



HAL
open science

Etude structurale du mécanisme d'échange de chaînes des dimères de la protéine HU d'E. coli

Rémy Le Meur

► **To cite this version:**

Rémy Le Meur. Etude structurale du mécanisme d'échange de chaînes des dimères de la protéine HU d'E. coli. Sciences agricoles. Université d'Orléans, 2015. Français. NNT : 2015ORLE2004 . tel-01424222

HAL Id: tel-01424222

<https://theses.hal.science/tel-01424222>

Submitted on 2 Jan 2017

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

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

ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

Centre de Biophysique Moléculaire, CNRS Orléans

THÈSE présentée par :
Rémy LE MEUR

Soutenue le : **23 Janvier 2015**

pour obtenir le grade de : **Docteur de l'université d'Orléans**
Discipline/ Spécialité : **Biologie Structurale et Fonctionnelle**

**Étude structurale du mécanisme d'échange de chaînes
des dimères de la protéine HU d'*E. coli***

THÈSE codirigée par :

Céline LANDON
Bertrand CASTAING

Chargée de Recherche, Centre de Biophysique Moléculaire
Directeur de Recherche, Centre de Biophysique Moléculaire

et coencadrée par :

Karine LOTH

Maître de Conférences, Université d'Orléans

RAPPORTEURS :

Bernhard BRUTSCHER
Jacques OBERTO

Directeur de Recherche, Institut de Biologie Structurale
Chargé de Recherche, Institut de Génétique et Microbiologie

JURY :

Frédéric BOCCARD
Franck BRIGNOLAS
Bernhard BRUTSCHER
Bertrand CASTAING
Céline LANDON
Karine LOTH
Jacques OBERTO
Nicolas WOLFF

Directeur de Recherche, Centre de Génétique Moléculaire
Professeur des universités, Université d'Orléans, Président du jury
Directeur de Recherche, Institut de Biologie Structurale
Directeur de Recherche, Centre de Biophysique Moléculaire
Chargée de Recherche, Centre de Biophysique Moléculaire
Maître de Conférences, Université d'Orléans
Chargé de Recherche, Institut de Génétique et Microbiologie
Directeur de Recherche, Institut Pasteur

Remerciements

Ce travail de thèse a été réalisé au Centre de Biophysique Moléculaire (CBM), au CNRS d'Orléans, et a été financé par le ministère de l'enseignement supérieur et de la recherche. Je remercie le Dr Jean-Claude Beloeil et le Dr Eva Jakab Toth de m'avoir accueilli au sein du laboratoire.

Je remercie le Dr Bernard Brutscher et le Dr Jacques Oberto d'avoir accepté d'être les rapporteurs de ma thèse. Je remercie également le Dr Frédéric Boccard, le Professeur Franck Brignolas, et le Dr Nicolas Wolff pour leur participation à mon jury de thèse.

Je tiens à remercier tous ceux qui ont participé de près ou de loin à ce travail :

Céline Landon et Bertrand Castaing, co-directeurs de ma thèse, pour m'avoir accueilli dans leurs équipes respectives. Pour m'avoir fait confiance et avoir encouragé mes initiatives. Pour leur soutien permanent et sans faille, et pour avoir toujours été à l'écoute et de bon conseil.

Karine Loth, pour m'avoir encadré au cours de cette thèse. Pour m'avoir appris et aidé dans tous les aspects de la RMN abordés dans ce travail, de l'attribution au calcul de structure et en passant par la dynamique. Pour m'avoir supporté dans son bureau pendant 3 ans, pour son soutien et ses conseils.

Françoise Culard, pour son aide inestimable dans le projet Fpg-HU. Pour avoir partagé son labo avec moi depuis le stage de master et m'avoir transmis une partie de son savoir-faire sur la caractérisation des complexes ADN-protéine en biochimie.

Franck Coste, pour m'avoir initié à la cristallographie des protéines (malgré le fait que je fasse de la RMN). Pour m'avoir fait découvrir les joies d'un « run » synchrotron. Pour ses conseils et pour son aide de la cristallogénèse à la collecte de données jusqu'à l'affinement des modèles. Et pour son intérêt pour mon travail et son avancement : « On en est où ? »

Norbert Garnier, pour la simulation en dynamique moléculaire en gros grain. De la mise en place, à la production jusqu'à l'analyse. Avec deux protéines et de multiples températures, cela représente pas mal de temps machine et humain, et surtout d'usure de la touche « Enter ».

Virginie Nadan, pour sa contribution dans le projet Fpg-HU avec les plus beaux gels retards que j'ai pu voir jusqu'à aujourd'hui.

Marianne Désir, pour sa contribution, au cours de ses stages de master, aux travaux préliminaires de RMN sur HU α_2 .

Nos collaborateurs de l'équipe de spectrométrie de masse du CBM : David Ginguène, doctorant de la même promotion que moi, pour la mise au point instrumentale et l'enregistrement des spectres de masse, en condition native, des (trop nombreux) échantillons de HU et Martine Cadène sa directrice de thèse et chef d'équipe.

Nos collaborateurs de l'Institut de Chimie des Substances Naturelles (ICSN) de Gif-sur-Yvette (TGIR RMN), qui m'ont accueilli lorsque nous n'avions plus de spectromètre RMN en fonctionnement. Nelly Morellet pour son accueil et son aide dans la mise en place des expériences. Ewen Lescop pour son aide et ses précieuses suggestions, en particulier l'utilisation des zz - ^{15}N -HSQC pour la caractérisation cinétique de l'échange.

Nos collaborateurs du laboratoire Polymères Biopolymères Surfaces (PBS) à Rouen : Annick Schaumann pour les tests préliminaires des anticorps anti-HU sur les biofilms bactériens. Emmanuelle De et Thierry Jouenne pour avoir accepté de travailler avec nous sur ce projet.

Je remercie Philippe Gaudu (INRA Jouy-en-josas) et Anne Bravard (CEA Fontenay-aux-roses) pour la souche MG1363 de *L. lactis* de laquelle nous avons cloné un gène codant pour HU.

Je tiens à remercier chaleureusement tous les membres, présents ou anciens, de mes équipes d'accueil, « RMN des biomolécules : structure, dynamique et interaction » et « Réparation de l'ADN : approches structurales et fonctionnelles », qui m'ont accompagné au cours de cette thèse : Artur Biela, Julie Berthault, Claudine Boiteux, Serge Boiteux, Abdennour Braka, Bertrand Castaing, Franck Coste, Kadija Essalhi, Norbert Garnier, Stéphane Gauffinont, Martine Guérin, Laurence Jouvensal, Françoise Keita-Paquet, Céline Landon, Karine Loth, Fanny Meindre, Hervé Meudal, Virginie Nadan, Charlotte Rieux et Violette Senille.

Ces trois années de thèses se sont écoulées dans le meilleur environnement possible tant du point de vue humain, notamment à savourer les pauses café et pauses déjeuner, que du point de vue scientifique, à apprendre et à progresser auprès de chacun d'entre vous, dans tous les domaines.

Je remercie Bertrand, Franck et Norbert pour les parties de ping-pong. Il faudrait équiper certains de crampons pour les glissades et d'autres de dragonnes pour les raquettes qui volent.

Je remercie également tous les doctorants et postdocs du laboratoire pour les bons moments passés au quotidien. Je souhaite du courage à ceux qui s'apprêtent à terminer leur thèse : Axelle, David et Safia. Également tous mes vœux de réussite aux nouvelles générations : Abdennour, Alba, Charlotte, Dounia. Merci pour votre amitié et les bons moments passés au laboratoire ou en dehors.

Je souhaite remercier Daniel Auguin, pour m'avoir fait découvrir et apprécier la biologie structurale dès les bancs de l'université. Je n'aurai peut-être pas suivi la même voie sans ton influence. La RMN n'est pas si terrifiante et inaccessible, même pour un biologiste.

Je souhaite enfin remercier ma famille (ma mère et ma sœur) et mes proches pour leur affection et leur soutien. Je ne suis pas arrivé jusqu'ici tout seul. Cette voie a demandé des sacrifices mais c'est celle que j'ai choisi et dans laquelle je compte poursuivre et m'épanouir.

Table des matières

Chapitre 1 : Étude bibliographique.....	17
1.1. La nécessité de compaction des génomes.....	17
1.2. Les acteurs protéiques de la compaction de l'ADN.....	19
1.2.1. Les Histones et les protéines de type Histone.....	19
1.2.2. Les Topoisomérases.....	20
1.3. La Compaction de l'ADN dans les 3 grands règnes du Vivant.....	23
1.3.1. La compaction de l'ADN chez les Eucaryotes.....	23
Généralités.....	23
Les Histones et le Nucléosome.....	26
Les protéines non histones.....	26
1.3.2. La compaction de l'ADN chez les Archées.....	30
Généralités.....	30
Les Histones d'Archées.....	31
Les Protéines non histones chez les Archées.....	33
Alba.....	33
Cren7.....	36
Sul7.....	36
CC1.....	36
MC1.....	37
HU.....	37
1.3.3. Compaction de l'ADN chez les Bactéries.....	38
Généralités.....	38
Les protéines architecturales du nucléoïde bactérien.....	40
SMC.....	40
Fis.....	42
H-NS.....	44
IHF.....	45
1.4. La protéine HU.....	47
1.4.1. Généralités.....	47
1.4.2. Les ligands de HU.....	47
1.4.3. Les fonctions biologiques de HU.....	47
1.4.4. Multiplicité des gènes codant pour HU.....	50
1.4.5. Régulation et fonctions des isoformes de HU d' <i>E.coli</i>	51
1.5. L'échange de chaînes, un mécanisme essentiel à la formation de l'hétérodimère de HU d' <i>E.coli</i>	54
1.6. Objectifs et plan de la thèse.....	58
Chapitre 2 : Matériels et Méthodes.....	61
2.1. Transformation bactérienne.....	61
2.2. Production des protéines.....	61
2.2.1. Expression des protéines de type sauvage non marquées.....	61
2.2.2. Expression des protéines de type sauvage avec marquage isotopique ¹⁵ N/ ¹³ C.....	61
2.2.3. Expression des protéines de type mutant fusionnées à la « Maltose Binding Protein » (MBP).....	62

2.3. Purification des protéines.....	62
2.3.1. Purification des protéines de type sauvage.....	62
2.3.2. Purification des protéines fusionnées avec une Étiquette 8-HIS.....	67
2.3.3. Purification des protéines fusionnées avec la MBP.....	68
2.4. Cristallographie des rayons X.....	69
2.4.1. Cristallogénèse.....	69
2.4.2. Collecte des données de diffraction des rayons X.....	70
2.4.3. Remplacement moléculaire et affinement.....	71
2.5. Résonance Magnétique Nucléaire.....	71
2.5.1. Attribution des résonances.....	71
2.5.2. Calculs des structures.....	72
2.5.3. Relaxation.....	72
2.5.4. ZZ- ¹⁵ N-HSQC : extraction des constantes de vitesse d'échange.....	72
2.5.5. Perturbations de déplacements chimiques (CSP).....	73
2.6. Bio-informatique et logiciels.....	74
2.6.1. Visualisation des structures et génération des figures.....	74
2.6.2. Graphiques et ajustement des courbes expérimentales.....	74
2.6.3. Criblage, préparation et analyse des structures.....	74
2.6.4. Alignement de séquences primaires.....	74
2.7. Échange de chaînes entre les dimères.....	75
2.8. Biochimie.....	75
2.8.1. Pontage chimique par le glutaraldehyde.....	75
2.8.2. Oxydation et purification de HU α_2 F79C en conditions natives.....	75
2.8.3. Electrophorèses en gel de polyacrylamide et d'agarose.....	75
2.9. Simulation de dynamique moléculaire.....	76
2.10. Spectrométrie de masse.....	77

Chapitre 3 : Stratégie de clonage pour la production et la purification des variants de HU.....79

3.1. Contraintes à résoudre.....	79
3.2. HU fusionnée à une étiquette Histidines.....	80
3.2.1. Construction des vecteurs pET28a-8HE-hupA/B.....	80
3.2.2. Tests d'expressions du vecteur pET28a-8HE-hupA/B.....	84
3.2.3. Reclonage des gènes <i>hupA/B</i> étiquetés 8His dans pJES.....	85
3.2.4. Tests d'expression du vecteur pJES-8HE-hupA/B.....	87
3.2.5. Purification de la protéine HU fusionnée avec une étiquette 8His.....	88
Lyse des bactéries (EB).....	88
1ère Chromatographie d'affinité (AF1).....	89
Clivage de l'étiquette 8His par l'entérokinase (EK).....	91
2ème chromatographie d'affinité (AF2).....	92
Chromatographie échangeuse d'ion (20CM).....	93
3.2.6. Contrôle de repliement par RMN.....	94
3.2.7. Conclusion.....	94
3.3. HU fusionnée avec la Maltose Binding Protein (MBP).....	96
3.3.1. Stratégies alternatives pour le clonage de gènes.....	96
3.3.2. Construction des vecteurs pMAL-c5E-hupA/B.....	101
3.3.3. Tests d'expressions.....	101
3.3.4. Purification de la protéine.....	102

Lyse des bactéries (EB).....	102
Chromatographie d'affinité (Amy).....	103
Clivage de l'étiquette MBP par l'entérokinase (EK).....	103
Chromatographie échangeuse d'ion SP-FF (FF).....	105
Tamissage Moléculaire (S200).....	107
Chromatographie échangeuse de cation (20CM).....	107
3.3.5. Conclusion.....	108
3.4. À retenir.....	109

Chapitre 4 : Caractérisation des structures natives.....111

4.1. Structure cristallographique de HU β_2	112
4.1.1. Cristallogénèse.....	112
4.1.2. Diffraction des rayons X et résolution de la structure cristallographique.....	113
4.2. La protéine HU de <i>Lactococcus lactis</i> (LlHU).....	116
4.2.1. Culture de la souche MG1363 de <i>Lactococcus lactis</i>	116
4.2.2. Clonage du gène <i>hup</i> de <i>L. lactis</i> dans le vecteur pMAL-c5e.....	117
4.2.3. Production et purification de la protéine HU de <i>L. lactis</i>	118
4.2.4. Cristallogénèse.....	119
4.2.5. Diffraction des rayons X et structure cristallographique préliminaire de LlHU.....	120
4.3. Analyse comparative des structures natives de HU.....	122
4.3.1. Criblage de la PDB.....	122
4.3.2. Analyse des surfaces d'interactions des dimères.....	125
4.3.3. Analyses des cavités hydrophobes des dimères.....	128
4.4. RMN : Attribution des dimères de HU de <i>E. coli</i> et caractérisation de l'homodimère HU α_2 natif (N ₂).....	133
4.4.1. Attribution des dimères de HU dans leur conformation native.....	133
4.4.2. Article I : Attribution du squelette peptidique des trois dimères de HU à 293K : EcHU α_2 , EcHU β_2 et EcHU $\alpha\beta$	135
4.4.3. Structures des homodimères en solution (en cours d'affinement).....	145
4.4.4. Relaxation des homodimères.....	148
4.4.5. Effet de la température sur les spectres ¹⁵ N-HSQC de HU α_2	150
Suivi des déplacements chimiques.....	150
Suivi des intensités.....	154
Analyse des pics apparaissant avec l'augmentation de la température.....	156
4.5. À retenir.....	158

Chapitre 5 : La conformation intermédiaire.....159

5.1. Caractérisation des conformations intermédiaires par RMN.....	161
5.1.1. Attribution de la conformation intermédiaire des homodimères.....	161
5.1.2. Calcul des constantes cinétiques d'échange.....	166
5.1.3. Perturbations de déplacement chimique entre les conformations N ₂ et I ₂	167
5.1.4. Prédiction des structures secondaires à partir des déplacements chimiques.....	169
5.1.5. Analyses des spectres NOESY des conformations natives et intermédiaires des homodimères.....	172
5.2. Étude de la transition N ₂ → I ₂ en simulation de dynamique moléculaire dite « en gros grains ».....	178
5.2.1. Analyse des trajectoires.....	180
Mouvements des hélices du HTH.....	180

Stabilité de l'hélice C-terminale.....	185
Conclusion.....	187
5.3. À retenir.....	189
Chapitre 6 : Le Tétramère Transitoire.....	191
6.1. Mise en évidence d'un tétramère par pontage chimique.....	192
6.1.1. Pontage des dimères de HU à concentration variable en glutaraldéhyde.....	193
6.1.2. Effets de la température sur le pontage.....	194
6.2. Détection du tétramère par spectrométrie de masse en conditions natives.....	196
6.3. Mise en place d'une stratégie pour isoler le tétramère.....	203
6.3.1. Mutagenèse dirigée et test <i>in silico</i>	204
6.3.2. Production et Purification du mutant F79C de HU α_2	206
6.3.3. Pontage des cystéines de la version mutante F79C de HU α_2	207
6.3.4. Suivi de la formation du pont disulfure de EcHU α_2 F79C par spectrométrie de masse	209
6.3.5. Analyse des dommages oxydatifs du mutant EcHU α_2 F79C.....	212
6.3.6. Purification et caractérisation du mutant F79C.....	214
6.4. Modèles putatifs de la structure du tétramère.....	217
6.5. À retenir.....	220
Chapitre 7 : Étude préliminaire de la Formation de l'hétérodimère.....	221
7.1.1. Mode opératoire.....	221
7.1.2. Cinétique de formation de l'hétérodimère suivie par spectrométrie de masse en conditions natives.....	221
7.1.3. Échange de chaînes entre homodimères identiques.....	224
7.1.4. Accord entre les mécanismes et constantes cinétiques.....	227
Description des mécanismes d'échange envisagés.....	227
Déterminations des étapes limitantes.....	229
Confrontation des modèles.....	229
7.2. Conclusion.....	230
7.3. À retenir.....	231
Conclusions & Perspectives.....	233
Chapitre Connexe : La protéine HU d'<i>Escherichia coli</i> est capable de moduler le système de Réparation de l'ADN par Excision de Bases.....	241
7.4. Introduction.....	241
7.5. Article : La protéine HU est capable de stimuler la réparation de la 8-oxoguanine médiée par la Formamidopyrimidine ADN Glycosylase.....	242
7.6. Conclusions et perspectives.....	279
Bibliographie.....	283
Annexes.....	307

Abréviations

3D	Tridimensionnelle
ADN	Acide désoxyribonucléique
AP	Site apurique/apyrimidique
ARN	Acide ribonucléique
BIS Tris	Bis-(2-hydroxyethyl)Imino-tris(hydroxymethyl)methane
CAPS	N-Cyclohexyl-3-aminopropanesulfonic acid
CD	Dichroïsme circulaire
CHES	2-(N-Cyclohexylamino)ethane Sulfonic acid
CV	Volume de colonne
CSP	Chemical Shift Perturbation
DTT	Dithiothréitol
HEPES	2-(4-(2-hydroxyethyl)-1-piperazinyl)ethanesulfonic acid
HRV3C	Human rhinovirus 3C
LB	Lysogeny broth
MD	Dynamique moléculaire
MES	2-(N-morpholino)ethanesulfonic acid
MPD	2,4-methyl-pentanediol
MS	Spectrométrie de masse
MT	Marqueur de taille
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
ORF	Cadre de lecture (Open Reading Frame)
PAGE	Poly-Acrylamide Gel Electrophoresis
pb	Paire de bases
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol

TEV	Tobacco Etch Virus
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
RDC	Couplages dipolaires résiduels
RMN	Résonance magnétique nucléaire
RMSD	Root Mean Square Deviation : deviation standard
RX	Diffraction des rayons X
SDS	Sodium dodecyl sulfate

Chapitre 1 : Étude bibliographique

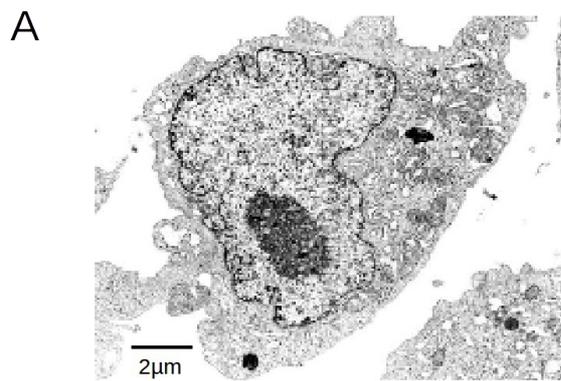
1.1. La nécessité de compaction des génomes

L'ADN constitue l'espèce moléculaire support de l'information génétique (génome) commune à tous les règnes du vivants (Eucaryotes, Bactéries et Archées) ainsi qu'aux organelles symbions (chloroplastes et mitochondries) et certains virus. La taille des génomes varie énormément en fonction des espèces. Chez les eucaryotes, le plus grand génome connu aujourd'hui est celui de *Paris japonica* avec environ 150 milliards de paires de bases, alors que le plus petit est attribué au parasite *Encephalitozoon intestinalis* avec 2,3 millions de paires de bases (pb).¹ Cependant, quelques soient les espèces, la taille des molécules non contraintes est souvent gigantesque en comparaison des volumes des compartiments dans lesquels elles sont contenues (**Figure 1, Tableau 1**).

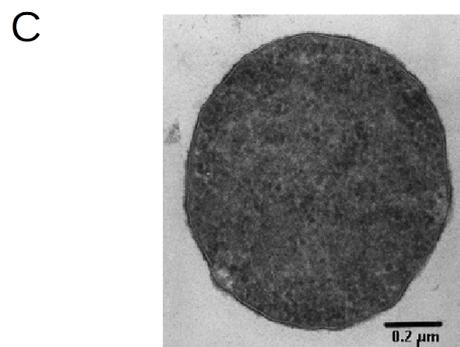
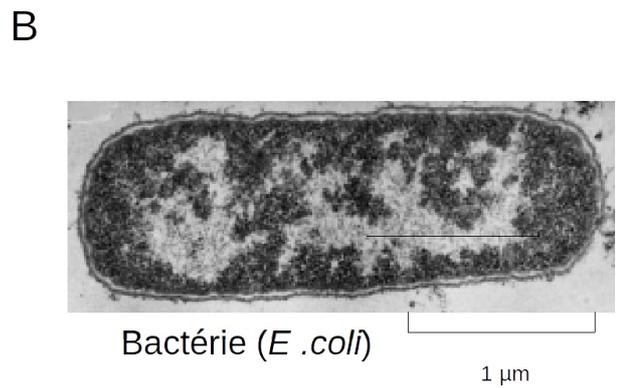
Si l'on met à part les virus, pour lesquels la contrainte est purement stérique, le défi chez la plupart des organismes vivants réside dans la conciliation de la compaction du génome dans un faible volume tout en maintenant les fonctions biologiques actives telles que la transcription des gènes, la réplication, la réparation ou la ségrégation des chromosomes entre les cellules filles.

