome formation on ssDNA.

The crystal structure of the *HpDnaB* dodecamer solved here provides significant information on *HpDnaB* dodecamer compared to our previous nsEM reconstruction (17). It reveals that the NTDs of the helicase are arranged in dilated collars that self-interact in a stack-twisted mode to generate a head-to-head double hexamer. The dodecamer assembly stabilizes the NTD-ring conformation but allows for the CTD-ring to remain sufficiently dynamic to accept structural changes associated with ATP hydrolysis or ssDNA binding. Our work establishes that if the topology of this bacterial dodecameric helicase (i.e. NTD–NTD in-

teractions) is reminiscent of Mcm2–7 and some MCMs, the structural and molecular determinants of the hexamers association are different. In *Methanobacterium thermoautotrophicum* MCM (*MtMCM*) (47), dodecamerization relies on interactions between loops that coordinate Zn atoms in the *MtMCM* B domain (Supplementary Figure S8). The structural basis for dodecamer formation is completely different in *HpDnaB* (Figure 2) and thus rules out a common evolutionary structural motif. Nevertheless, it is interesting to note that, despite diverse evolutionary routes, dodecameric replicative helicases might exist in all kingdoms of life, suggesting that some selective advantages might be associated with this organization (48).

It is as yet unclear how *HpDnaB* is loaded onto the replication forks and if an additional factor is required for this process (Figure 5C). In this regard, we found that the double hexameric *HpDnaB* does not interact efficiently with dsDNA contrary to ssDNA, suggesting that unwound origins of replication might be better substrates (Figure 3A). The *HpDnaB* dodecamer positions two hexamers in opposite directions but the two rings cannot unwind dsDNA simultaneously in the 5' to 3' direction if a single strand passes through the double ring. Accordingly, the use of a longer ssDNA oligonucleotide did not increase the affinity of the *HpDnaB* for DNA, suggesting that only one of the two rings of the dodecamer is efficiently bound to ssDNA (Figure 3B).

We found that the *HpDnaB* dodecamer is indeed poorly active for DNA unwinding and that the protein unwinds dsDNA as a hexamer, similar to other DnaBs (12). ATP hydrolysis and ssDNA binding are not sufficient to separate the two rings (Figure 3C and Table 2), suggesting that an additional factor is required to initiate DNA unwinding at the replication fork. Our study suggests that the primase might be such a factor since the interaction with *HpDnaG*^{HBD} separates the *HpDnaB* double hexamer (Figures 3D, 5C and Table 2). This can be explained by the crystal structure of His-*HpDnaB* in which the dodecamer interface mediated by NTD collars overlaps the primase interaction site described in *BstDnaB* (4) and conserved in *HpDnaB* (43). *HpDnaG* association with *HpDnaB* not only separates the dodecamer, but also increases its ATPase activity, its affinity for ssDNA and its DNA unwinding capacity. These findings suggest that the separation of the dodecamer by *HpDnaG* drives DNA unwinding by one hexamer in the 5' to 3' direction allowing replication to initiate (Figure 5C). Mechanistically, the separation of the two helicase rings by interaction with the primase might also be relevant to other systems such as archaeal or eukaryotic systems, which also exploit double ring helicases during replication initiation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

Thanks are due to G. Fox and A. Thompson on beamline PROXIMA2, J. Perez and P. Roblin for help on the SWING beamline from SOLEIL and ESRF staff on beamline ID14EH4. We acknowledge the use of the UMS3444

Protein Science platform, the platforms of the Grenoble Instruct Center (ISBG : UMS 3518).

FUNDING

This work was supported by the ATIP-Avenir program (L.T. and J.T.) of CNRS and by a fellowship Region Rhône-Alpes from the program CIBLE 2011 (A.B.). The Grenoble Instruct Center is supported by FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble PSB. The EM facility is supported by the Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER, the Centre National de la Recherche Scientifique (CNRS), the CEA, the University of Grenoble, EMBL, and the GIS-Infrastructures en Biologie Sante et Agronomie (IBISA). Funding for open access charge: CNRS.

Conflict of interest statement. None declared.