Tableau 1: Taille de quelques génomes représentatifs

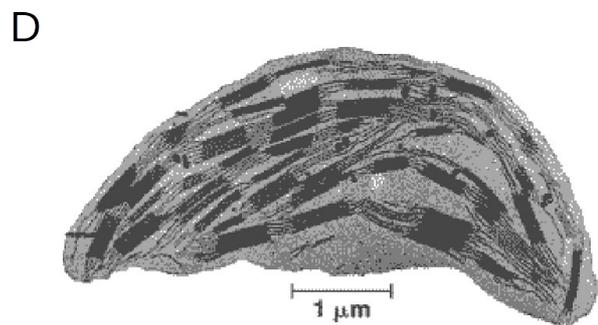
Espèce/organite	Taille du génome (pb)	Longueur du génome linéaire non contraint (ADN : 0,34 nm/pb)
<i>Paris japonica</i>	~150 000 000 000	51 m
<i>Homo sapiens</i>	3 326 743 047	1,13 m
<i>Escherichia coli</i>	4 641 652	1,57 mm
<i>Encephalitozoon intestinalis</i>	~2 300 000	782 µm
<i>Pyrococcus abyssi</i>	1 768 812	601 µm
Chloroplaste (<i>Arabidopsis thaliana</i>)	154 478	52 µm
Mitochondrie (<i>Homo sapiens</i>)	16 569 (6 à 10 copies)	5,6 µm (33,6 à 56 µm)
Bactériophage λ	48502	16,4 µm



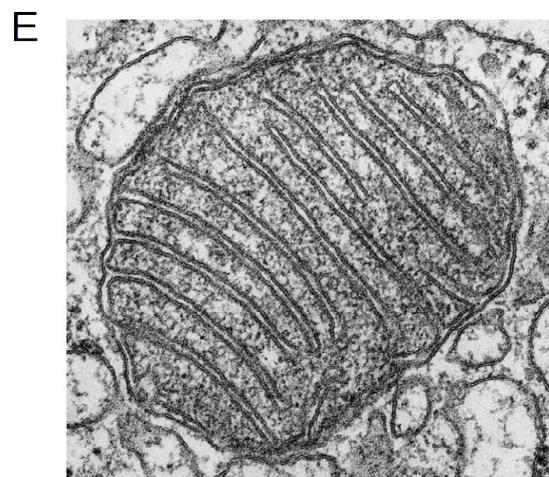
Cellule eucaryote



Archée (*T. acidophilum*)

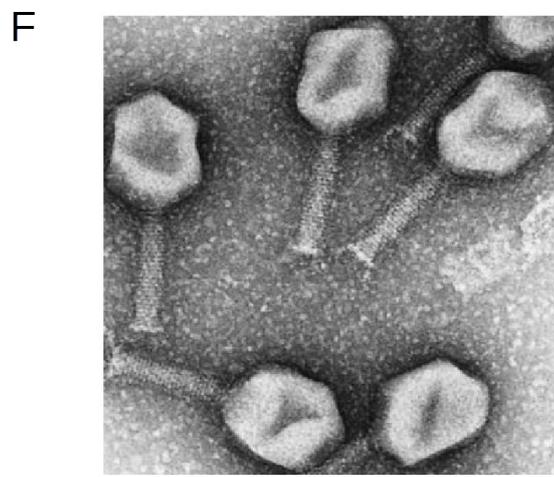


Chloroplaste



Mitochondrie

500 nm



Bactériophage λ

100 nm

Figure 1: Micrographes des différents compartiments contenant des génomes compactés

(sources : B/, E/ et F/ adaptés de « Molecular Biology of the Cell », 5^e Edition, Alberts B., Johnson A., Lewis J., Raff M., Roberts K. (Garland Science). A/ provient de Science photo library, auteur : Gopal Murti.. C/ auteur : Linda Stannard. (UCT/Photo Science Library). D/ <http://www.quia.com/jg/1782313list.html>)

1.2. Les acteurs protéiques de la compaction de l'ADN

La compaction et la décompaction fonctionnelle des génomes est le résultat de l'action concertée de 2 grands types de protéines : les protéines qui contraignent et/ou maintiennent des structures locales de l'ADN telles que les histones et les protéines de type histone et les enzymes de topologie, les topoisomérases.

1.2.1. Les Histones et les protéines de type Histone

Ce sont de petites protéines basiques dépourvues d'activité enzymatique mais capable d'interagir fortement avec l'ADN. Ces protéines peuvent être subdivisées en 3 grands types en fonction du mode de fixation. On distingue les **ponteurs** (« Bridgers ») qui fixent 2 sites distants d'ADN (inter- ou intra-moléculaire), des **enrouleurs** (« wrappers ») qui provoquent un sur-enroulement de l'ADN en le fixant sur leur pourtour, et des **courbeurs** (« benders ») qui imposent une courbure à l'ADN par fixation (côté concave ou convexe de la courbure) (**Figure 2**). L'association de ces protéines avec l'ADN constitue la **chromatine**.

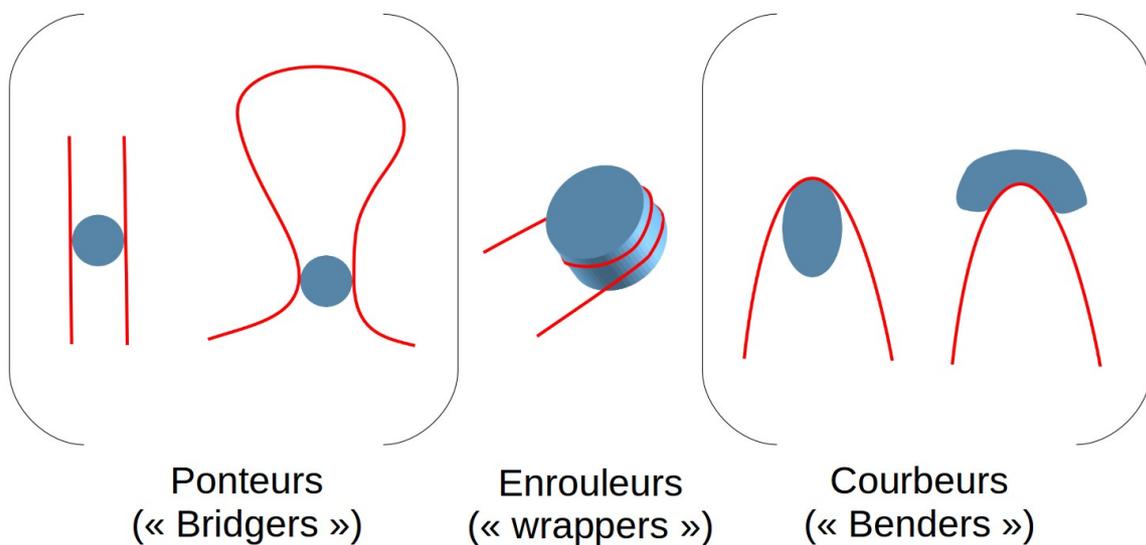


Figure 2: Modes de compaction des protéines de type histone En bleu, les protéines de type histone et en rouge des fragments d'ADN.

1.2.2. Les Topoisomérases

La compaction et certains processus métaboliques tels que la réplication, la transcription ou la recombinaison génèrent des contraintes topologiques sur l'ADN.^{2,3} Le sur-enroulement, positif ou négatif, est régulé par une classe d'enzymes qu'on appelle les **topoisomérases** qui sont également capables de résoudre des problèmes d'entrelacements intramoléculaires (nœuds) ou de réaliser la décaténation entre deux molécules d'ADN lors de la réplication des chromosomes bactériens par exemple (**Figure 3**).

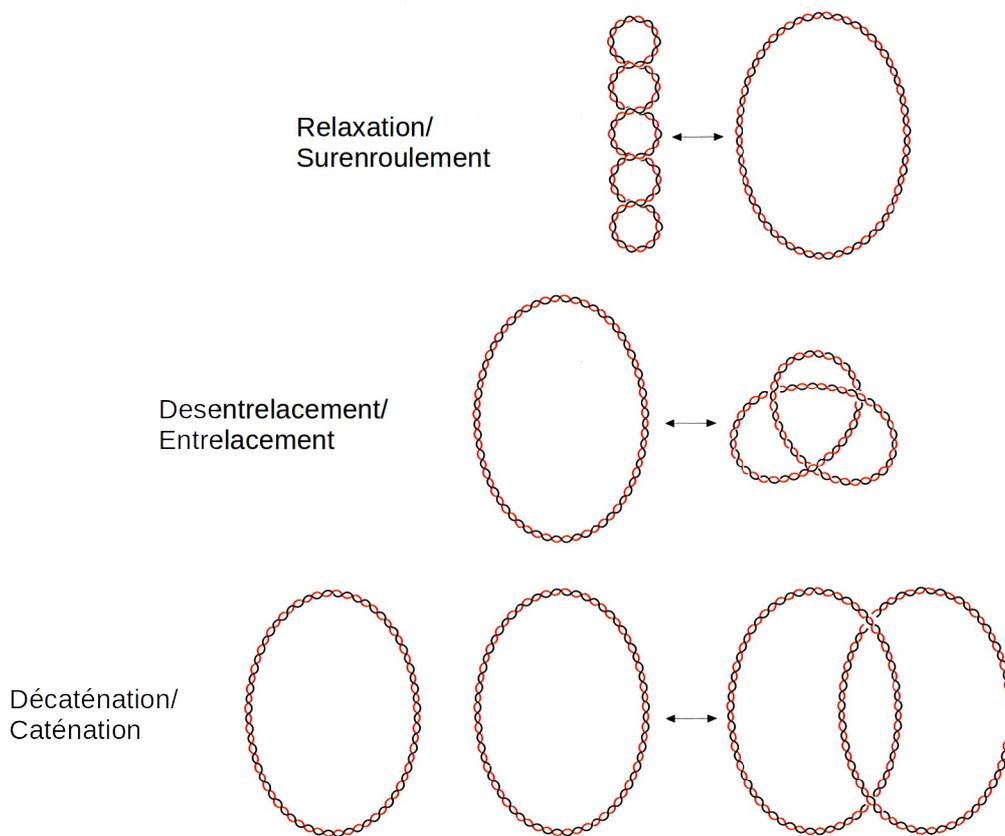


Figure 3: Fonctions des Topoisomérases (Adapté de A.D. Bates et A. Maxwell)^{164,265}

Les topoisomérases clivent l'ADN en formant un complexe covalent transitoire relié par une liaison phosphodiester entre un acide aminé tyrosine de la protéine et une extrémité de l'ADN clivé. La topology de l'ADN est modifiée alors que la protéine est toujours liée covalamment à l'ADN. Elle n'est libérée que lorsque l'ADN est religué.

On distingue les topoisomérases de type I qui ne clivent qu'un brin d'ADN et les topoisomérases de type II qui forment une cassure double brin transitoire (**Figure 4**). Les topoisomérases de type I sont subdivisées en deux types : le type IA pour les topoisomérases qui forment une liaison covalente avec l'extrémité 5' de l'ADN clivé et le type IB pour celles qui se lient covalamment avec l'extrémité 3' de l'ADN clivé.

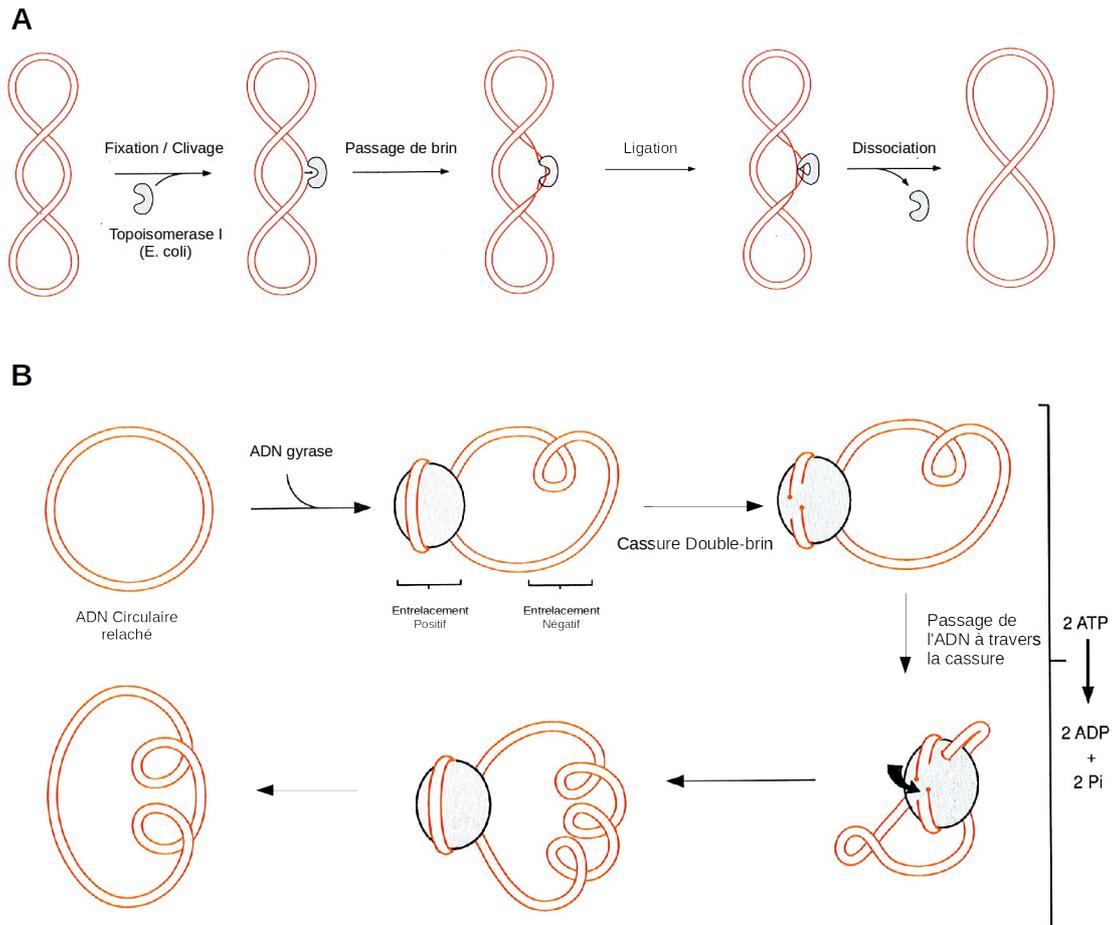


Figure 4: Schémas des mécanismes d'action de la topoisomerase I (Type I) et de l'ADN gyrase (Type II) d'*E. coli* (Adapté de A.D. Bates et A. Maxwell)¹³⁹

Les topoisomérases de type IA sont principalement monomériques. Elles peuvent relaxer l'ADN surenroulé négativement mais pas positivement. Leur activité requiert la présence de magnésium et d'une région simple brin exposée. La structure de la topoisomérase I d'*E. Coli*, du type IA, est présentée **Figure 5**.

Les topoisomérases de type IB présentent une homologie structurale avec les tyrosine-recombinases telles que Cre (bacteriophage P1), XerD (*E. coli*) ou certaines intégrases de phages.⁴ Elles ne possèdent pas d'homologie structurale ou séquentielle avec les autres types de topoisomérases. Elles sont capables de relaxer l'ADN sur-enroulé positivement ou négativement, et ne nécessitent ni magnésium ni région simple brin. La structure de la topoisomérase I humaine, du type IB, est présentée **Figure 5**.

Les topoisomérases de type II sont également subdivisées en types IIA et IIB sur la base de considérations structurales. Celles de type IIA présentent un repliement conservé à l'exception de l'extrémité C-terminale qui peut varier d'un enzyme à l'autre. Un des membres les mieux caractérisés de cette famille est la gyrase capable d'introduire des super-tours négatifs en utilisant l'ATP comme source d'énergie (**Figure 4 B**). La structure de la gyrase A de *M. tuberculosis* est présentée **Figure 5**. Le type IIB a été introduit par la mise en évidence de la topoisomérase VI d'archée.⁵ Cette protéine possède très peu d'homologie de séquence ou de structure avec d'autres topoisomérases (à l'exception du domaine ATPasique) mais possède une homologie avec la protéine Spo11 impliquée dans l'initiation de la recombinaison homologue durant la méiose. Cependant, au contraire des topoisomérases, Spo11 ne religue pas les brins d'ADN après avoir généré la cassure double brin. La structure de la topoisomérase VI de *M. mazei* est présentée **Figure 5**.

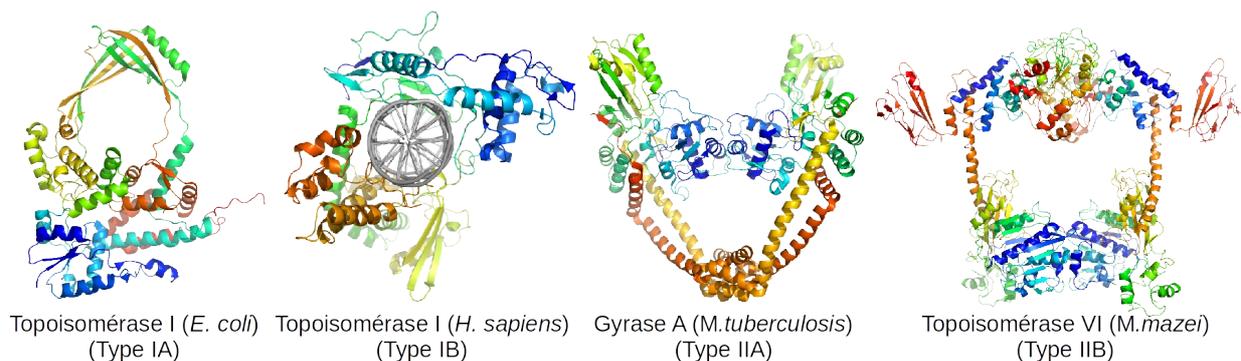


Figure 5: Structures de différents types de topoisomérases Le dégradé de couleur reflète la position dans la séquence primaire (Du bleu au rouge, de l'extrémité N-terminal à C-terminal respectivement). (Codes PDB de la gauche vers la droite : 1CY1, 1A35, 3ILW et 2QE)

1.3. La Compaction de l'ADN dans les 3 grands règnes du Vivant

1.3.1. La compaction de l'ADN chez les Eucaryotes

Généralités

Il existe différents niveaux de compaction de l'ADN nucléaire (**Figure 6**). Le premier niveau de compaction correspond à la formation d'un complexe nucléoprotéique, le **nucléosome**. Il est observable par microscopie électronique à transmission et présente un aspect caractéristique dit en « collier de perles » où chaque « perle » correspond à un nucléosome (flèche noire) relié par des segments d'ADN nus (flèche blanche) (**Figure 7**). Le nucléosome est constitué de **deux copies de chaque histone (H2A, H2B, H3, H4) formant un octamère** enroulant 147 pb d'ADN (soit 1,65 tours par nucléosome). La distance inter-nucléosomique est de l'ordre de 10 à 80 pb. Cette structuration de la chromatine la moins compacte (~11 nm de diamètre) est aussi appelée **euchromatine**. Les régions de l'ADN activement transcrites sont toujours dans l'euchromatine.

Une fibre plus compacte (~30 nm de diamètre) est obtenue par l'association des nucléosomes par l'intermédiaire d'autres protéines appelées les histones de liaison (« linker histones » tels que l'histone H1). Bien que souvent décrite *in vitro*, l'existence de cette fibre *in vivo* et sa structuration sont sujettes à controverse.⁶ A ce jour, différents modèles sont proposés pour la structuration de cette fibre de 30 nm : une structuration de type solénoïde à dominantes d'interaction $i/i+1$ (interactions internucléosomiques),⁷⁻⁹ une structuration zig-zag à dominantes d'interaction $i/i+2$,^{10,11} ou encore une combinaison de ces deux modèles. Cette fibre condensée est aussi appelée **hétérochromatine** et correspond à une région faiblement transcrite. L'hétérochromatine peut être soit constitutive pour des régions particulières (centromères, télomères) soit facultative pour des régions où l'expression des gènes a été réprimée.

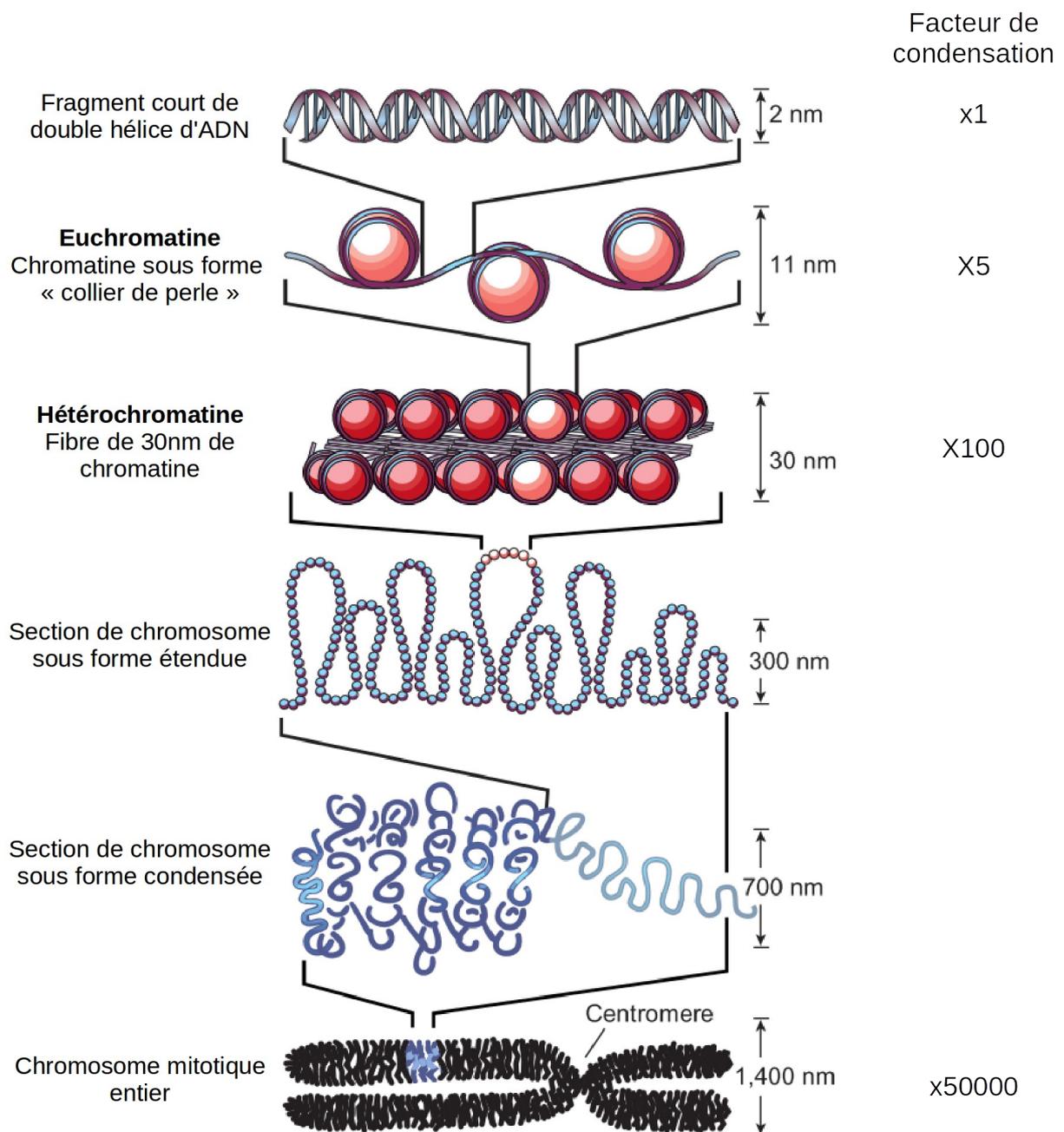


Figure 6: Niveaux de compaction de l'ADN nucléaire Schéma adapté de Felsenfeld G. et Groudine M.¹⁶⁶

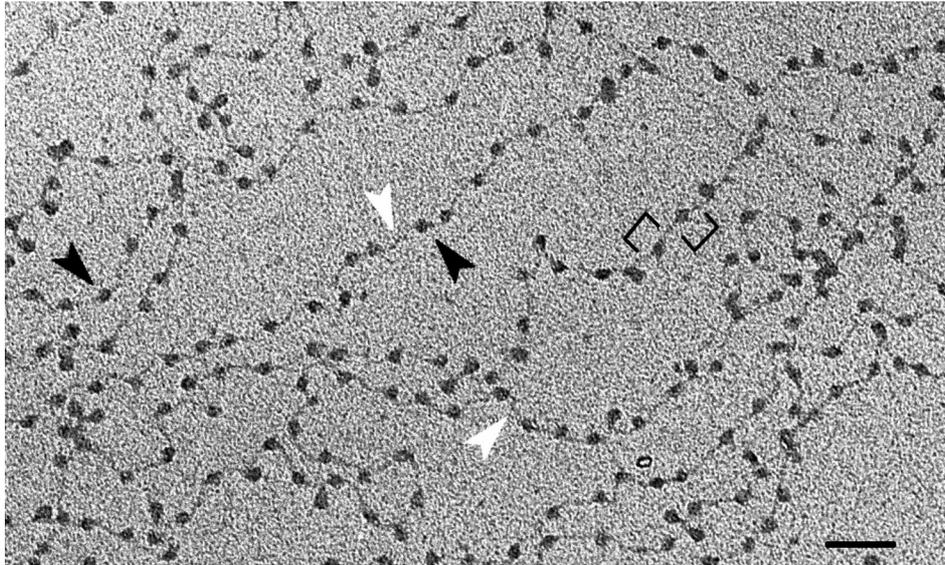


Figure 7: Micrographe électronique de la fibre nucléosomique Image de microscopie électronique à transmission de chromatine décondensée d'érythrocyte de poulet. Les flèches noires indiquent des nucléosomes et les flèches blanches des régions inter-nucléosomiques. La barre représente 50nm. (source : région sélectionnée de l'image obtenue à l'adresse suivante: <http://www.cellimagelibrary.org/images/709>, dans le domaine public , auteur : Chris Woodcock)

L'état euchromatine ou hétérochromatine peut être contrôlé par des modifications post-traductionnelles des histones (acétylation/méthylation des lysines, méthylation des arginines, ubiquitinylation, phosphorylation des sérines). Ces marques épigénétiques sur les histones influent sur la condensation de l'ADN et sur les fonctions biologiques associés à la chromatine, notamment la transcription mais aussi la réparation de l'ADN.¹² C'est le code histone proposé par Junwein T. et collaborateurs.¹³

Outre les modifications post-traductionnelles, le remplacement de certaines histones par des variants peut influencer sur l'état de condensation de l'ADN et sur la transcription des gènes. Les variants H3.3 et H2A.Z par exemple sont des marqueurs de chromatine active d'un point de vue transcriptionnel alors que le variant CENP-A de H3 est associé à la région centromérique.

Cette fibre de 30 nm peut encore être condensée d'avantage jusqu'à l'état le plus compact correspondant au chromosome mitotique (**Figure 6**).¹⁴

Les Histones et le Nucléosome

La structure du nucléosome a été déterminée par diffraction des rayons X pour différentes espèces : le xénope, la levure, la drosophile et l'Homme.¹⁵⁻¹⁷ La structure atteignant la plus haute résolution (1,9 Å) est celle du nucléosome de xénope résolue par Davey et collaborateur¹⁶ (**Figure 8 A**). L'octamère d'histones est constitué d'un corps hétérotétramérique (H3-H4)₂ flanqué de deux hétérodimères H2A-H2B. Cet octamère peut être subdivisé en dimères qui possèdent des repliements semblables (**Figure 8 B**).

Les éléments de structures secondaires dans les histones sont presque exclusivement des hélices α et des portions non structurées. Les protomères de chaque dimère sont enchevêtrés et interagissent l'un avec l'autre *via* leurs hélices. L'association des chaînes protéiques deux à deux semble donc nécessaire pour une structuration correcte d'un protomère. Effectivement, Sperling R. et Bustin M.¹⁸ ont montré que les dimères constituent des intermédiaires stables en solution (en absence d'ADN) avant la formation de l'octamère nucléosomique (en présence d'ADN). De plus, le mélange de dimères d'histones *in vitro* (sauf H1) conduisait à un échange d'histones entre ces dimères. Ces dimères peuvent s'associer entre eux *via* un motif hélice-tour-hélice pour former des oligomères d'ordre supérieur allant jusqu'à l'octamère.

Ces histones présentent également de grandes régions N-terminales très basiques et non structurées qui émergent hors de la structure nucléosomique (**Figure 8**). C'est dans ces régions que se trouve la majorité des sites de modifications post-traductionnelles (acétylation/méthylation/ubiquitinylation) à l'origine du code histone.

Les protéines non histones

Les protéines HMG (« High mobility group ») représentent la majorité des protéines non histones impliquées dans l'architecture de l'ADN. La désignation « high mobility group » est due à leur propriétés électrophorétiques. Ces protéines étant de petites tailles et très chargées, elles migrent rapidement sur gel d'électrophorèse. Les protéines à boîtes HMG peuvent être divisées en trois sous-familles : HMGA, HMGB et HMGN.¹⁹

La famille HMGA, ou HMG-AT-hook, est constituée de quatre protéines HMGA1a et HMGA1b et HMGA1c, issues de l'épissage alternatif du gène HMGA1, ainsi que HMGA2 qui est issue d'un gène différent. Ces protéines présentent toutes un motif « AT-hook » qui reconnaît les régions de

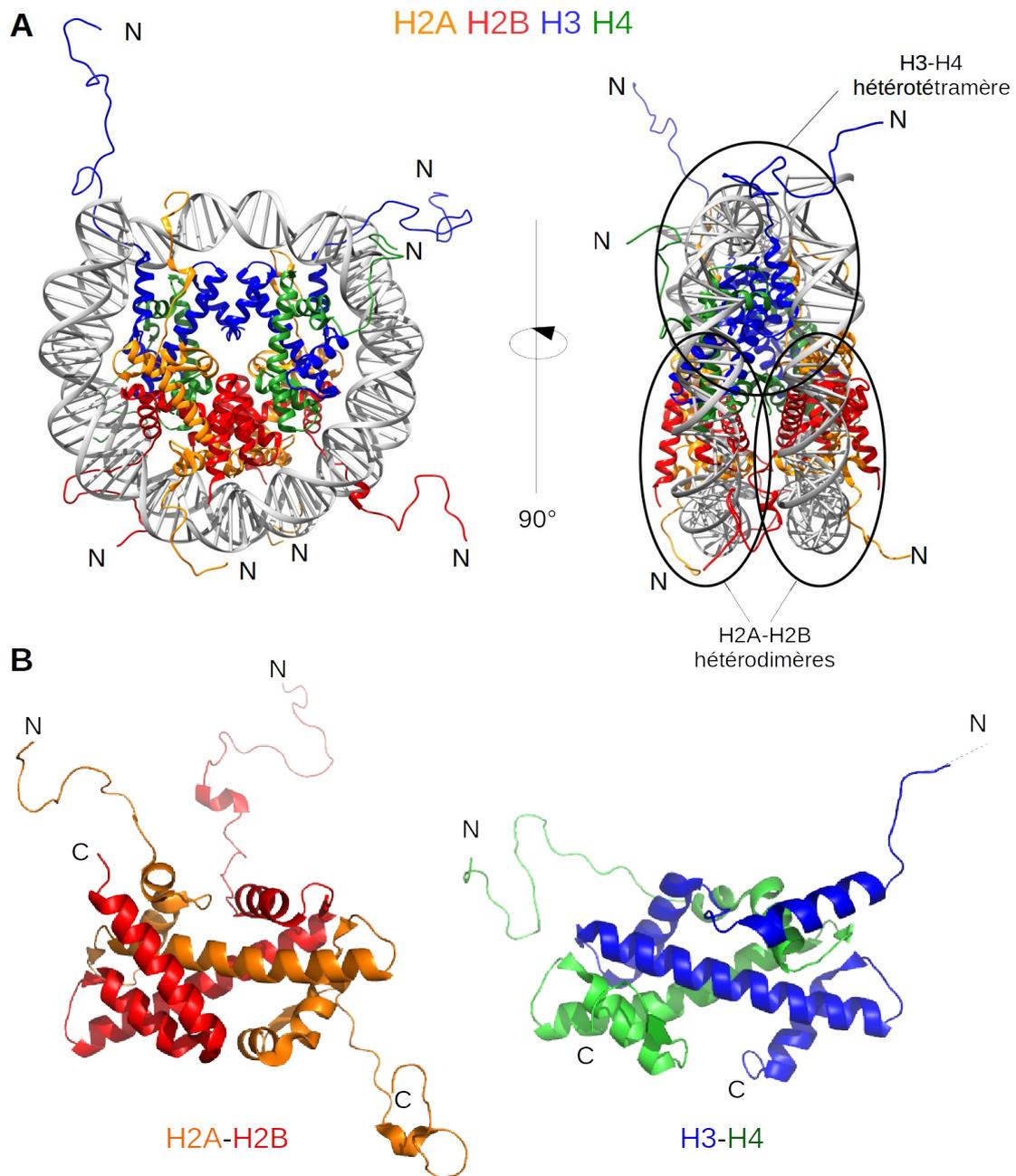


Figure 8: Le nucléosome A- Structure cristallographique du nucléosome de *Xenopus laevis* (code pdb : 1KX5). L'ADN est coloré en gris, les histones H2A, H2B, H3, H4 sont colorées en orange, rouge, bleu et vert, respectivement. Les régions N-terminales non structurées sont indiquées (« N »). B- Structure des hétérodimères H2A- H2B et H3-H4 du nucléosome.

l'ADN riches en paires A/T.²⁰ Ces régions forment un petit sillon étroit qui permet leur reconnaissance structurale par ces protéines. Elles sont également capables de reconnaître spécifiquement des jonctions quatre voies²¹ et des ADN courbés ou surenroulés.²² En plus de leur reconnaissance de l'ADN, les protéines HMGA sont capables d'interagir avec de nombreux partenaires protéiques ayant des fonctions liées au métabolisme des acides nucléiques telles que la compaction, la transcription et la réparation de l'ADN.²³ Parmi ses partenaires d'interaction, les protéines HMGA sont notamment capables d'interagir avec les nucléosomes.²⁴ Aucune structure n'a été rapportée pour cette famille. En effet, les protéines HMGA sont intrinsèquement désordonnées lorsqu'elles sont libres en solution. Ce caractère est supposé à l'origine de la capacité de ces protéines à interagir avec ses nombreux partenaires protéiques.

Les protéines HMGB présentent un motif structural conservé constitué de un à six domaines « boîtes HMG » en tandem (~80 acides aminés par boîte) pouvant être terminé par une queue C-terminale acide (~30 acides aminés). Chaque boîte HMG est composée d'un enchaînement de trois hélices α formant une structure en L avec un angle d'environ 80° (**Figure 9 A**). La queue C-terminale acide est désordonnée.²⁵ L'interaction avec l'ADN se fait du côté concave du L (côté convexe de la courbure) au niveau du petit sillon et provoque une forte courbure d'environ 90° (**Figure 9 B**). Les protéines HMGB sont également capables de reconnaître, avec une haute affinité, certaines structures de l'ADN présentant des courbures telles que des jonctions quatre voies, des hernies, des coudes ou des ADN modifiés contenant par exemple des adduits du cis-platine.²⁶ Les protéines HMGB sont également capables d'interagir avec le nucléosome.²⁷ La fixation de HMGB1 sur le nucléosome permet le recrutement de facteurs de remodelage de la chromatine (ACF/CHRAC) antagonisant l'effet de l'histone H1 en augmentant la dynamique du nucléosome.²⁸

Les protéines HMGN sont composées des protéines HMGN1 à 5 que l'on retrouve uniquement chez les vertébrés.²⁹ Les fonctions de ces protéines semblent intimement liées à la différenciation cellulaire.³⁰ L'expression des gènes codant ces protéines diminue au cours du développement embryonnaire.³¹ Cependant cette diminution ne semble pas survenir dans les cellules constamment renouvelées.³² Les protéines HMGN sont capables de faciliter la décondensation de l'hétérochromatine afin de rendre l'ADN accessible aux processus de transcription, de réplication et de réparation.²⁹ Un des mécanismes proposés pour expliquer cette décondensation est la compétition de ces protéines avec l'histone de liaison H1.³³ Il n'y a pas de structure disponible pour les protéines de cette famille. Les protéines HMGN sont intrinsèquement désordonnées. Elles sont constituées

d'une région de localisation nucléaire (NLS), suivie d'une région de fixation au nucléosome d'environ 30 acides aminés et terminée par une queue C-terminale acide.²⁹

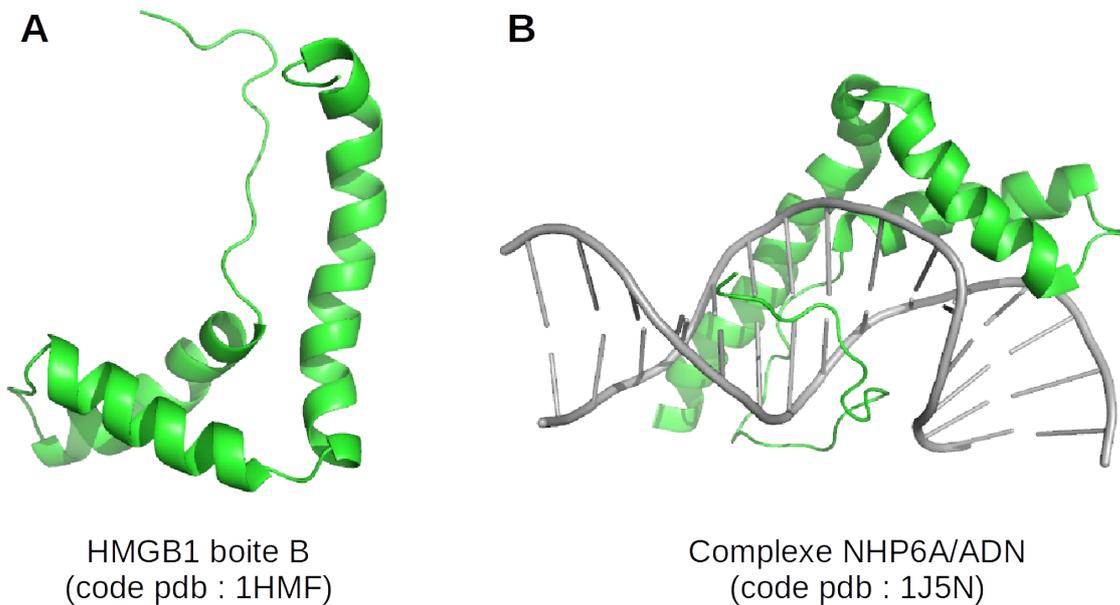


Figure 9: Structures de boîtes HMG libre (A) ou en complexe avec l'ADN (B) de protéines de la familles HMGB

Les trois familles de protéines HMG décrites ici jouent un rôle dans le remodelage de la chromatine, soit par interaction directe avec l'ADN, soit par interaction avec les nucléosomes. Ces dernières entrent en compétition avec la fixation d'histones de liaison telle que H1 ce qui conduit à la décondensation de la chromatine. chez les eucaryotes, l'équilibre entre l'action des protéines non-histones et les histones de liaisons permet une régulation fine de l'expression des gènes par l'intermédiaire de la condensation de l'ADN.

1.3.2. La compaction de l'ADN chez les Archées

Généralités

Des trois grands règnes du Vivant, les archées représentent les organismes les moins bien caractérisés. D'abord considérés comme des bactéries vivant dans des milieux extrêmes (alors nommées archéobactéries), les archées présentent des morphologies semblables à celles des bactéries. Leur taille (de l'ordre du μm) et celle de leur génome (quelques Mbp) sont similaires et, étant tous deux des procaryotes, elles ne présentent pas de compartimentation de leur génome dans un noyau. L'ADN est compacté dans une structure chromatinienne appelée le nucléoïde. Néanmoins les machineries de réplication et de transcription des archées sont bien plus proches de celles des eucaryotes.

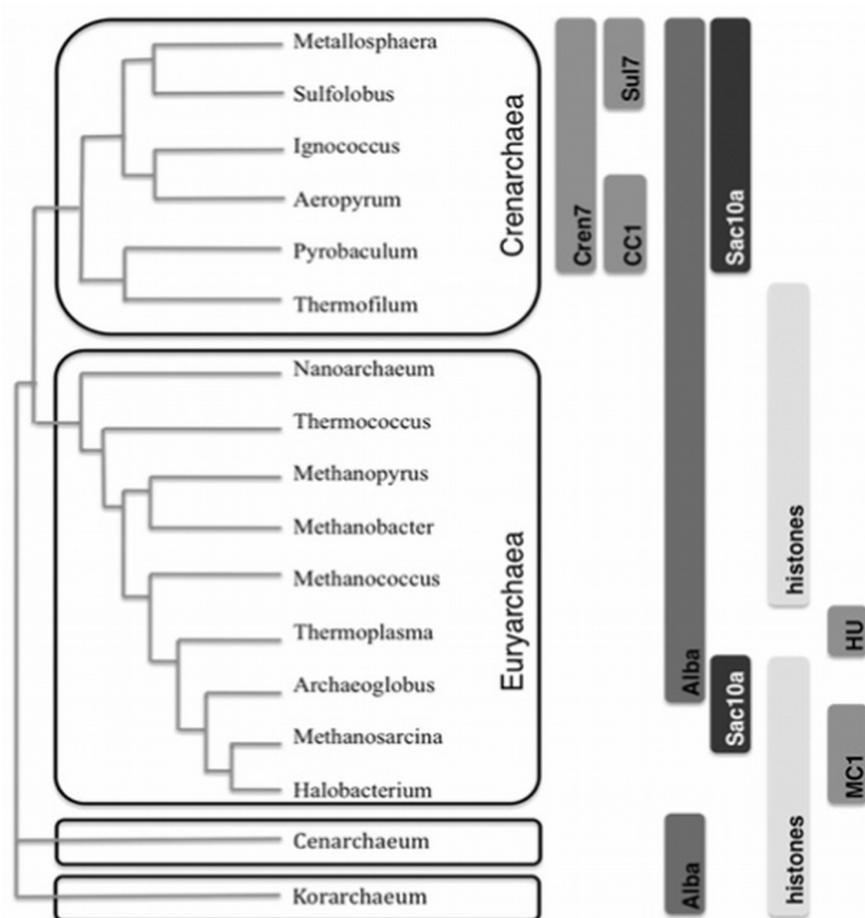


Figure 10: Distribution phylogénique des protéines majeures de compaction de l'ADN chez les Archées (de Driessen et Dame 2011)¹⁸⁵

La compaction de l'ADN chez les archées est réalisée par des protéines de type histone qui diffèrent en fonction des *phyla* (**Figure 10**). On distingue les protéines histones, principalement présentes chez les euryarchées, qui possèdent une homologie de séquence et de structure avec les histones eucaryotes, des protéines de type histones qui ne possèdent qu'une homologie fonctionnelle avec les histones à l'instar des bactéries.

Les Histones d'Archées

La plupart des histones d'archées découvertes à ce jour possèdent uniquement le motif structuré dépourvu des extensions N- ou C-terminales déstructurées typiquement eucaryotes (**Figure 11**). La plupart des modifications post-traductionnelles, chez les eucaryotes, ayant lieu sur ces portions de séquence, il est probable que la régulation fine de la compaction par le code histone ne soit pas employée par les archées sur ces protéines. Il existe cependant des variants chez certaines espèces qui présentent des extensions C-terminales, mais celles-ci n'ont pas d'homologie de séquences avec les queues d'histones eucaryotes et ont davantage été associées à un gain de thermostabilité.³⁴

Les histones d'archées restent sous forme dimérique en solution et s'assemblent sous forme tétramérique en présence d'ADN.^{35,36} Les expériences de circularisation sur l'histone HMfb de *methanothermus fervidus* menées par Bailey et collaborateurs³⁶ montrent un enroulement de l'ADN sur ~80 pb. Cet enroulement est similaire à celui de l'hétérotétramère H3-H4 au sein du nucléosome eucaryote et suggère que les archées possédant des histones peuvent former des structures similaires aux nucléosomes eucaryotes autour d'un tétramère d'histone (**Figure 12 B**). En effet, des structures en « collier de perle » ont été observées par microscopie électronique sur des extraits d'ADN génomique de *Methanobacterium thermautotrophicus* (**Figure 12 A**).