REFERENCES

- Costa, A., Hood, I.V. and Berger, J.M. (2013) Mechanisms for initiating cellular DNA replication. *Annu. Rev. Biochem.*, **82**, 25–54.
- Bailey, S., Eliason, W.K. and Steitz, T.A. (2007) The crystal structure of the *Thermus aquaticus* DnaB helicase monomer. *Nucleic Acids Res.*, **35**, 4728–4736.
- Wang, G., Klein, M.G., Tokonzaba, E., Zhang, Y., Holden, L.G. and Chen, X.S. (2008) The structure of a DnaB-family replicative helicase and its interactions with primase. *Nat. Struct. Mol. Biol.*, **15**, 94–100.
- Bailey, S., Eliason, W.K. and Steitz, T.A. (2007) Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science*, **318**, 459–463.
- Strycharska, M.S., Arias-Palomo, E., Lyubimov, A.Y., Erzberger, J.P., O'Shea, V.L., Bustamante, C.J. and Berger, J.M. (2013) Nucleotide and partner-protein control of bacterial replicative helicase structure and function. *Mol. Cell*, **52**, 844–854.
- Davey, M.J. and O'Donnell, M. (2003) Replicative helicase loaders: ring breakers and ring makers. *Curr. Biol.*, **13**, R594–R596.
- Soultanas, P. (2012) Loading mechanisms of ring helicases at replication origins. *Mol. Microbiol.*, **84**, 6–16.
- Kobori, J.A. and Kornberg, A. (1982) The *Escherichia coli* dnaC gene product. III. Properties of the dnaB-dnaC protein complex. *J. Biol. Chem.*, **257**, 13770–13775.
- Karplus, P.A. and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science*, **336**, 1030–1033.
- Arias-Palomo, E., O'Shea, V.L., Hood, I.V. and Berger, J.M. (2013) The bacterial DnaC helicase loader is a DnaB ring breaker. *Cell*, **153**, 438–448.
- Liu, B., Eliason, W.K. and Steitz, T.A. (2013) Structure of a helicase-helicase loader complex reveals insights into the mechanism of bacterial primosome assembly. *Nat. Commun.*, **4**, 2495.
- Itsathitphaisarn, O., Wing, R.A., Eliason, W.K., Wang, J. and Steitz, T.A. (2012) The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell*, **151**, 267–277.
- Makowska-Grzyska, M. and Kaguni, J.M. (2010) Primase directs the release of DnaC from DnaB. *Mol. Cell*, **37**, 90–101.
- Bird, L.E., Pan, H., Soultanas, P. and Wigley, D.B. (2000) Mapping protein-protein interactions within a stable complex of DNA primase and DnaB helicase from *Bacillus stearothermophilus*. *Biochemistry*, **39**, 171–182.
- Robinson, A., Causer, R.J. and Dixon, N.E. (2012) Architecture and conservation of the bacterial DNA replication machinery, an underexploited drug target. *Curr. Drug Targets*, **13**, 352–372.
- Soni, R.K., Mehra, P., Mukhopadhyay, G. and Dhar, S.K. (2005) *Helicobacter pylori* DnaB helicase can bypass *Escherichia coli* DnaC function in vivo. *Biochem. J.*, **389**, 541–548.
- Stelter, M., Gutsche, I., Kapp, U., Bazin, A., Bajic, G., Goret, G., Jamin, M., Timmins, J. and Terradot, L. (2012) Architecture of a dodecameric bacterial replicative helicase. *Structure*, **20**, 554–564.

18. Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.
19. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.
20. Collaborative Computational Project-4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
21. McCoy, A.J., Storoni, L.C. and Read, R.J. (2004) Simple algorithm for a maximum-likelihood SAD function. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 1220–1228.
22. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
23. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1948–1954.
24. Kashav, T., Nitharwal, R., Abdulrehman, S.A., Gabdoulkhalov, A., Saenger, W., Dhar, S.K. and Gourinath, S. (2009) Three-dimensional structure of N-terminal domain of DnaB helicase and helicase-primase interactions in *Helicobacter pylori*. *PLoS One*, **4**, e7515.
25. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
26. Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 470–478.
27. Blanc, E., Roversi, P., Vornrhein, C., Flensburg, C., Lea, S.M. and Bricogne, G. (2004) Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2210–2221.
28. Ludtke, S.J., Baldwin, P.R. and Chiu, W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, **128**, 82–97.
29. Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.*, **142**, 334–347.
30. Heymann, J.B., Cardone, G., Winkler, D.C. and Steven, A.C. (2008) Computational resources for cryo-electron tomography in Bsoft. *J. Struct. Biol.*, **161**, 232–242.
31. Shaikh, T.R., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A. and Frank, J. (2008) SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nat. Protoc.*, **3**, 1941–1974.
32. van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R. and Schatz, M. (1996) A new generation of the IMAGIC image processing system. *J. Struct. Biol.*, **116**, 17–24.
33. David, G. and Perez, J. (2009) Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.*, **42**, 892–900.
34. Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
35. Schneidman-Duhovny, D., Hammel, M., Tainer, J.A. and Sali, A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.*, **105**, 962–974.
36. Pelikan, M., Hura, G.L. and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.*, **28**, 174–189.
37. Panuska, J.R. and Goldthwait, D.A. (1980) A DNA-dependent ATPase from T4-infected *Escherichia coli*. Purification and properties of a 63,000-dalton enzyme and its conversion to a 22,000-dalton form. *J. Biol. Chem.*, **255**, 5208–5214.
38. Biswas, T. and Tsodikov, O.V. (2008) Hexameric ring structure of the N-terminal domain of *Mycobacterium tuberculosis* DnaB helicase. *FEBS J.*, **275**, 3064–3071.
39. Lo, Y.H., Tsai, K.L., Sun, Y.J., Chen, W.T., Huang, C.Y. and Hsiao, C.D. (2009) The crystal structure of a replicative hexameric helicase DnaC and its complex with single-stranded DNA. *Nucleic Acids Res.*, **37**, 804–814.
40. Soni, R.K., Mehra, P., Choudhury, N.R., Mukhopadhyay, G. and Dhar, S.K. (2003) Functional characterization of *Helicobacter pylori* DnaB helicase. *Nucleic Acids Res.*, **31**, 6828–6840.
41. Cheung, K.L., Huen, J., Kakihara, Y., Houry, W.A. and Ortega, J. (2010) Alternative oligomeric states of the yeast Rvb1/Rvb2 complex induced by histidine tags. *J. Mol. Biol.*, **404**, 478–492.
42. Lyubimov, A.Y., Costa, A., Bleichert, F., Botchan, M.R. and Berger, J.M. (2012) ATP-dependent conformational dynamics underlie the functional asymmetry of the replicative helicase from a minimalist eukaryote. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 11999–12004.
43. Abdul Rehman, S.A., Verma, V., Mazumder, M., Dhar, S.K. and Gourinath, S. (2013) Crystal structure and mode of helicase binding of the C-terminal domain of primase from *Helicobacter pylori*. *J. Bacteriol.*, **195**, 2826–2838.
44. Johnson, S.K., Bhattacharyya, S. and Griep, M.A. (2000) DnaB helicase stimulates primer synthesis activity on short oligonucleotide templates. *Biochemistry*, **39**, 736–744.
45. Bhattacharyya, S. and Griep, M.A. (2000) DnaB helicase affects the initiation specificity of *Escherichia coli* primase on single-stranded DNA templates. *Biochemistry*, **39**, 745–752.
46. Wu, C.A., Zechner, E.L. and Marians, K.J. (1992) Coordinated leading- and lagging-strand synthesis at the *Escherichia coli* DNA replication fork. I. Multiple effectors act to modulate Okazaki fragment size. *J. Biol. Chem.*, **267**, 4030–4044.
47. Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M. and Chen, X.S. (2003) The structure and function of MCM from archaeal *M. Thermoautotrophicum*. *Nat. Struct. Biol.*, **10**, 160–167.
48. Sclafani, R.A., Fletcher, R.J. and Chen, X.S. (2004) Two heads are better than one: regulation of DNA replication by hexameric helicases. *Genes Dev.*, **18**, 2039–2045.
49. Svergun, D., Barberato, C. and Koch, M.H.J. (1995) CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.