Certaines histones d'archées, comme HMk chez *Methanopyrus kandleri* (**Figure 11**) ou HHb chez *Halobacterium salinarum* présentent un double motif histone formant un pseudo-dimère similaire aux dimères d'histones eucaryotes. Les histones de *Methanothermus fervidus*, HMfA et HMfB, comptent parmi les histones d'archées les mieux étudiées. Bien que possédant plus de 80 % d'identité entre elles, la compaction de l'ADN induite par HMfB est plus importante que celle induite par HMfA. Par ailleurs, le dimère HMfA est plus abondant en phase exponentielle de croissance que HMfB, ce dernier devenant plus important en phase stationnaire de croissance. Tout ceci est cohérent avec une expression des gènes réduite en phase stationnaire.³⁷ Il est à noter que HMfA et HMfB peuvent former des homodimères et des hétérodimères après chauffage à 95°C.³⁷

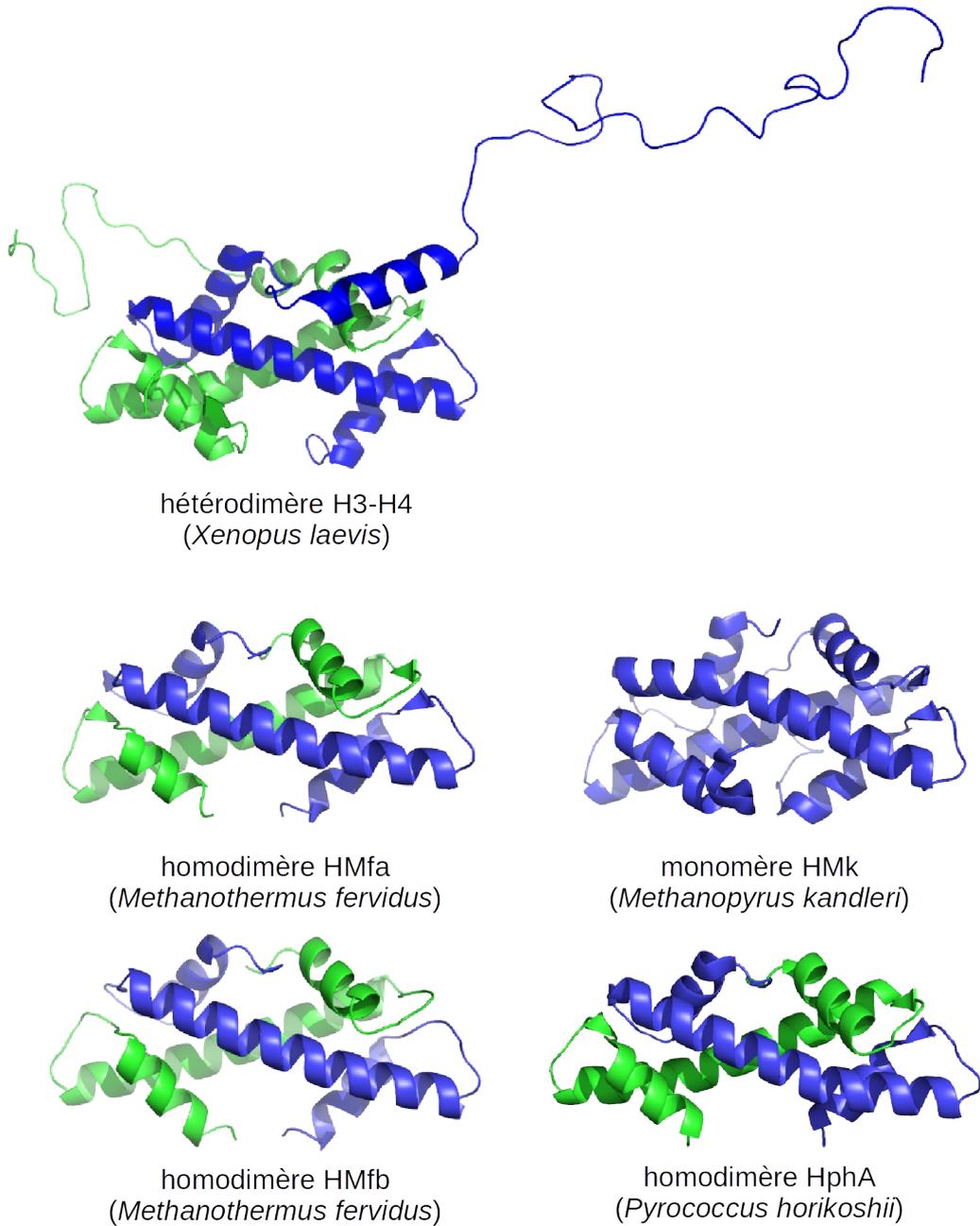


Figure 11: Comparaison des structures d'histones d'archées avec la structure de l'hétérodimère H3-H4 eucaryote Structures des histones H3-H4, HMfA, HMfB, HMk et HphA (code pdb : 1KX5, 1B67, 1A7W, 1F1E et 1KU5, respectivement).

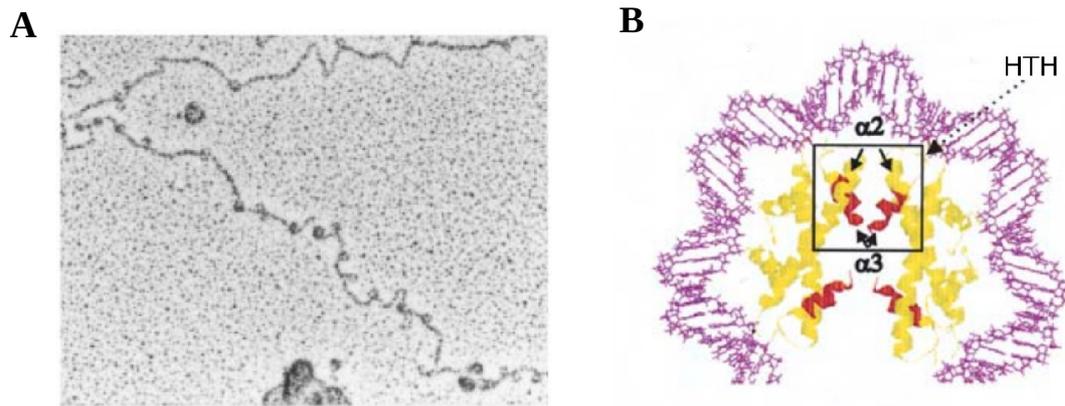


Figure 12: Formation de structures analogues au nucléosome chez les Archées A- Image de microscopie électronique d'ADN génomique de *Methanobacterium thermautotrophicus* (euryarchée) mettant en évidence une structuration en « collier de perle » (Pereira *et al* 1997).⁷³ B- Modèle de nucléosome d'archée (adapté de Reeve *et al* 2004).²⁶⁴

Les Protéines non histones chez les Archées

Il existe d'autres protéines que les histones participant à la compaction de l'ADN chez les archées, en particulier chez les Crenarchées qui, pour la plupart, ne possèdent pas d'histones. Nous allons décrire rapidement les mieux caractérisées d'entre elles.

Alba

Alba est la protéine associée au nucléoïde la plus répandue à travers les *phyla* d'archées (**Figure 10**). Cette protéine de 10 kDa est aussi décrite sous le nom de Sac10b, sso10b et Ssh10b du nom des souches d'archées d'où elle a été extraite (*Sulpholobus acidocaldarius*, *Sulpholobus solphataricus* et *Sulpholobus shibatae*). Cette protéine extrêmement abondante ($\sim 500 \mu\text{M}$)³⁸ doit son nom à sa particularité d'avoir son affinité pour l'ADN modulée par acétylation (« Acetylation Lowers Binding Affinity »).³⁹ En effet, Bell S. et collaborateurs³⁹ ont montré que ssSir2, une désacétylase, interagit avec Alba *in vitro*. Alba est naturellement acétylée *in vivo* (en N-terminal et au niveau de sa lysine 16) et la présence ou l'absence de ces groupements acétyls module l'affinité de Alba pour l'ADN d'un facteur 30. Ce travail suggère que des protéines non histones peuvent être régulées par acétylation *in vivo* à l'instar des histones eucaryotes.

Alba est une protéine dimérique en solution, mais des expériences de pontage covalent ont montré que des oligomères d'ordre supérieur pouvaient exister. Alba présente un réplètement α/β homologue à celui du facteur de transcription bactérien IF3. L'interface inter-dimère est principalement

hydrophobe et un bras en β -hairpin (semblable à la protéine HU) est supposé être impliqué dans l'interaction avec l'ADN.

Alba est capable de provoquer un sur-enroulement négatif de l'ADN à température élevée.³⁸ Cette protéine peut fixer indifféremment l'ADN double brin et simple brin ainsi que l'ARN simple brin avec des constantes de dissociation de l'ordre de 0,1 à 0,3 μM .⁴⁰ Malgré sa désignation comme une protéine associée au nucléoïde, le rôle véritable d'Alba dans la compaction de l'ADN est controversé. En effet, le travail de Bohrmann et collaborateurs a montré que la plupart des copies de la protéine Alba sont localisées dans le cytosol plutôt qu'au nucléoïde.⁴¹ De plus, à faible concentration, Alba est capable de ponter deux brins d'ADN (un par monomère) et ainsi de participer à la compaction de l'ADN en tant que ponteur (**Figure 13 B**), en formant une boucle dans l'ADN. Cependant, à haute concentration, Alba recouvre complètement l'ADN formant une fibre linéaire et rigide (**Figure 13 C**).

Il est intéressant de noter que chez *Sulfolobus*, deux gènes codent pour la protéine Alba. Le gène *alba1* code pour la protéine la plus conservée et *alba2* la protéine divergente.⁴² L'abondance de la protéine Alba2 est très faible par rapport à la protéine Alba1 (1:20 en phase stationnaire) et existe probablement principalement sous forme hétérodimérique.

Différentes études de dénaturation thermique ont été menées sur ces protéines par microcalorimétrie différentielle à balayage et dichroïsme circulaire.^{43,44} Biyani et collaborateurs⁴³ ont montré que l'homodimère Alba2 de *Sulpholobus solphataricus* pouvait exister en équilibre entre 2 conformations dimériques alternatives à température proche de la température de vie de cet organisme. D'après des expériences de dichroïsme circulaire, cette conformation alternative présente une structuration en hélice α diminuée alors que le contenu en feuillet β semble augmenter.⁴³ Par ailleurs, le travail de Wu et collaborateurs⁴⁴ sur Alba1 de *Sulfolobus shibatae* a mis en évidence au moins une transition pour cette protéine. Les auteurs ont conclu à un modèle de dénaturation à deux états $N_2 \rightarrow 2D$ mais l'indépendance du T_m à la concentration et la faible gamme de température explorée ($<95^\circ\text{C}$) pour une protéine thermostable conduit à penser que seule la transition $N_2 \rightarrow I_2$ a été observée et que la dissociation vers deux monomères dénaturés n'est observable qu'à plus haute température.

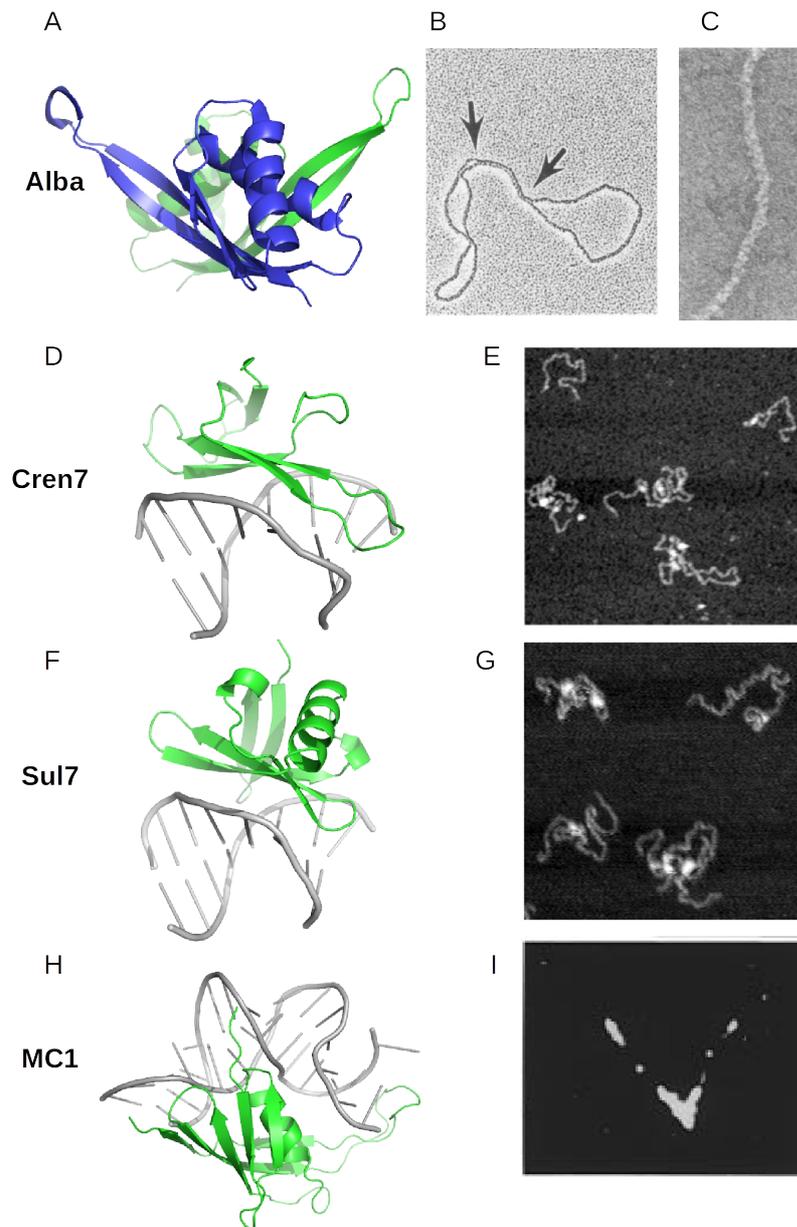


Figure 13: Protéines non histones d'archées A/D/F – Structures cristallographiques des protéines Alba1 (code pdb : 1H0X), Cren7 (code pdb : 3KXT), et sul7 (code pdb : 1AZP), respectivement. B/C- Image de microscopie électronique de Alba1 fixée sur de l'ADN double brin formant des pontages inter-brins (B) ou une fibre nucléo-protéique étendue et non compacte (C). E- Image de microscopie à force atomique de Cren7 en interaction avec de l'ADN double brin.⁴⁶ G- Image de microscopie à force atomique de Sul7 en interaction avec de l'ADN double brin.²⁶³ H- Structure RMN de MC1 complexée avec un ADN double-brin. (publication en préparation, Paquet *et al*). I- Image de microscopie électronique de MC1 avec un ADN de 176 pb possédant un site préférentiel pour MC1 en son centre.¹⁴

Cren7

Cren7 est une protéine architecturale de l'ADN présente chez la plupart des crenarchées. Cette protéine très abondante (1 % des protéines cellulaires) de 7 kDa ne possède pas de spécificité de séquence et est capable de fixer l'ADN double brin avec une affinité de l'ordre de 0,1 μM et l'ADN simple brin avec une affinité bien moindre, de l'ordre de 2 μM .⁴⁵ Sa structure 3D présente un repliement de type SH3 (SRC homologue 3) avec un bras en β -hairpin qui s'intercale dans le petit sillon de l'ADN (**Figure 13 D**) de manière analogue à la protéine HU. Le mode de fixation sur l'ADN est similaire à la protéine Sul7 bien que ces 2 protéines ne partagent pas d'homologie de séquence. La fixation de Cren7 provoque une courbure d'environ 53° sur l'ADN, notamment *via* l'intercalation de résidus hydrophobes du bras dans le petit sillon. La compaction induite par la fixation de cren7 sur l'ADN est également observable en microscopie à force atomique (**Figure 13 E**).⁴⁶ Cren7 peut être méthylée par la protéine KMT4 ce qui peut suggérer un contrôle post-traductionnel de la compaction de l'ADN par Cren7.⁴⁷

Sul7

Sul7 est une protéine associée au nucléoïde que l'on retrouve uniquement chez les archées du genre *Sulpholobus* et *Metallosphaera* (**Figure 10**). Cette protéine de 7 kDa est très abondante (5 % des protéines cellulaires). A l'instar de Cren7, elle est capable de courber l'ADN d'un angle d'environ 60° par fixation du côté convexe de la courbure (**Figure 13 F**). Sa structure est assez similaire à Cren7 avec une hélice α C-terminale supplémentaire. Les images de microscopie à force atomique révèlent un comportement similaire au niveau de la compaction d'un ADN double brin (**Figure 13 G**). Sul7 peut être méthylée au niveau de ses résidus lysines. Néanmoins, ces modifications ne semblent pas modifier son affinité pour l'ADN. Il semble que la méthylation de Sul7 soit augmentée lors de chocs thermiques, ce qui peut suggérer un rôle de cette modification dans la thermostabilité de cette protéine.⁴⁸

CC1

CC1 (Crenarchaeal Chromatin protein 1) est une petite protéine monomérique et basique de 6 kD présente chez les archées du genre *Aeropyrum* et *Pyrobaculum* (**Figure 10**). Elle a été mise en évidence lors d'un criblage à la recherche de protéine se fixant à l'ADN chez les archées.⁴⁹ Elle est capable de se lier indifféremment sur l'ADN simple brin ou double brin avec des affinités similaires. Des caractérisations structurales préliminaires par CD et RMN montrent un haut pourcentage de structuration secondaire en feuillet β ce qui est similaire d'une part à d'autres protéines

architecturales de l'ADN (telles que Cren7, Sul7 ou MC1 : figure D, F et H, respectivement) et d'autres part à l'« OB fold » caractéristique des protéines de fixation de l'ADN simple brin (SSB).⁵⁰ A ce jour, le rôle de CC1 en tant que protéine architecturale de l'ADN ou SSB reste à être documenté par d'autres investigations.

MC1

MC1 est une petite protéine monomérique⁵¹ basique de 93 acides aminés, présente uniquement chez les archées du genre *Methanoarcina* et *Halobacterium*. MC1 présente une haute affinité pour l'ADN double-brin, en particulier s'il est surenroulé négativement.⁵² MC1 est également capable de reconnaître spécifiquement certaines structures de l'ADN telles que des jonctions 4 voies.⁵³ MC1 est un courbeur d'ADN, mais contrairement aux autres protéines architecturales de l'ADN chez les archées, elle courbe l'ADN en se fixant du côté concave de la courbure (**Figure 13 H**). Cette forte courbure impose un angle d'environ 115° à l'ADN double brin comme on peut l'observer sur la structure en solution du complexe et sur les images de microscopie électronique (**Figure 13 H et I**).^{54,55} Cette forte courbure est due à un double mode de compaction par des résidus chargés au niveau du corps rigide de la protéine, tamponnant les phosphates du squelette phosphodiester d'une part, et par l'intercalation de résidus hydrophobes du bras flexible d'autres part.

HU

HU (ou Hta pour histone like protein from *Thermoplasma acidophilum*) est une protéine qui existe principalement chez les bactéries mais que l'on peut aussi trouver chez les archées du genre *Thermoplasma* (**Figure 10 p 30**). Comme on ne trouve cette protéine que dans un seul genre d'archée et qu'elle est très semblable aux protéines bactériennes, il est probable que la présence du gène codant cette protéine chez les archées *Thermoplasma* soit issue d'un transfert horizontal. Il est intéressant de noter que *Thermoplasma* est le seul genre des euryarchées qui ne possède pas d'histones, suggérant une redondance, voir une compétition entre les deux types de protéines de compaction. La structure et les fonctions de la protéine HU seront mieux détaillées plus loin dans ce chapitre. Néanmoins, on peut noter quelques particularités de la protéine de *Thermoplasma acidophilum* par rapport à ses homologues bactériens. D'après des expériences de séquençage menées par Delange R. et collaborateurs,^{56,57} il semble que la majorité des protéines HU chez *Thermoplasma acidophilum* perde leur méthionine N-terminale alors que cette méthionine est strictement conservée chez ses homologues bactériens. On note également un nombre plus important de phenylalanines (5 par monomère contre 3 à 4 chez les bactéries) qui, si le repliement et

l'orientation des chaînes latérales sont conservés, pourraient toutes interagir entre elles dans le cœur hydrophobe du dimère et contribuer à sa plus grande stabilité.

1.3.3. Compaction de l'ADN chez les Bactéries

Généralités

Comme chez les archées, les bactéries ne possèdent pas de noyau et l'ADN non compartimenté est compacté dans une structure nucléo-protéique appelé le **nucléoïde (Figure 14)**. Les bactéries ne possèdent pas d'histones semblables aux histones eucaryotes et on n'observe d'ailleurs pas de structuration en « collier de perle » comme chez les eucaryotes (**Figure 8 p 27**) ou les euryarchées (**Figure 12 p 33**) dans les extraits de chromatine bactérienne. Ces extraits présentent une structuration avec des boucles surenroulées appelées **plectonèmes (Figure 14 A)**.

Contrairement aux plasmides qui peuvent être relaxés par une seule cassure simple brin, la relaxation complète du nucléoïde requiert un nombre important de cassures. Cela suggère que le nucléoïde est subdivisé en domaines topologiques indépendants.⁵⁸ Les expérimentations menées par Postow et collaborateurs⁵⁹ ont montré que ces domaines topologiques mesuraient 10 kb en moyenne, représentant environ 400 domaines topologiques pour un nucléoïde de 4 Mbp. Dans le modèle présenté sur la **Figure 14 B**, ce sont les plectonèmes qui représentent ces domaines topologiques. Les délimitations de ces domaines sont flexibles et peuvent être modulées en fonction du contexte biologique.

Au niveau en deça de ces domaines topologiques, le nucléoïde peut être subdivisé en différentes régions de 800 kb à 1 Mbp que l'on appelle des **macro-domaines**. Ces macro-domaines ont d'abord été identifiés par des expériences de FISH (fluorescence *in situ* hybridation) et de microscopie à fluorescence qui ont permis de déterminer que ces macro-domaines, dénommés Ori (incluant l'oriC) et Ter (incluant la région de terminaison de réplication), avaient une localisation contrôlée dans l'espace de la cellule.^{60,61} Ces domaines sont localisés à l'opposé l'un de l'autre au niveau des pôles dans les nouvelles cellules, alors qu'il migrent vers le milieu de la cellule à l'initiation de la réplication. L'existence de ces macro-domaines a également été mise en évidence par des expériences de recombinaison. Dans ces expériences, Valens M. et collaborateurs⁶² ont mis en évidence des macro-domaines supplémentaires dénommés « Left » et « Right » ainsi que deux régions moins structurées de part et d'autre du macro-domaine Ori.

Ces structurations en macro-domaines du nucléoïde semblent jouer un rôle dans le bon déroulement de la division cellulaire chez les bactéries. En effet, des mutations affectant la localisation des macro-domaines conduisent à la génération de nombreuses cellules dépourvues de nucléoïdes.⁶¹ Cependant les bases moléculaires de la structuration de ces macro-domaines restent mal connues.

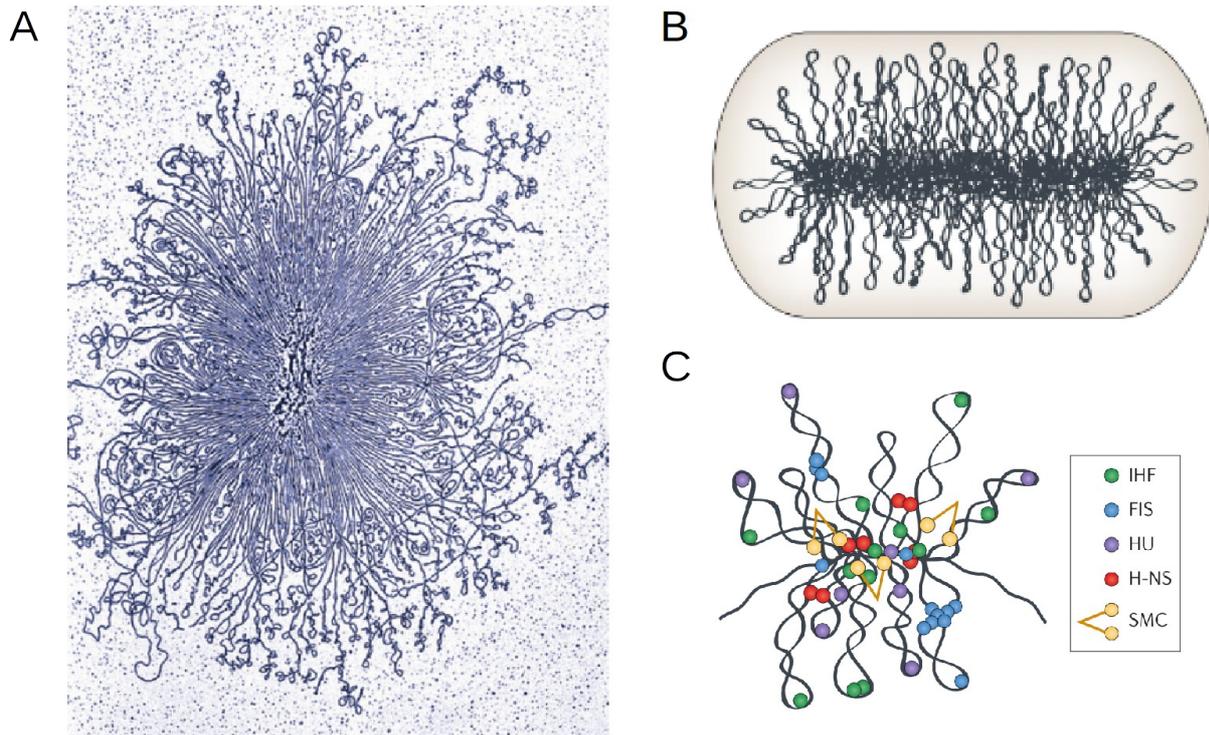


Figure 14: Structuration du nucléoïde bactérien A- Image de microscopie électronique du nucléoïde bactérien extrait d'*Escherichia coli* et coloré à l'acétate d'uranyle (longueur ~20 µm) B- Schéma de nucléoïde dans la cellule bactérienne. C- Représentation schématique des interactions des protéines d'architecture de l'ADN bactérien formant le nucléoïde. (sources: image et schémas reproduit de wang et collaborateur)²⁶⁷

En absence d'histone, la compaction du génome chez les bactéries est opérée par l'action conjointe des topoisomérases (telles que la gyrase) qui sur-enroulent négativement l'ADN (comme on peut le constater au niveau des plectonèmes) et des protéines associées au nucléoïdes (NAPs) qui courbent et pontent des boucles sur-enroulées (**Figure 14 C**).

Les protéines architecturales du nucléoïde bactérien

Les NAPs sont généralement des petites protéines, basiques, abondantes et souvent multimériques. Une quinzaine d'entre elles ont pu être mis en évidence chez les bactéries.^{63,64} Nous allons en décrire quelques unes parmi celles qui sont les mieux caractérisées

SMC

Les protéines SMC (« Structural Maintenance of Chromosomes ») représentent une famille de protéines que l'on retrouve à la fois chez les eucaryotes et les procaryotes, et qui jouent un rôle dans l'organisation et la ségrégation des chromosomes. Contrairement à la plupart des protéines d'architecturales de l'ADN, les protéines SMC forment des complexes de grandes tailles. La protéine SMC de *E.coli*, MukB, constitue un homodimère de 340 kDa et peut s'associer à d'autres facteurs, MukE et MukF, formant le complexe MukBEF. Le travail de Danilova et collaborateurs⁶⁵ a montré que MukB est associé au macro-domaine Ori et que des souches mutantes K.O. pour MukB sont thermosensibles et présentent des anomalies de réplication qui génèrent des cellules anucléées.

Le repliement typique des protéines SMC est une structure en V (**Figure 15 A et B**), avec une région dite « charnière » responsable de la dimérisation. Cette région est prolongée par de long bras flexibles en « coiled-coil » formant un angle pouvant varier de 0 à 180°. Ces longs bras sont terminés par des domaines globulaires ATPasiques qui peuvent former des interactions supplémentaires intra- ou inter-dimères *via* un processus ATP-dépendant.⁶⁶ La structure du domaine charnière prolongé d'une partie du bras coiled-coil a été résolue par Li Y. et collaborateurs⁶⁷ (**Figure 15 B**).

MukB est capable de condenser l'ADN *in vitro* et *in vivo*.⁶⁸⁻⁷¹ Néanmoins les bases moléculaires de cette compaction sont encore mal connues, bien que l'ATP semble pouvoir moduler ce processus de condensation.⁷² Cette protéine apparaît comme la seule protéine d'architecture de l'ADN possédant une activité enzymatique (en excluant les topoisomérases). En l'absence de faces basiques d'interaction avec l'ADN, différents modèles de compaction du chromosome bactérien par MukB, ont été élaborés et sont présentés **Figure 15 C**.⁷³ Dans ces modèles, les brins d'ADN sont piégés entre les bras d'un ou plusieurs dimères de MukB, contraignant ainsi plusieurs régions distantes de l'ADN chromosomique à la manière d'un ponteur d'ADN.

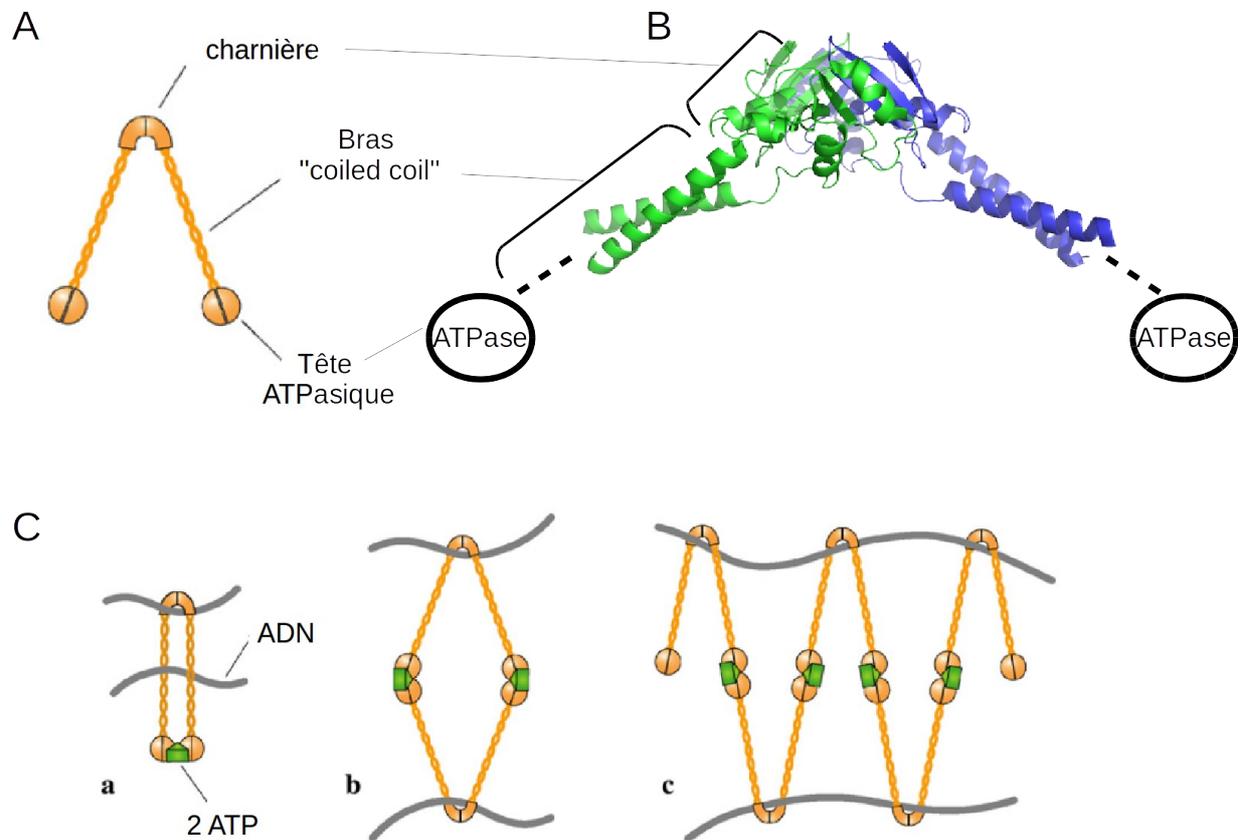


Figure 15: Structure et compaction de l'ADN par MukB (SMC) A/B : Schéma/structure du dimère MukB d'*E. coli*. Une partie du bras et le domaine ATPasique ont été tronqués (code pdb : 2WMM). C- Modèles ATP-dépendant de la compaction de l'ADN par MukB (pontage). Des formes dimériques (a), tétramériques (b), oligomériques (c) sont proposées. (figure adaptée de Thanbichler M. et Shapiro L.)³⁵

Fis

Fis (Factor for Inversion Stimulation) est une des protéines d'architecture du nucléoïde les plus abondantes de la bactérie en phase exponentielle (30000 protéines par bactérie).⁷⁴ Elle a initialement été identifiée pour son rôle dans la recombinaison site-spécifique.⁷⁵⁻⁷⁷ Fis est une protéine homodimérique de 22 kDa.⁷⁵ Cette protéine est capable de lier l'ADN de manière séquence spécifique suivant le consensus : GNNYRNNWNNYRNNC (avec N :G, A,T ou C ; Y : C ou T ; R A ou G).⁷⁸ Cependant Fis est également capable de se fixer sur l'ADN de manière non séquence spécifique. Les structures de Fis libre et en complexe avec l'ADN ont été résolues par cristallographie.^{79,80} Cette protéine est caractérisée par un « β -hairpin » N-terminal suivi de quatre hélices α (**Figure 16 A**). Le β -hairpin et les deux premières hélices sont impliqués dans la dimérisation. Les hélices 1 et 2 de chaque protomère sont enchevêtrées et interagissent ensemble majoritairement *via* des résidus hydrophobes. Les deux hélices C-terminales forment un domaine hélice-tour-hélice (HTH) responsable de l'interaction de Fis avec l'ADN. Fis est capable d'induire une courbure de l'ADN de 65° en se fixant du côté concave de cette courbure (**Figure 16 A**).⁸⁰ La fixation de Fis sur de l'ADN plasmidique entraîne sa condensation (**Figure 16 B**).^{81,82} Il a également été montré que Fis pouvait stabiliser des points de jonctions de brins d'ADN et des boucles.⁸² Ces études montrent que Fis agit de manière directe dans la compaction de l'ADN par fixation. De plus, Fis est capable de moduler l'action des topoisomérases telles que la gyrase et la topoisomérase I *in vitro* et *in vivo*.^{83,84} Cette modulation peut être directe par interférence avec la fixation des topoisomérases sur l'ADN, ou indirecte par modulation de l'expression de leur gène *in vivo*. En effet, Fis est capable de réguler l'expression de nombreux gènes.^{81,85-87}

Il est intéressant de noter certaines propriétés de l'homodimère de Fis. Des études de dénaturation-renaturation par des agents chaotropiques ont été menées par Topping T. et collaborateurs⁸⁸ et mettent en évidence un intermédiaire dimérique transitoire de dénaturation (I_2) de Fis. Les auteurs présentent cette propriété comme possiblement conservée chez les protéines dimériques présentant des domaines hélices enchevêtrées. Cet état intermédiaire dimérique de Fis peut être stabilisé ou supprimé par mutagenèse dirigée.^{89,90} De plus, des protéines de types mutant et sauvage de Fis ont la propriété de pouvoir échanger leurs chaînes pour former des hémiduplexes.⁹¹

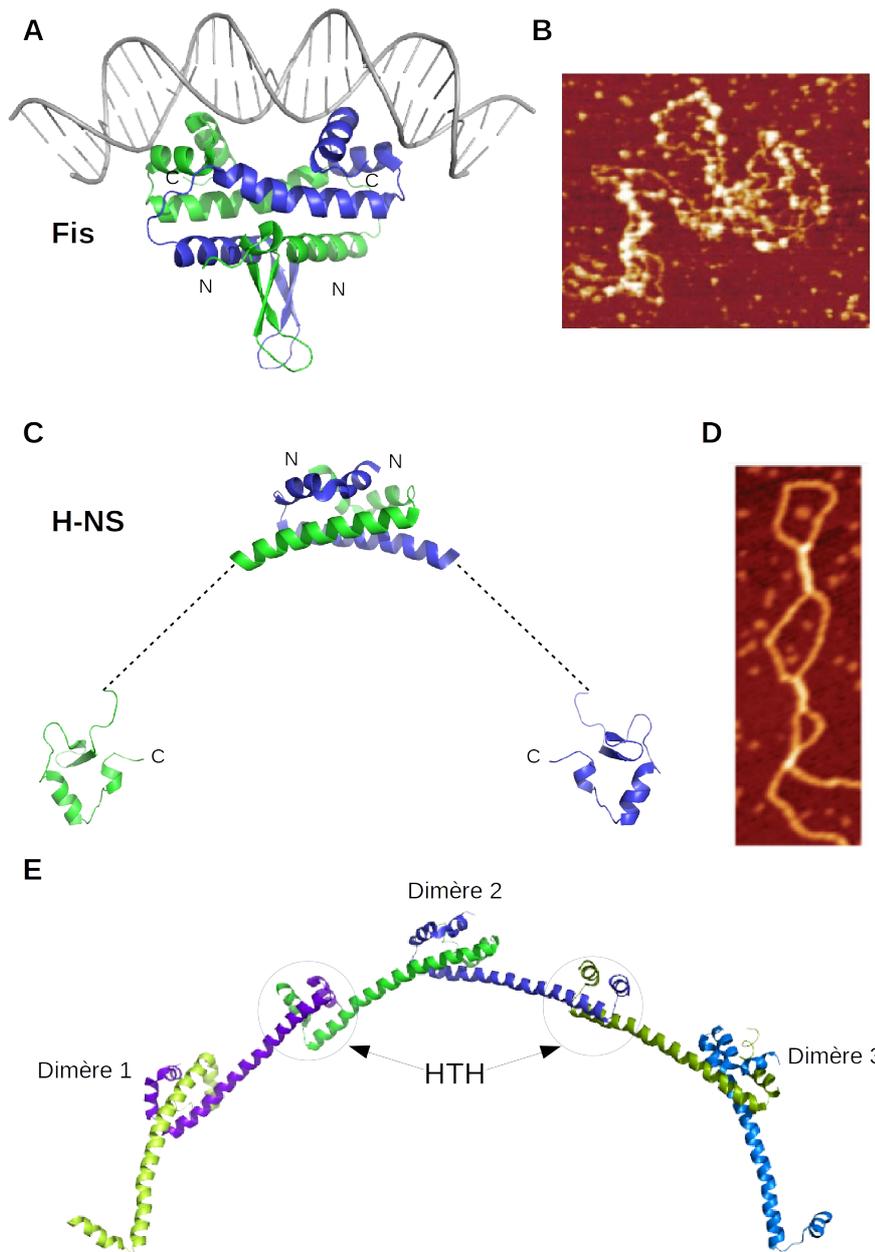


Figure 16: Fis et H-NS: Structure, oligomérisation et compaction de l'ADN A- Structure du complexe Fis-ADN (code pdb : 4IHW). B- Image AFM de Fis compactant un ADN plasmidique (~ 1 dimère / 100 pb). C- Structure du domaine de dimérisation N-ter (code pdb : 1NI8) et du domaine de fixation à l'ADN C-ter (code pdb : 1HNS) de H-NS de *E.coli*. D- Image AFM d'une boucle d'ADN stabilisée par H-NS (source : luijsterberg et collaborateurs¹⁰⁰). E- Oligomérisation supérieure de H-NS suggérée par l'empilement cristallin de la structure du domaine de dimérisation de H-NS de *Salmonella tiphimurium* (code pdb : 3NR7)

H-NS

H-NS (« histones-like nucleoid structuring protein ») est une protéine homodimérique de 31kDa impliquée dans la structuration du nucléoïde bactérien. Cette protéine est divisée en 2 domaines, un domaine de dimérisation et un domaine de fixation à l'ADN tous deux reliés par un segment flexible. Les structures des deux domaines ont été résolues pour la protéine H-NS d'*E. coli*.^{92,93} Le domaine de fixation à l'ADN (côté C-terminal 90-137) présente un motif « β -hairpin » suivi de deux courtes hélices α .⁹² Le domaine de dimérisation (côté N-terminal 1-46) présente une structuration avec trois hélices enchevêtrées, les hélices 3 de chaque protomère faisant des interactions de type « coiled-coil » (**Figure 16 C**).⁹³ Certaines études ont montré que H-NS pouvait former des oligomères d'ordres supérieurs aux dimères.⁹⁴⁻⁹⁶ L'empilement cristallin de la structure du domaine de dimérisation chez *Salmonella typhimurium* (fragment N-terminal 1-82, bien plus long que le fragment cristallisé chez *E. coli*) propose un modèle de cette oligomérisation d'ordre supérieur. Chaque dimère pouvant s'associer à deux autres dimères par l'intermédiaire des domaines hélice-tour-hélice émergeant du « coiled-coil » de dimérisation (**Figure 16 E**).⁹⁷ L'arc de cercle formé par cette structure suggère une stabilisation des boucles plectonémiques par ces grands oligomères de H-NS. Il a été également montré que H-NS était capable de ponter deux brins d'ADN, probablement *via* ses deux domaines de fixation de l'ADN.^{98,99} L'existence de ces deux modes de compaction de l'ADN explique la conformation induite par H-NS sur l'ADN observée *in vitro* (**Figure 16 D**).^{100,101} La protéine H-NS est moins abondante que d'autres protéines d'architecture de l'ADN telles que Fis ou HU. On compte au maximum 20000 H-NS par cellule en phase exponentielle, et ce nombre diminue à 12000 en phase stationnaire. Néanmoins, l'effet « ponteur » d'ADN sur l'architecture globale du nucléoïde ne nécessite pas un nombre aussi important de molécules que les « courbeurs » ou « enrouleurs » d'ADN. Une quantité trop importante de H-NS, comme dans le cadre d'une sur-expression, entraîne une surcompaction du nucléoïde, qui adopte une forme sphérique, et conduit au final à la mort de la bactérie.^{102,103}

En plus de sa fonction de compaction de l'ADN, H-NS est capable de réguler la transcription de nombreux gènes.¹⁰⁴⁻¹⁰⁷ Outre les gènes intrinsèques de la bactérie, certains gènes étrangers issus de transferts horizontaux sont inhibés par H-NS.¹⁰⁴ Une implication de cette protéine dans la recombinaison a également été mise en évidence. Il semble qu'elle soit capable d'inhiber l'intégration de génomes étrangers par recombinaison¹⁰⁸ mais qu'elle stimule la voie de réparation par recombinaison.¹⁰⁹ Les bases moléculaires de ces deux fonctions apparemment antagonistes sont encore mal connues.

IHF

IHF (Integration Host Factor) est une protéine exclusivement hétérodimérique de 20kDa qui a été initialement caractérisée pour son rôle dans l'intégration du génome de bactériophages (λ , μ).¹¹⁰ Elle est capable de se lier à l'ADN de façon séquence spécifique en suivant le consensus : WATCARNNNNTTR (W : A ou T ; R : A ou G ; N : A,T, G ou C) avec une préférence pour une région riche en A/T en amont de cette séquence.¹¹¹ Il y a trois ordres de grandeur de différence entre la fixation séquence spécifique ($K_d \sim 1\text{nM}$) et la fixation non spécifique ($K_d \sim 1\mu\text{M}$) de IHF sur l'ADN.^{111,112}

Les deux chaînes de IHF de *E.coli*, α (94 acides aminés) et β (99 acides aminés), possèdent 25 % d'identité de séquence. La structure de l'hétérodimère présente un motif hélice-tour-hélice N-terminal, suivit d'un feuillet β à trois brins, qui comporte un long bras en β -hairpin entre le brin 2 et le brin 3 et se termine par une petite hélice α C-terminale (**Figure 17 A**).¹¹³ Une proline très conservée au niveau de chaque bras en « β -hairpin » s'intercalent entre les bases de l'ADN et provoque une courbure locale. Les deux courbures induites par l'intercalation des prolines et les interactions électrostatique avec les chaînes latérales des résidus de la « selle » (surface du dimère entre les bras) et du corps (HTH) provoquent une très forte courbure en U pouvant aller jusqu'à 180° .¹¹³ Cette courbure est également observable sur des images de microscopie à force atomique (**Figure 17 C**).¹¹⁴

La courbure induite par IHF au niveau des sites attP aide à la formation de l'intrasome pour la recombinaison.¹¹⁰ L'importance d'IHF a également été rapportée pour d'autres processus telles que la transcription de certains gènes,¹¹⁵⁻¹²¹ la réplication,¹²²⁻¹²⁴ et bien sûr la compaction de l'ADN.^{114,125} Le nombre de dimères d'IHF est estimé à 6000 par cellule en phase exponentielle et augmente jusqu'à environ 27000 en début de phase stationnaire.⁷⁴ Avec environ 1000 sites de fixation suivant le consensus de IHF dans le génome de *E.coli*, il est probable que IHF contribue à la structuration globale du nucléoïde non seulement par des interactions séquence spécifiques mais également par des interactions non spécifiques.