Résumé

Durant la réplication du chromosome bactérien, le désappariement du double brin d'ADN est réalisé par l'hélicase hexamérique DnaB. Chez *Escherichia coli*, le positionnement de l'hexamère de DnaB sur l'ADN simple brin dans le sens 5'-3' est permis par le facteur de chargement. La primase DnaG interagit ensuite avec l'hélicase pour former le primosome. Chez *Helicobacter pylori*, aucun facteur de chargement n'a été identifié, ce qui est également le cas pour la majorité des espèces bactériennes. De plus, DnaB d'*H. pylori* (HpDnaB) peut compléter des souches mutantes d'*E. coli* dnaB^{ts} et dnaC^{ts} suggérant que HpDnaB peut jouer le rôle des deux protéines. Pour mieux comprendre le mode d'action de HpDnaB, nous avons résolu sa structure cristallographique à une résolution de 6.7 Å. Celle-ci révèle que la protéine s'assemble en dodécamère, formé par deux hexamères interagissant par leurs domaines N-terminaux (NTD). Nos expériences en diffusion des rayons X aux petits angles (SAXS) montrent que le dodécamère de HpDnaB adopte une conformation modifiée et dynamique en solution. Nous avons ensuite étudié la structure de HpDnaB après interaction avec $\text{HpDnaG}^{\text{HBD}}$ et/ou l'ADN simple brin par chromatographie d'exclusion stérique couplée à la diffusion de la lumière multi-angles (SEC-MALS) et par SAXS. Ces expériences suggèrent qu'après interaction avec $\text{HpDnaG}^{\text{HBD}}$, le double hexamère est dissocié en simples hexamères formant un complexe avec $\text{HpDnaG}^{\text{HBD}}$. De plus, HpDnaB forme des hexamères avec l'ADN simple brin en présence d'AMP-PNP. L'ensemble de nos résultats suggère que la formation du primosome d'*H. pylori* conduit à la dissociation du dodécamère en deux complexes $\text{HpDnaB}_6 \bullet \text{HpDnaG}_3$.

Mots clés: Réplication de l'ADN, hélicase, *Helicobacter pylori*, cristallographie aux rayons X

Abstract

During bacterial chromosomal replication, unwinding of double stranded DNA is performed by the hexameric helicase DnaB. In *Escherichia coli*, the positioning of DnaB hexamers onto replication forks in the 5' to 3' direction is dedicated by helicase loader. DnaB then interacts with the DnaG primase helicase binding domain (DnaG^{HBD}) to form the primosome. *Helicobacter pylori* does not encode for a DnaC homologue, which is also the case of most bacterial species. Moreover, *H. pylori* DnaB (HpDnaB) could complement two temperature-sensitive mutants of *E. coli* dnaB^{ts} and dnaC^{ts} , suggesting that the HpDnaB was able to bypass DnaC in these cells. To gain insights into HpDnaB mode of activation, we have solved the crystal structure of HpDnaB at 6.7 Å resolution. The structure reveals a novel dodecameric organisation where HpDnaB assembles as planar stack-twisted double hexamers via N-terminal domain (NTD)-rings interactions. Small angle X-ray scattering analysis (SAXS) demonstrates that HpDnaB adopts a modified and dynamic structure in solution but maintains dodecameric architecture. We have then investigated the structure of HpDnaB upon interaction with $\text{HpDnaG}^{\text{HBD}}$ and/or ssDNA using size exclusion chromatography coupled to multi-angle light scattering and SAXS. These experiments show that upon interaction with $\text{HpDnaG}^{\text{HBD}}$, HpDnaB double hexamer dissociates into single hexamers to form a complex with $\text{HpDnaG}^{\text{HBD}}$. Moreover, we found that HpDnaB also forms hexamers in complex with ssDNA in the presence of AMP-PNP. Collectively, these data suggest that primosome assembly in *H. pylori* results in the dissociation of the double hexamer into two $\text{HpDnaB}_6 \bullet \text{HpDnaG}_3$ sister primosomes.

Keywords: DNA replication, helicase, *Helicobacter pylori*, X-ray crystallography