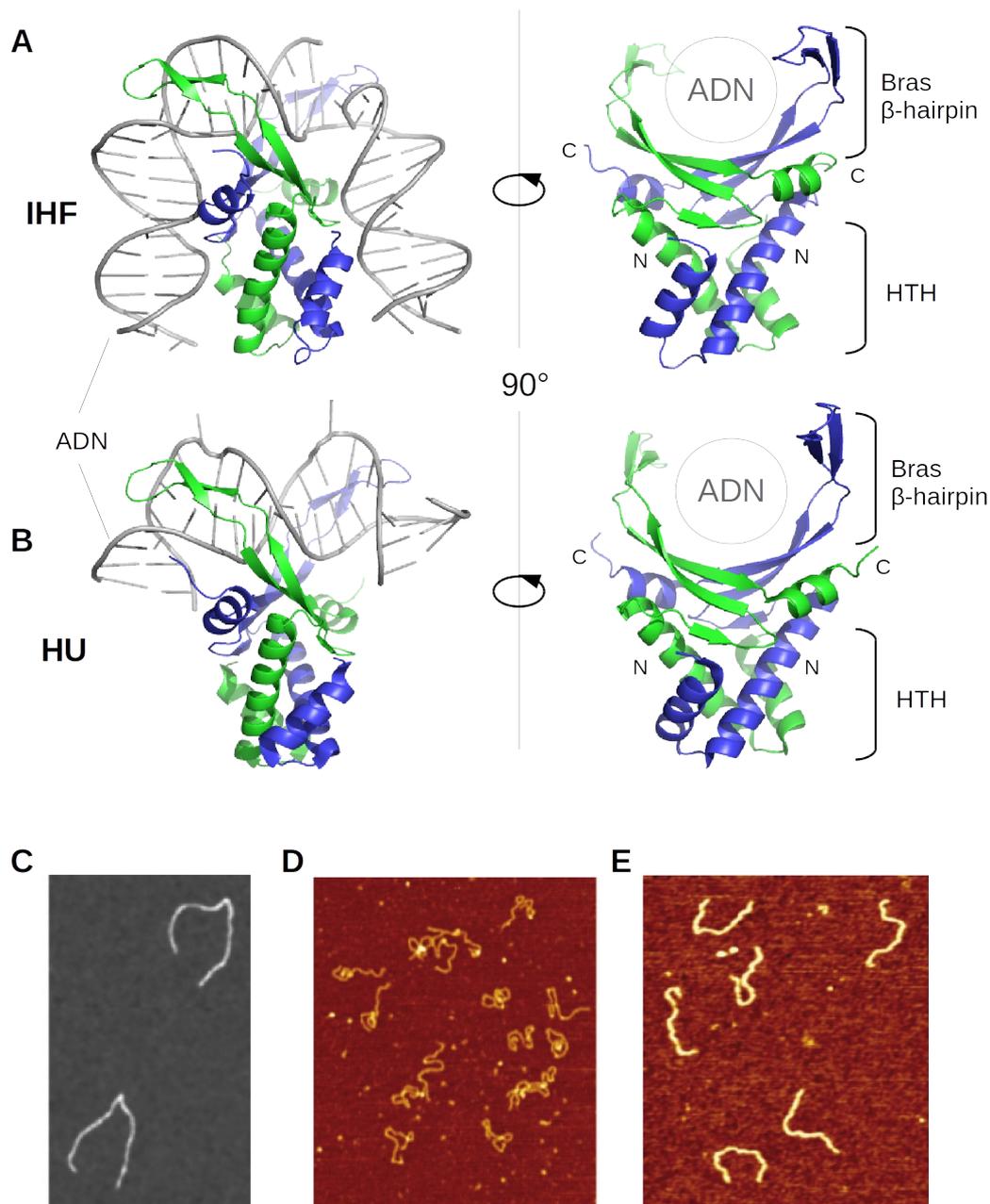


Figure 17: IHF et HU: structure et courbure de l'ADN A- Structure de IHF de *E.coli* avec un ADN comportant une cassure simple brin (code pdb : 1IHF). B- Structure de HU de la cyanobactérie *Anabaena* PCC7120 avec un ADN comportant deux hernies (code pdb : 1P71). C- Image de microscopie à force atomique de l'interaction de IHF avec un ADN de site spécifique de fixation en son centre (source : Dame R.T. et collaborateurs)²⁶² D/E- Images de microscopie à force atomique de complexes HU/ADN (D : 1 dimère / 92 pb ;E : 1 dimère / 1,8 pb) (source : van Noort J. et collaborateurs⁵⁵).

1.4. La protéine HU

1.4.1. Généralités

HU (H pour « histone like » et U pour U93 du nom de la souche depuis laquelle elle a été extraite pour la première fois) est une protéine dimérique de 19 kDa qui possède de nombreuses similarités avec IHF. Elle est très abondante avec 30000 molécules par bactérie en phase exponentielle de croissance.⁷⁴ Le repliement global et la courbure induite à l'ADN est très semblable entre les HU et IHF (**Figure 17 A et B**). Les deux prolines d'intercalation aux extrémités des bras en β -hairpin sont conservées entre HU et IHF. Mais, si IHF reconnaît des bases de sa séquence consensus par des liaisons hydrogène avec certaines de ses arginines, HU forme uniquement des interactions types ponts salins ou liaisons hydrogène avec le squelette phosphodiester ou les désoxyriboses, ce qui explique l'absence de spécificité de séquence dans la reconnaissance de l'ADN par HU.

1.4.2. Les ligands de HU

Bien que HU ne possède pas de réelle spécificité de séquence pour sa fixation à l'ADN, une fixation préférentielle sur les séquences riches en A/T a été observée.^{126,127} HU est également capable de reconnaître avec une haute affinité certaines structures de l'ADN présentant une haute flexibilité (cassure simple-brin et une lacune d'un ou deux nucléotides)¹²⁸⁻¹³⁰, ou encore des portions courbées (fourches et jonctions quatre voies complètes ou incomplètes)^{129,131} qui minimisent l'énergie nécessaire à la torsion de l'ADN par cette protéine. L'affinité de HU pour ces structures est de l'ordre du nanomolaire alors que l'affinité de HU pour un ADN-B double brin quelconque est de l'ordre du micromolaire. HU est également capable de fixer l'ADN simple-brin, ainsi que les ADN double-brins contenant des hernies (**Figure 17 B**) ou des régions mésappariées. Il y a également des évidences de la fixation de HU sur des ARN et des hybrides ARN/ADN ainsi que sur des protéines mimiques de l'ADN.¹³²⁻¹³⁴

1.4.3. Les fonctions biologiques de HU

HU est associée à de nombreux processus ayant trait à l'organisation et au métabolisme de l'ADN chez la bactérie tels que la compaction, la réplication, la transcription et la réparation de l'ADN.

La compaction de l'ADN est la principale fonction attribuée à HU mais représente également la plus controversée.¹³⁵ En effet, en fonction de la souche d'origine et des conditions d'incubation

(concentration en protéine, force ionique, rapport protéine:ADN), différents effets ont été observés sur la conformation de l'ADN *in vitro*.¹³⁶ Les études employant des protéines issues d'organismes thermophiles présentent une compaction de l'ADN accrue avec l'augmentation de la concentration en HU.^{137,138} D'autres études employant la protéine HU d'*E.coli* présentent différents effets sur l'ADN en fonction de la concentration.¹³⁹ A faible rapport HU:ADN, on observe une compaction globale de l'ADN après fixation de la protéine, probablement due à une courbure induite par sa fixation (**Figure 17 D**). Lorsque que l'on augmente le rapport HU:ADN, l'ADN adopte une conformation étendue et rigide avec une augmentation de l'épaisseur apparente (**Figure 17 E**). Les auteurs interprètent ces observations par une polymérisation de la protéine autour de l'ADN, le rigidifiant sans le compacter. Cette observation est comparable à la fibre également formée par la protéine d'archée, Alba, à des rapports protéine:ADN élevés (**Figure 13 C p35**). Une troisième forme d'apparente relaxation de l'ADN a également été observée sur de l'ADN plasmidique qui apparaît parfaitement circulaire en microscopie à force atomique. Cet état semble stimulé en cas de présence de cassures simple brin. Dans ce cas HU s'oppose à l'effet d'autres protéines de compaction telles que H-NS.¹³⁵ D'autres études ont montré que le type de complexe ADN/HU était fortement dépendant de la force ionique et que dans des conditions de salinité physiologique (150 à 200 mM NaCl), seule la fonction de compaction était observée.¹⁴⁰ *In vivo*, la sur-expression de HU n'entraîne pas de sur-condensation du chromosome bactérien et l'absence de gène codant pour HU ne provoque pas sa décondensation.^{102,141} Néanmoins, une double mutation du gène *hupA* (E38K, V42L) altère la conformation du chromosome bactérien qui adopte alors une forme sphérique fortement compactée.¹⁴²

Toutes ces études mettent en avant des rôles contradictoires de HU sur la compaction de l'ADN. Toutefois, on peut considérer HU comme un facteur de remodelage de la chromatine dont l'action peut être modulée en fonction des conditions. La compétition apparente entre les modes de compaction des différentes protéines de type histone pourrait permettre de réguler l'expression différentielle des gènes bactériens en fonction du contexte biologique.

HU module également **la transcription** de nombreux gènes chez la bactérie. Des études comparatives des profils transcriptomiques en présence et en absence des gènes codant pour HU ont été réalisées chez *E. coli* et *S. typhimurium*.^{143,144} Ces études ont montré que HU régule la transcription de gènes impliqués dans la réponse à différents stress : stress acide, stress thermique, stress osmotique, anaérobiose et induction de la réponse SOS. Une partie de cette réponse peut être

indirecte par la régulation de la traduction de l'ARNm de Rpos(σ 38).¹³²

HU est également impliquée dans la régulation du métabolisme. Elle est notamment requise avec GalR dans la répression de l'opéron Gal.¹⁴⁵⁻¹⁴⁸ La courbure de l'ADN induite par HU permet le rapprochement des deux sites de fixation de GalR permettant ainsi sa dimérisation et donc l'inhibition de la transcription de cet opéron. HU régule également des gènes impliqués dans la virulence chez *S.typhimurium*.¹⁴⁴ On peut également noter qu'une double mutation du gène *hupA* chez *E.coli* (E38K-V42L) modifie drastiquement son profil de transcription et peut rendre virulente une souche commensale d'*E.coli*.^{142,149}

Un des principaux phénotypes observés chez les souches K.O. pour les gènes codant pour HU est la formation de cellules anucléées qui s'apparentent à des défauts de réplication et de ségrégation de la chromatine. En effet, des études ont montré que HU est impliquée dans la formation du complexe de réplication au niveau de l'origine de réplication.^{123,150} Des études phylogénétiques ont également montré que HU a suivi une évolution parallèle à d'autres protéines impliquées dans **la réplication et la ségrégation** telles que Dam, SeqA et MukB.¹⁵¹

On peut également souligner le rôle de HU dans la **réparation de l'ADN**. Les souches mutantes K.O. pour les gènes codant pour HU présentent une sensibilité accrue au rayonnement γ et UV.^{152,153} Ce phénotype peut être dû à une protection passive de l'ADN par la protéine ou à une régulation distante de gènes impliquée dans la réparation. Un rôle direct et actif dans les processus de réparation n'est pas exclu. En effet, HU est capable de reconnaître bon nombre d'intermédiaires de réparation : les lacunes, les cassures simple brin, les mésappariements, les fourches et les jonctions 4 voies.^{128,131,154} Un homologue de HU (Hlp), chez *Mycobacterium smegmatis*, participe à la réparation des cassures double-brin en stimulant la religation.¹⁵⁵ Mais cette fonction est associée à une extension C-terminale que l'on ne retrouve pas chez les protéines de *E. coli*. Une autre étude menée par Hashimoto O. et collaborateurs¹⁵⁶ propose un rôle protecteur de HU dans la réparation des lésions proches sur des brins opposés (sites dit « multi-lésés »). Les auteurs proposent que HU permette une réparation séquentielle de ce type de dommages en prenant en charge une cassure simple brin pendant que l'autre dommage est réparé. HU préviendrait alors la formation de cassures double-brin accidentelles.

1.4.4. Multiplicité des gènes codant pour HU

Si IHF est hétérodimérique, HU est un homodimère chez la plupart des eubactéries qui ne possèdent qu'un seul gène codant pour cette protéine. Mais, chez les enterobactéries telles qu'*E. coli*, il existe deux gènes, *hupA* et *hupB*, codant pour les chaînes protéiques HU α et HU β , respectivement.^{157,158} Deux hypothèses ont été émises pour expliquer la dualité des gènes codant pour HU chez les entérobactéries.¹⁵⁸ La première repose sur la duplication du gène codant pour HU chez un ancêtre commun, de façon analogue à la duplication des gènes codant pour IHF. Les deux gènes résultants, *hupA* et *hupB*, ayant ensuite évolués séparément vers les gènes codant actuellement pour les différentes protéines HU chez les enterobactéries. La seconde hypothèse implique un transfert horizontal entre deux bactéries ancestrales possédant chacun un gène différencié codant pour HU, les ancêtres des gènes *hupA* et *hupB*. Ainsi, on aurait trois groupes descendant de cet événement : un groupe possédant un seul gène de type *hupA*, un autre possédant un seul gène de type *hupB* et un troisième possédant les deux gènes *hupA* et *hupB*. Chacun de ces groupes ayant évolué séparément. L'hypothèse du transfert horizontal est privilégiée par rapport à la duplication car elle permet d'expliquer l'existence de HU « de type α » ou « de type β » chez les bactéries possédant un seul gène codant pour HU selon leur similarité pour l'une ou l'autre des protéines HU entérobactériennes.¹⁵⁸

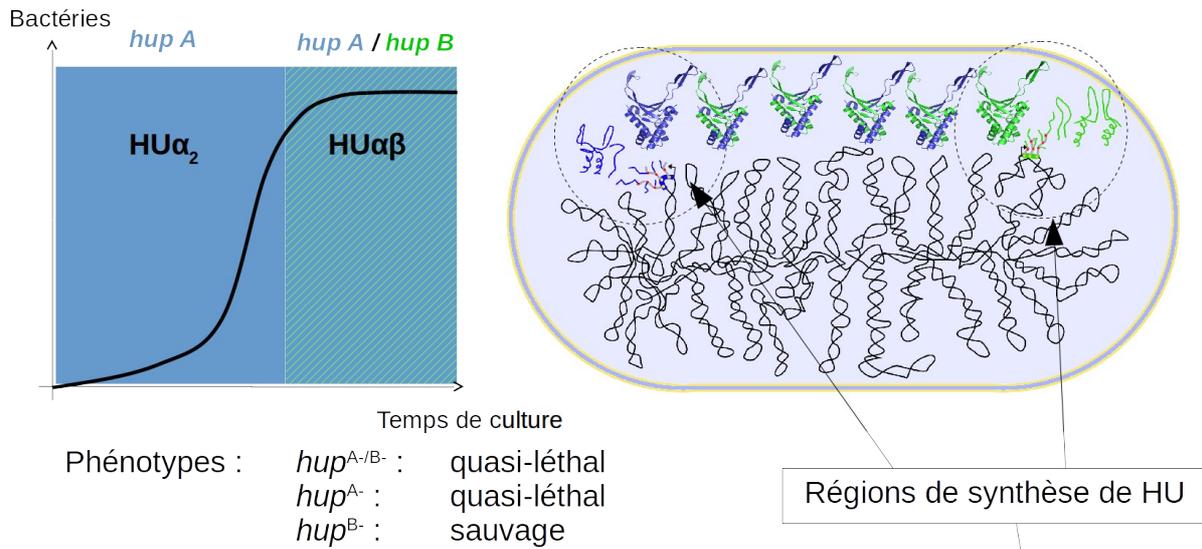
La redondance des gènes codant pour HU a également été mise en évidence chez les pseudomonales. En particulier, trois gènes codant pour des protéines HU-like (en plus de ce codant pour IHF) ont été mis en évidence chez *Pseudomonas putida* : *hupB*, *hupN* et *hupP*.^{159,160} Si *hupB* semble avoir une origine commune avec le gène du même nom chez *E.coli*, les auteurs présentent *hupN* et *hupP* comme des produits de duplication de *hupB* chez un ancêtre commun possédant ce seul gène codant pour HU. Les fonctions de ces nouvelles protéines HU chez *Pseudomonas putida* sont encore mal connues. Les souches K.O. pour un seul des deux gènes *hupB* ou *hupN* ne présentent pas de phénotypes particuliers et ont des comportements de croissance et de viabilité similaires à la souche de type sauvage. Cependant la double délétion de ces deux gènes est létale pour cette bactérie.¹⁶⁰ Ceci suggère que les deux protéines issues des gènes *hupN* et *hupB* possèdent des fonctions redondantes essentielles à la survie de la bactérie qui ne sont pas compensées par la protéine codée par le gène *hupP* dont la fonction a divergé de ses protéines paralogues, et reste encore inconnue.

1.4.5. Régulation et fonctions des isoformes de HU d'*E.coli*

Contrairement aux HU de *Pseudomonas putida*, les trois formes dimériques de HU d'*E. coli* ont fait l'objet de nombreuses études qui ont démontré des fonctions et une régulation différente pour chaque dimère. Les gènes *hupA* et *hupB* ne sont pas organisés en opéron et sont positionnés sur le chromosome bactérien à 90 et 10 minutes, respectivement (séparés par environ 900kpb).^{161,162} Les deux gènes sont placés sous le contrôle de promoteurs différents. *hupA* possède un unique promoteur alors que *hupB* présente quatre promoteurs alternatifs (P1 à P4).¹⁶³ L'analyse des transcrits a montré que les ARNm issus de la transcription à partir des promoteurs P2 et P3 sont les plus abondants alors que ceux issus de P4 le sont peu et ceux issus de P1 ne sont pas détectables. La transcription des deux gènes *hupA* et *hupB* est modulée par Fis et CRP. Dans le cas de *hupA*, Fis et CRP sont des activateurs de transcription et semblent avoir des effets additifs. Dans le cas de *hupB*, CRP a toujours un effet activateur mais FIS devient inhibiteur de la transcription.¹⁶³ Cette différence provient probablement du positionnement différent des sites de fixation de Fis pour les deux gènes. Fis étant très abondante en phase exponentielle uniquement, on comprend aisément que l'expression de chaque gène de HU dépend de la phase de croissance de la bactérie (**Figure 18 A**). En phase exponentielle, seul le gène *hupA* est activé et l'homodimère HU α_2 est le plus abondant.¹⁶⁴ En phase stationnaire, Fis devient très peu abondant, et les deux gènes sont activement transcrits, activés par CRP. L'hétérodimère devient l'espèce majoritaire et améliore la survie de la bactérie en phase stationnaire tardive, lorsque les nutriments se raréfient.¹⁶⁴ Dans des conditions normales de culture, à une température optimale de 37°C, l'homodimère HU β_2 n'est pas indispensable aux bactéries. La viabilité n'est pas altérée pour les souches K.O. pour *hupB* cultivées à 37°C. Toutefois, ces souches ne sont pas viables en cas de choc froid. C'est dans ces conditions que l'homodimère HU β_2 devient l'homodimère majoritaire. Le gène *hupA* est alors réprimé tandis que *hupB* est activement transcrit (**Figure 18 B**). Il semble donc que HU β_2 soit nécessaire à l'adaptation aux chocs froids chez *E. coli*, comme l'est l'ADN gyrase.¹⁶⁵

La structure quaternaire des dimères de HU d'*E. coli* est dépendante des conditions de culture et de l'environnement des bactéries. Chaque dimère possède des propriétés différentes de fixation sur l'ADN¹³⁰, de régulation des gènes¹²⁷ et de stabilité thermodynamique¹⁶⁶. Cette dernière a été étudiée par Dichroïsme Circulaire (CD) et microcalorimétrie différentielle à balayage (DSC) et montre que chaque dimère suit une dénaturation thermique en deux étapes (N₂ → I₂ → 2D).¹⁶⁶ Chaque étape possède une température de fusion caractéristique (T_{m1} et T_{m2}) (**Tableau 2**).

A- Conditions standard : culture à 37°C



B- Conditions de choc froid : 37°C → 10°C

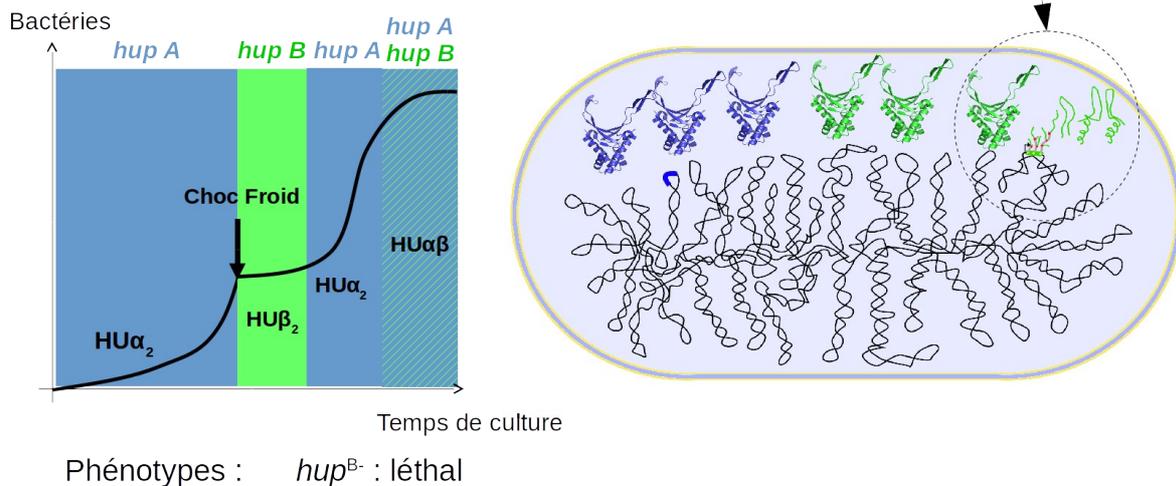


Figure 18: Expression différentielle des gènes codant pour HU et hétérodimérisation en fonction des conditions de cultures et de la phase de croissance Représentation schématique des courbes de croissance de populations bactériennes cultivées suivant deux conditions différentes sont présentées. Les gènes transcrits codant pour HU sont indiqués au dessus des graphiques (bleu: *hupA* seul ; vert: *hupB* seul et hachures vert/bleu : *hupA+hupB*). Les protéines majoritaires associées à chaque période de temps sont indiquées en noir. L'état de transcription des gènes,¹⁵⁷ les protéines majoritaires⁴⁶ et les phénotypes associés à l'absence des gènes codant pour HU dans les différentes conditions¹⁶⁷ proviennent de données issues de la littérature. Des schémas de bactéries présentent l'état de transcription des gènes *hup* (*hupA* en bleu et *hupB* en vert) et la formation de l'hétérodimère (ou non) en fonction de la température de culture.

Le T_{m1} , indépendant de la concentration en protéine, correspond à la transition de la structure dimérique dite « native » N_2 en une structure dimérique alternative dite « intermédiaire » (I_2). Le T_{m1} proches de la température de croissance optimale (voir très inférieure pour $HU\beta_2$) montre que les deux conformations dimériques, N_2 et I_2 , sont probablement présentes chez la bactérie. Le T_{m2} correspond à la dissociation du dimère intermédiaire en monomères déstructurés. Les valeurs élevées de T_{m2} , à des concentrations et forces ioniques physiologiques (**en gras tableau 2**), montrent que la forme monomérique doit être absente ou extrêmement minoritaire dans la bactérie (le monomère est forcément très déstructuré et donc très sensible aux protéases).

Tableau 2: Température de fusion des dimères de HU chez *E. coli* Avec T_{m1} la température de fusion $N_2 \rightarrow I_2$ et T_{m2} la température de fusion $I_2 \rightarrow 2D$. Données provenant du travail de Ramstein et collaborateurs.¹⁶⁶ Les conditions les plus proches des conditions physiologiques sont indiquées en gras (1 mg/ml \approx 50 μ M \approx 30000 dimères par cellule d'1 μ m³).

	Températures de fusion		Conditions	
	T_{m1}	T_{m2}	[HU]	[NaCl]
EcHU α_2	41°C	50°C	1 mg/mL	0,2 M
	40°C	51°C	4,5 mg/mL	0,2 M
	59°C	69°C	1 mg/mL	1 M
EcHU β_2	27°C	56°C	1 mg/mL	0,2 M
	26°C	59°C	4,5 mg/mL	0,2 M
	50°C	65°C	1 mg/mL	1 M
EcHU $\alpha\beta$	38°C	53°C	1 mg/mL	0,2 M
	41°C	62°C	4,5 mg/mL	0,2 M
	60°C	72°C	1 mg/mL	1 M

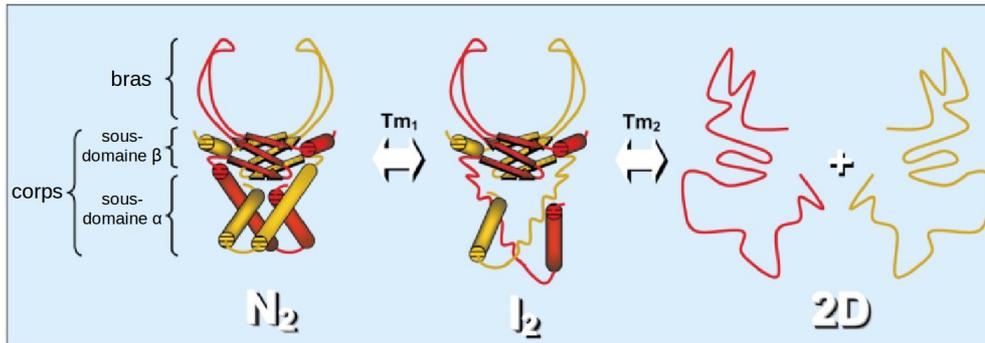
1.5. L'échange de chaînes, un mécanisme essentiel à la formation de l'hétérodimère de HU d'*E. coli*

Comme on l'a vu précédemment, la modulation de la structure quaternaire de HU chez *E. coli* est contrôlée par le stade de croissance des bactéries et les conditions de culture. Chaque gène est contrôlé pour être transcrit au moment utile afin d'obtenir le bon dimère de HU. Cependant le mécanisme qui permet la formation (ou non) de l'hétérodimère reste inconnu.

En effet les gènes *hupA* et *hupB* sont très distants en terme de séquence (à 90,5min et 9,9min respectivement, soit ~900kpb d'écart) et appartiennent à des macro-domaines différents (ORI et NS1 respectivement). Or, comme la transcription et la traduction sont couplées chez la bactérie, la première protéine rencontrée par un monomère de HU néosynthétisé est le monomère synthétisé précédemment. Ceci signifie que les premières espèces dimériques de HU formées sont des homodimères dans les régions de synthèse de HU (**Figure 18**). Ainsi, le moyen le plus simple envisagé pour former l'hétérodimère à partir de ces homodimères seraient une dissociation-réassociation des chaînes des deux homodimères pour former des hétérodimères. Néanmoins, les études de dénaturation thermique des protéines HU d'*E. coli* présentent des températures de dissociation (T_{m2}) bien au-delà des températures optimales de croissance de cet organisme (14 à 19°C de plus, **Tableau 2**). *In vitro*, des expériences de mélange des homodimères purifiés ont mis en évidence que l'hétérodimère se forme à des températures aussi basses que 27°C.¹⁶⁶ Cependant la vitesse d'échange diminue à basse température.

Ainsi, la formation de l'hétérodimère à partir des homodimères, *in vivo* ou *in vitro*, doit suivre un mécanisme différent qu'il reste à élucider. Les dimères de HU d'*E. coli* sont en équilibre entre deux conformations dimériques alternatives à température physiologique : N_2 et I_2 . Un modèle d'échange à partir des formes dimériques I_2 a été proposé par Garnier N. et collaborateurs.¹⁶⁷ Dans ce modèle, les auteurs proposent que la diminution de structuration en hélice α soit localisée au niveau du sous-domaine HTH du « corps » (**Figure 19 A**) et que des déterminants structuraux d'interaction ainsi libérés soient nécessaires pour former un hétéro-tétramère. Ce tétramère se dissocierait ensuite pour former des hétérodimères (**Figure 19 B**).

A



B

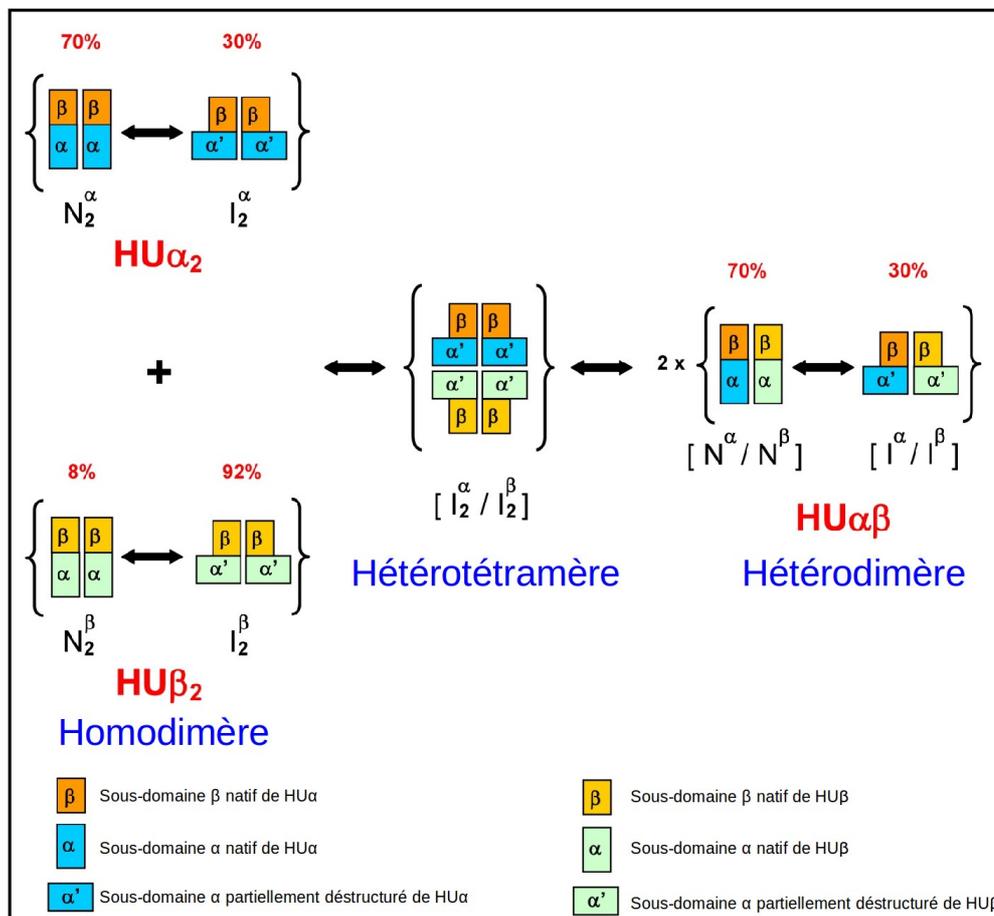


Figure 19: Modèles de dénaturation thermique et d'échange de chaînes des dimères de HU d'*E. coli* A- Modèle de dénaturation thermique $N_2 \rightarrow I_2 \rightarrow 2D$. (adapté de Ramstein J. et collaborateurs 2003)²⁶² B- Modèle d'échange de chaînes à 310K à partir des conformations intermédiaires de HU formant un tétramère transitoire. (adapté de Garnier et collaborateurs 2011)²⁶⁶

Des études d'échange de chaînes ont été réalisées par Andera et collaborateurs sur un homologue de HU appelé TF1, qui provient du phage lytique de *Bacillus subtilis* SPO1. Dans cette étude, les auteurs ont montré que la protéine sauvage pouvait échanger ses chaînes avec une protéine de type mutant (F61R), formant un hétéroduplex. Les dimères de types sauvage et mutant présentent des mobilités électrophorétiques différentes sur gel retard avec un ADN double brin. Ces différences de mobilité sont probablement dues au fait que chaque homodimère (sauvage ou mutant) ne courbe pas l'ADN de la même manière. L'hétéroduplexe présente des propriétés intermédiaires et migre à une position moyenne entre les deux formes homodimériques. Les auteurs ont observé que la cinétique d'échange était dépendante de la température et inhibée par un excès d'ADN. Ils ont également observé que la cinétique d'échange était dépendante de la concentration à une puissance supérieure à 1, ce qui est incompatible avec un modèle d'échange par dissociation/réassociation. Toutes ces observations sont cohérentes avec le modèle d'échange proposé pour les HU de *E. coli*. On peut imaginer que ce mécanisme d'échange soit commun aux protéines présentant ce type de structure. La structure de TF1 a été résolue et est présentée (**Figure 20 A**) en comparaison avec la structure de HU α_2 de *E. coli* (**Figure 20 B**). TF1 présente une homologie structurale forte avec HU et est constituée d'un domaine Hélice-Tour-Hélice, suivi d'un feuillet β terminé par une petite hélice C-terminale. L'échange de sous-unités a également été étudiée entre deux HU orthologues provenant des organismes *Bacillus stearotherophilus* et *Bacillus subtilis* par spectrométrie de masse en conditions natives.¹⁴² Ces deux protéines partagent 86 % d'identité de séquence mais la formation de leur hétéroduplexe est très défavorisée et prend plusieurs jours pour atteindre l'équilibre, sauf dans le cas où l'on chauffe le mélange à 80°C. Les auteurs proposent un mécanisme d'échange par dissociation-réassociation qui semble très inefficace pour l'échange de chaînes entre ces dimères. Ce modèle est cohérent pour ces protéines qui semblent dépourvues de forme dimérique alternative I₂ d'après des expériences de microcalorimétrie et de dichroïsme circulaire.^{143,144}

Bien qu'un mécanisme actif d'échange de chaînes soit uniquement pertinent chez des organismes possédant plusieurs isoformes de HU (enterobactéries, vibrio ou pseudomonales), il est assez étonnant que des protéines avec des structures et des séquences si semblables aient des propriétés thermodynamiques si éloignées. La structure globale de HU de *Bacillus stearotherophilus* (BstHU) (**Figure 20 C**) est très similaire à celle de HU de *Escherichia coli* (EchU) (**Figure 20 B**) (RMSD <1Å) et la séquence est identique entre 50 et 60 % entre les protéines de *Escherichia coli* et celles de *Bacillus stearotherophilus* ou *Bacillus subtilis* (**Figure 20 D**). Néanmoins des mutations ponctuelles ont déjà permis d'observer un changement de chemin de dénaturation (passage de

$N_2 \rightarrow 2D$ à $N_2 \rightarrow I_2 \rightarrow 2D$ ou l'inverse) chez d'autres dimères de protéines histones. Fis possède une forme intermédiaire cinétique I_2 qui peut être stabilisée par une mutation ponctuelle, P61A, ou être supprimée par la formation d'un pont disulfure intermoléculaire chez le mutant V58C.^{65,66} Ces travaux montrent qu'une simple mutation ponctuelle peut profondément altérer les propriétés thermodynamiques d'une protéine telle que la présence ou non d'un intermédiaire dimérique de dénaturation.

Des formes I_2 ont également été mises en évidence chez d'autres protéines dimériques se fixant sur l'ADN comme le répresseur de l'opéron tryptophane Trp, ou plus récemment chez l'hétérodimère d'histone H2A/H2B.^{145,146} Un des points communs entre toutes ses protéines est la dimérisation par des domaines hélices enchevêtrées permettant des contacts hydrophobes inter-chaînes. On peut imaginer que le rôle de l'intermédiaire dans l'échange de chaînes chez HU puisse être extrapolé à d'autres protéines oligomériques dont les fonctions biologiques sont modulées par leur structuration quaternaire. Ce pourrait être le cas par exemple de l'échange des dimères d'histones eucaryotes avec des variants au sein du nucléosome eucaryote.

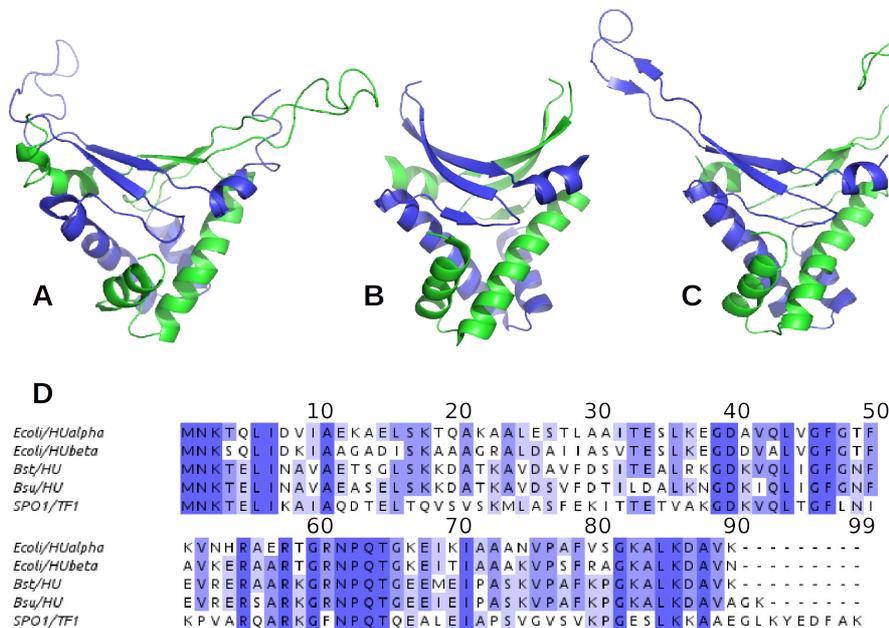


Figure 20: Comparaison des structures et des séquences de HU et de ses homologues qui présentent (ou non) un échange de chaînes A- Structure de l'homodimère TF1 du phage SPO1 (code pdb : 1WTU). B- Structure de l'homodimère $HU\alpha_2$ d'*Escherichia coli* (code pdb : 1MUL). Les bras hautement dynamiques ne sont pas visibles C- Structure de homodimère HU de *Geobacillus stearothermophilus* (code pdb : 1HUE). D- Alignements des structures primaires. Coloration par pourcentage d'identité.

1.6. Objectifs et plan de la thèse

L'objectif principal de ce travail de thèse est d'élucider au niveau atomique les bases moléculaires du mécanisme d'échange de chaînes entre les homodimères de la protéine HU pour former l'hétérodimère. Pour atteindre cet objectif, nous avons adopté une approche multidisciplinaire en utilisant différentes techniques de caractérisation biophysiques (cristallographie des rayons X, RMN, simulation par dynamique moléculaire, spectrométrie de masse) et biochimique (pontage, mutagenèse dirigée) pour caractériser les différentes étapes du mécanisme proposé (**Figure 21**).

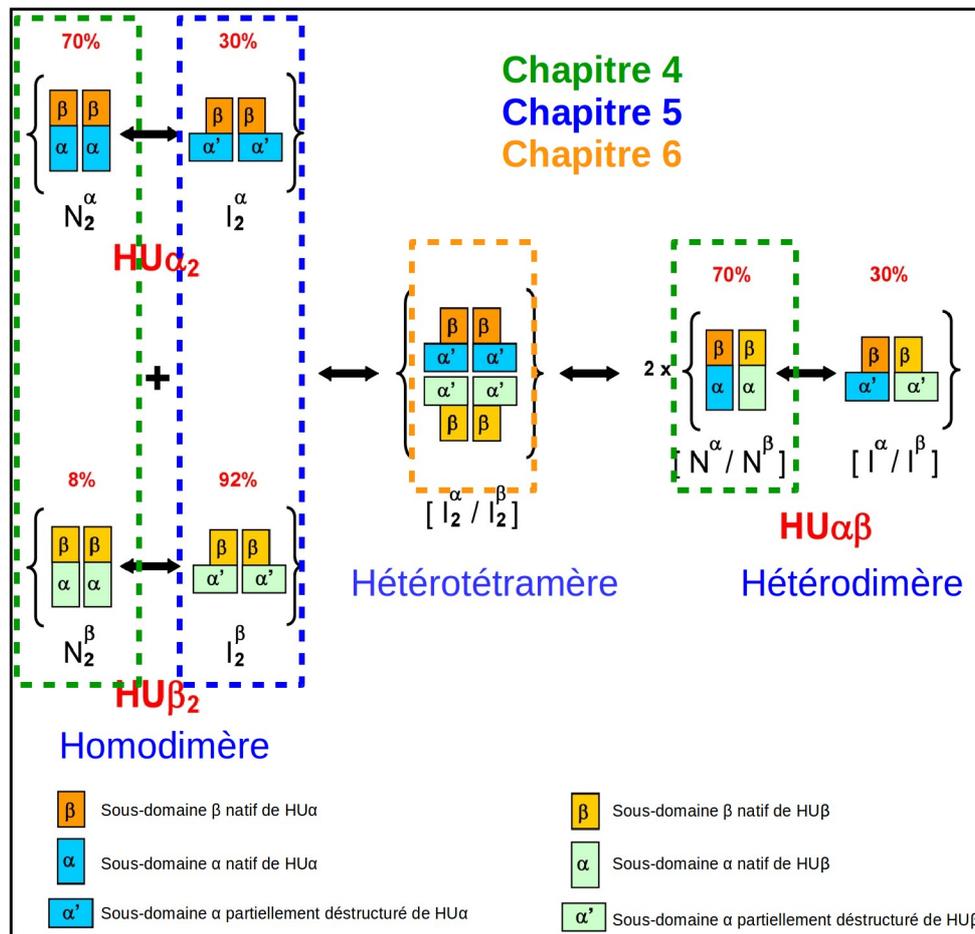


Figure 21: Mécanisme d'échange de chaînes des dimères de la protéine HU (adapté de Ramstein et collaborateurs)¹⁶⁶

La production de versions mutantes des HU d'*E.coli* et des HU orthologues a été initiée pour mettre en évidence les résidus clefs nécessaires à l'échange. La stratégie et le travail de biologie moléculaire réalisé pour cet objectif sont présentés dans le **chapitre 3**.

Le **chapitre 4** correspond à l'analyse structurales des protéines HU dans leurs conformations natives. Aux données disponibles dans la PDB s'ajoute la structure de deux nouvelles HU résolues par cristallographie des rayons X au cours de cette thèse : la protéine HU β_2 de *E.coli* et la protéine HU de *L.lactis*. L'analyse comparative de toutes ces structures et des analyses RMN des homodimères permettront de suggérer des paramètres structuraux pouvant conduire à la formation d'une conformation intermédiaire et à l'échange de chaîne.

Le **chapitre 5** décrit l'analyse de la transition des HU vers la conformation intermédiaire I₂. Les paramètres cinétiques et structuraux de cette transition sont caractérisés par RMN. Les données de RMN et dynamique moléculaire sont exploitées pour proposer un modèle de cette conformation.

Le **chapitre 6** correspond à la mise en évidence du tétramère par pontage chimique et par spectrométrie de masse native. Un mutant cystéine a été mis au point dans l'optique de stabiliser ce tétramère de manière homogène par un pont disulfure. Un protocole d'oxydation et de pontage a été mis au point et sera transposable à d'autres cibles de mutagenèse.

Le **chapitre 7** correspond à la caractérisation cinétique de l'échange de chaîne suivie par spectrométrie de masses en conditions natives.

Dans un chapitre supplémentaire (**chapitre connexe**), nous nous sommes intéressés au rôle de HU dans la réparation de l'ADN par la stimulation du « *turn-over* » de la Formamidopyrimidine-ADN glycosylase (Fpg), une glycosylase bi-fonctionnelle impliquée dans le système de Réparation par Excision de Base (BER). Des analyses biochimiques et un modèle du mécanisme sont présentés.

Chapitre 2 : Matériels et Méthodes

2.1. Transformation bactérienne

50 µL de bactéries calcium-compétentes sont incubés avec 10-100 ng de plasmide pendant 30 min dans la glace. On applique un choc thermique en incubant 45 secondes à 42 °C. On replace ensuite rapidement les échantillons dans la glace. Après 2 min d'équilibration, on supplémente à 1 mL avec du milieu LB et on place la culture à 37 °C sous agitation. Après 1 h d'incubation, on étale 100 µL de culture non diluée ou diluée au 10^e sur des boîtes de Pétri contenant du LB-Agar supplémenté d'ampicilline (100 µg/mL). On place les boîtes dans des étuves à 37 °C sur la nuit. Les colonies obtenues sont ensuite utilisées pour ensemercer des cultures plus importantes (Expressions de protéines ou préparations de plasmides).

2.2. Production des protéines

2.2.1. Expression des protéines de type sauvage non marquées

Les surproductions des protéines HU de type sauvage d'*E.coli* a été précédemment décrite.¹⁶⁸ On ensemence du milieu LB (100-200 mL) supplémenté d'ampicilline (100 µg/mL) avec des bactéries BL21-DE3 (New England Biolabs) transformées par le plasmide pJES-hupA/B. On laisse pousser sur la nuit à 37°C en erlenmeyer (1/10 du volume) sous agitation.

On utilise cette pré-culture pour ensemercer la culture de travail (4-8 x 1 L) à une turbidité de 0,1 UA_{600nm} que l'on incube à 37°C sous agitation. Entre 0,7 et 1 UA_{600nm}, on induit l'expression par ajout de 0,5mM final d'IPTG. Après 2h d'incubation, les bactéries sont collectées par centrifugation (6000rpm, 20min, 4°C). Les culots sont resuspendus dans du tampon de lavage (10 mM Tris/Hcl pH 7,5, 100 mM EDTA, 100 mM NaCl) et transférés dans des tubes falcon de 50mL. On centrifuge à 5000rpm pendant 30min à 4°C puis on élimine le surnageant. Les culots humides sont conservés à -20°C.

2.2.2. Expression des protéines de type sauvage avec marquage isotopique ¹⁵N/¹³C

Comme précédemment, on ensemence du milieu LB (100-200 mL) supplémenté d'ampicilline (100µg/mL) avec des bactéries BL21 (DE3) transformées par le plasmide pJES-hupA/B. On laisse

pousser sur la nuit à 37°C en erlenmeyer (1/5 du volume maximum) sous agitation.

On utilise cette pré-culture pour ensemer la culture de travail (4-8x 1 L) à une turbidité de 0,1 UA_{600nm} que l'on incube à 37 °C sous agitation. Entre 0,7 et 1UA_{600nm}, on collecte les bactéries par centrifugation (6000 rpm rotor SLA-3000, 20 min, 4 °C). On élimine le surnageant puis on resuspend les culots dans du milieu M9 base (on dilue ainsi les traces de milieu LB non éliminé). On collecte par centrifugation puis on resuspend dans du milieu M9 complet contenant du ¹⁵NH₄Cl et du ¹³C-glucose comme seules sources d'azote et de carbone, respectivement. Après une heure d'incubation à 37°C sous agitation on induit par ajout de 0,5mM final d'IPTG. Après 2h d'incubation, les bactéries sont collectées par centrifugation (6000 rpm rotor SLA-3000, 20min, 4°C). Les culots sont resuspendus dans du tampon de lavage (10mM Tris/Hcl pH7,5, 100mM EDTA, 100mM NaCl) et transférés dans des tubes falcon de 50mL. On centrifuge à 5000rpm pendant 30min à 4°C puis on élimine le surnageant. Les culots humide sont conservés à -20°C.

2.2.3. Expression des protéines de type mutant fusionnées à la « Maltose Binding Protein » (MBP)

La production des variants des isoformes de HU d'*E. coli* fusionnées à la MBP se fait de manière analogue à celle utilisée pour les protéines sauvages à l'exception prêt qu'on utilise du milieu LB supplémenté par du glucose (2 % m/V final). L'utilisation d'un milieu riche en glucose permet d'inhiber la production d'amylase qui pourrait dégrader la résine d'amylose lors des premières étapes de purification de la protéine fusionnée.

2.3. Purification des protéines

2.3.1. Purification des protéines de type sauvage

Les protéines de type sauvage ont été purifiées en adaptant le protocole publié par Coste et collaborateurs¹⁶⁸. Les culots bactériens ont été repris dans du tampon de lyse (25 mM Tris/HCl pH 7,5, 400 mM NaCl, 1 mM EDTA, 0,1 mM PMSF, 2 % Glycérol, 10 mM β-mercaptoéthanol) à raison de 10 mL par 1 g de culot bactérien. La suspension cellulaire est supplémentée par du triton X-100 (0,1 % final) et du lysosyme (1 mg/mL final). Après 10 min d'incubation en bain marie à 30°C, on réalise 3 cycles de congélation (azote liquide) /décongélation (bain-marie 30°C). Le lysat est centrifugé 1h à 35000 rpm (rotor 45Ti) à 4°C. Le surnageant est récupéré et constitue l'extrait brut (« EB »).

L'extrait brut est peu à peu complétement à 40 % de saturation en sulfate d'ammonium et précipité pendant 2 h à 4°C sous agitation. La suspension est centrifugée 1 h à 4°C et 35000 rpm (rotor 45Ti). Le surnageant est filtré à 0,2 µm et dialysé 1h contre du tampon SP-A +5% SPB (SP-B = SP-A + 1 M NaCl) puis sur la nuit contre du tampon SP-A +10%B (100 mM NaCl final).

Le dialysat (« Dia ») est filtré à 0,45 µm puis chargé sur une colonne de chromatographie échangeuse de cation SP-FF préalablement équilibrée dans du tampon SP-A +10%B. Après 10 CV (volumes de colonne) de lavage, l'élution est réalisée par un gradient linéaire de 10 à 60 % SP-B (600 mM NaCl) sur 25 CV. Le pic majoritaire en absorption à 280nm correspond au lysosyme (~15kDa) (HU coélue partiellement avec le lysosyme). Les fractions collectées au cours de l'élution sont analysées par SDS-PAGE et les fractions contenant HU (**en rouge Figure 22**) sont rassemblées (« FF »).

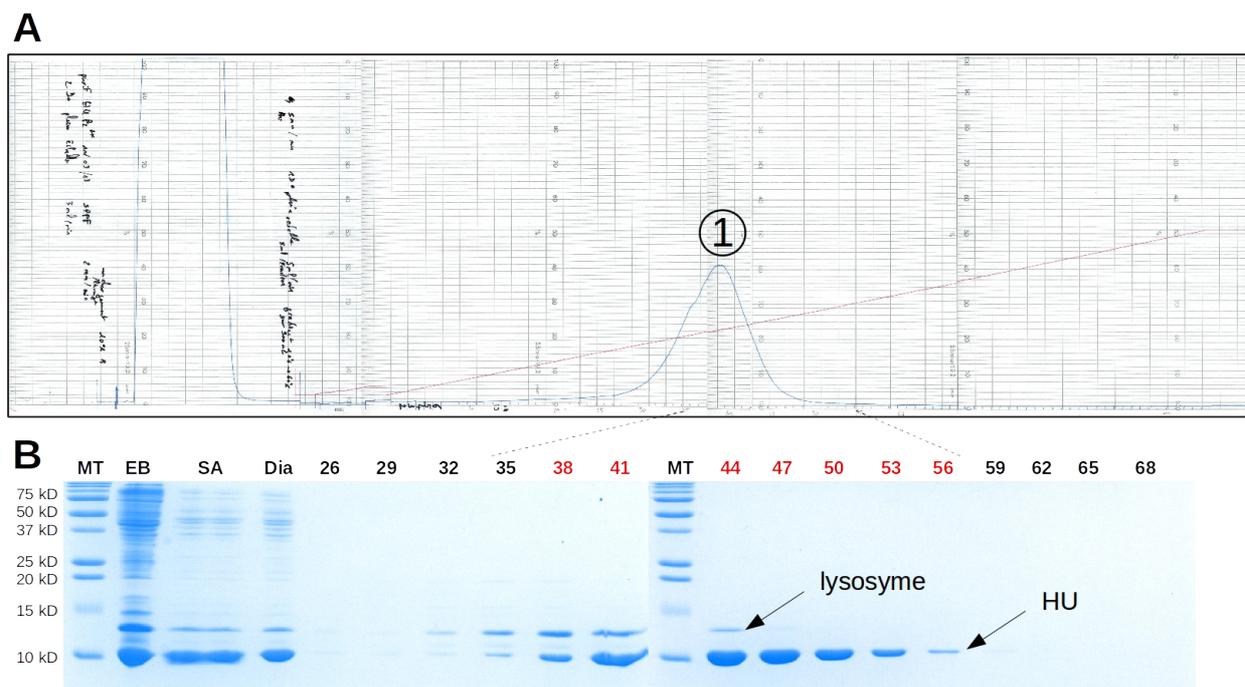


Figure 22: Analyse de l'élution de la colonne SP-FF A- Chromatogramme de l'élution avec, en rouge la proportion de tampon SP-B (gradient en NaCl), et en bleu l'absorbance à 280nm. B- Analyse en gel SDS-PAGE 16% avec « MT », le marqueurs de taille, « EB », l'extrait brut et « Dia », le dialysat. Les numéros sous les pistes correspondent aux numéros des fractions. Ceux en rouges correspondent aux fractions qui ont été rassemblées pour les prochaines étapes de purification.

La fraction FF est concentrée à l'aide d'un concentrateur Centricon™ Amicon™ (10 kD) à 4 mL (+1 mL de rinçage) et chargée sur une colonne de tamisage moléculaire AcA54 (préalable équilibrée dans du tampon SP-B) à 0,5 mL/min sur la nuit. Les fractions collectées sont analysées par SDS-PAGE et celles contenant la protéine (**en rouge Figure 23**) sont rassemblées (« AcA »).

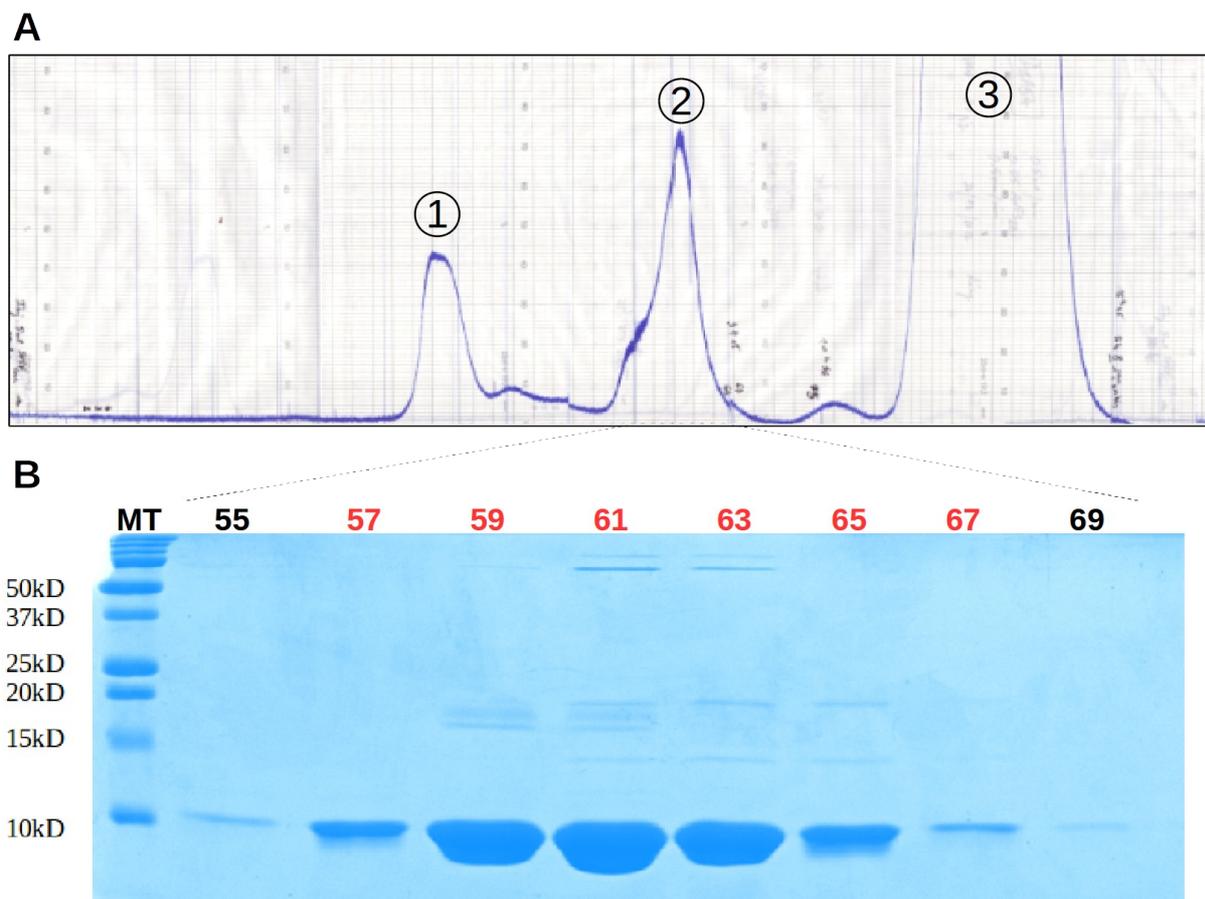


Figure 23: Analyse de l'élué de la colonne AcA54 A- Chromatogramme où le pic 1 correspond au volume d'exclusion comportant des protéines de hauts poids moléculaires. Le pic 2 correspond à la région où est éluée HU. Le pic 3 correspond au lysosyme. C- Analyse en gel SDS-PAGE 16% avec « MT », le marqueurs de taille, « EB », l'extrait brut, « Dia », le dialysat et « FF » le pool récolté après la SP-FF. Les numéros sous les pistes correspondent aux numéros des fractions. Ceux en rouges ont été rassemblés pour les prochaines étapes de purifications.

La fraction AcA est diluée avec du tampon SP-A pour atteindre une concentration en NaCl finale de 100 mM et est chargée sur une colonne de chromatographie échangeuse d'ion SP-HP préalablement

équilibrée dans du tampon SP-A +10%B. Après 10 CV de lavage, l'éluion est réalisée par un gradient linéaire de 10 à 50 % SP-B (500mM NaCl) sur 500mL. Les fractions collectées au cours de l'éluion sont analysées par SDS-PAGE et les fractions contenant HU sont rassemblées (« HP »).

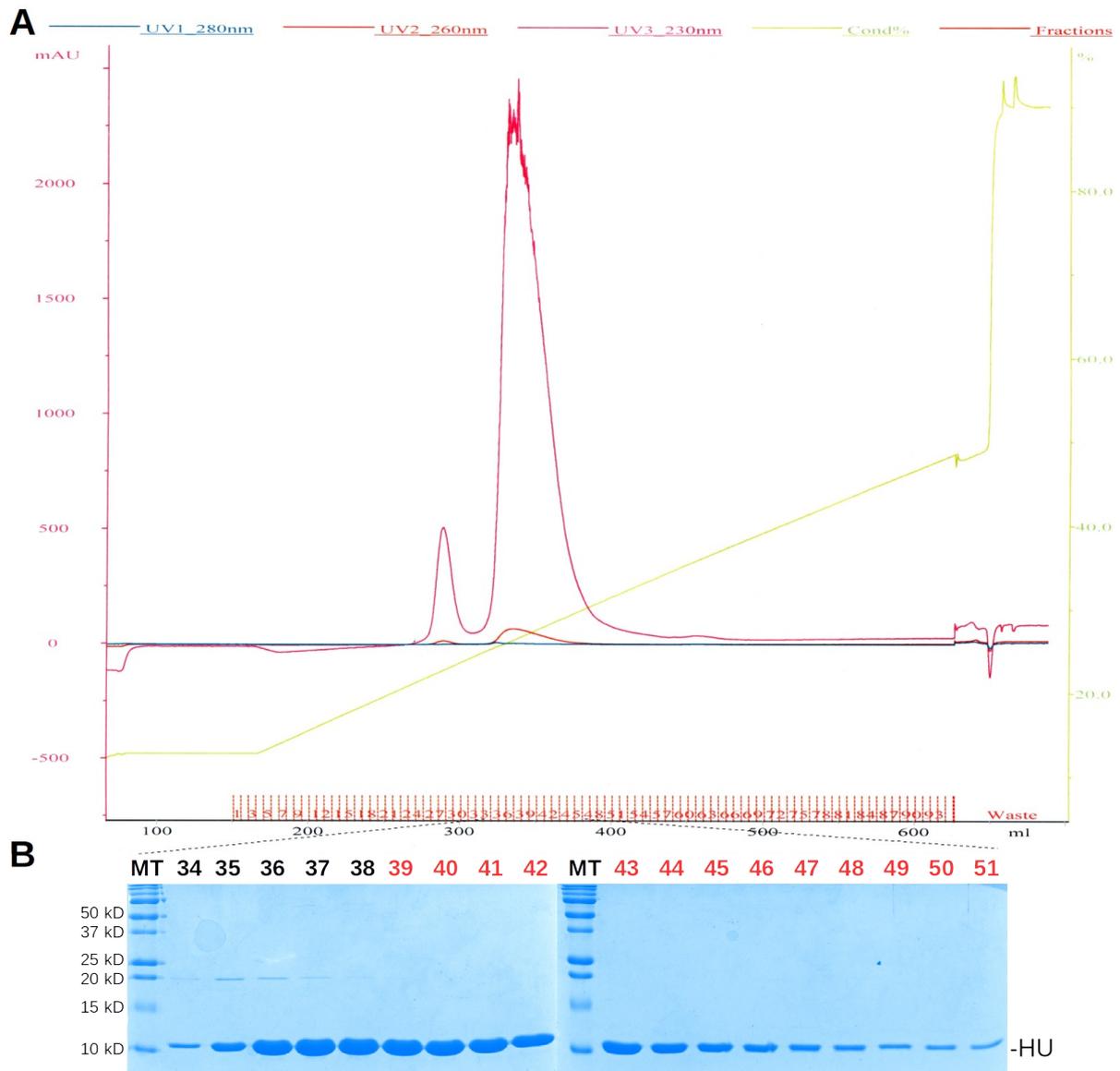


Figure 24: Analyse de l'éluion de la colonne SPHP A- Chromatogramme avec l'absorbance à 230, 260 et 280nm en magenta, rouge et bleu, respectivement. La conductimétrie est indiquée en vert. B- Analyse SDS-PAGE des fraction d'éluion contenant la protéine HU. Les fractions en rouge, pour lesquelles la protéine est pure, sont rassemblées.

En fonction de l'état de pureté à l'issue de cette étape, une étape de chromatographie échangeuse de cations faible (20CM) peut être réalisée. La fraction HP est échangée par Centricon contre le tampon CM-A (25 mM Tris-HCl pH7,5, 1 mM EDTA, 0,1 mM PMSF, 2,5 mM β-mercaptoéthanol) exempt de glycérol. Puis on charge sur une colonne 20CM préalablement équilibrée avec un tampon CM-A +5 % CM-B (CM-B =CM-A + 1M NaCl). On lave par 10 CV avec du tampon CM-A +5% CM-B. On élue par un gradient linéaire de 5 à 75 % CM-B sur 80 CV.

Un gel SDS-PAGE présentant un échantillon à chaque étape de la purification est présenté **Figure 25**.

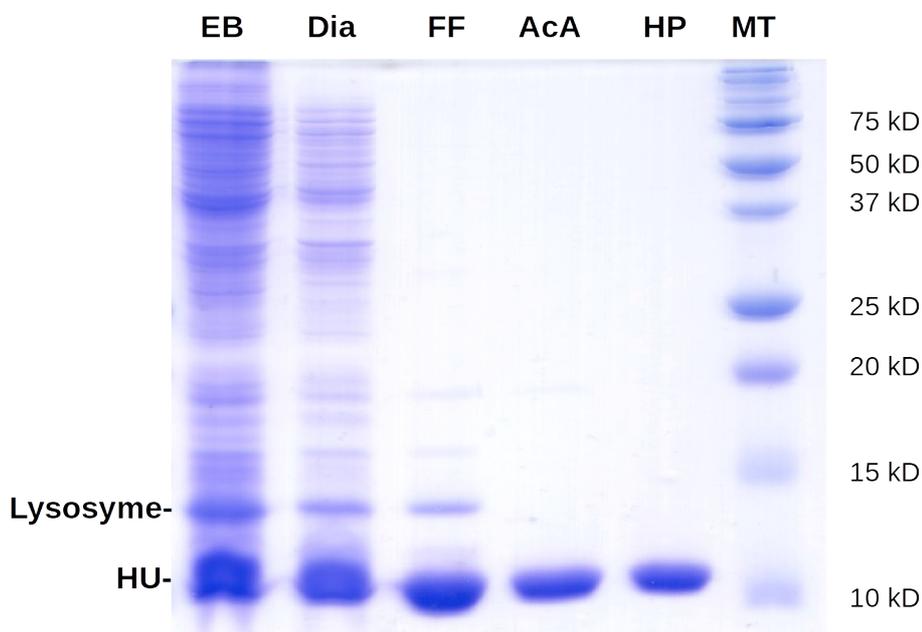


Figure 25: Gel SDS-PAGE résumé de la purification HU de type sauvage « MT » : Marqueur de taille ; « EB » : Extrait Brut ; « Dia » : dialysat ; « FF »:SP-FF ; « Aca » : AcA54 ; « HP » : SP-HP.

Ce protocole de purification est celui qui présente le meilleur rendement pour la purification des protéines sauvage avec entre **30 et 60mg de protéine purifiée pour 10g de culot bactérien**.

2.3.2. Purification des protéines fusionnées avec une Étiquette 8-HIS

Remarque : des chromatogrammes et gels SDS-PAGE correspondant à cette purification sont présentés dans le chapitre 3 (partie 3.2.5 p 88 à 94).

Les culots bactérien sont repris à raison de 10 mL par gramme de cellule dans le tampon de lyse (25 mM Tris/HCl pH 7,5 ; 800mM NaCl ; 1 mM EDTA ; 0,1 mM PMSF). La suspension est supplémentée par du lysosyme (1 mg/mL final) et incubée 15min à 30°C sous agitation. Puis, On réalise 6 à 10 cycles de sonication (1min : 5"pulse : 5"repos ; 40 % puissance d'un sonificateur 750 Watt équipé d'une sonde 13 mm). La suspension doit passer par une phase visqueuse (libération de l'ADN génomique) puis redevenir fluide lorsque l'ADN est suffisamment fractionné. Le lysat est ensuite centrifugé 1 h à 4°C et à 8000 rpm (rotor SS-34). Le surnageant est filtré à 0,2µm et constitue l'extrait brut (« EB »)

L'extrait brut est chargé sur une colonne Ni-NTA préalablement équilibrée dans un tampon His-A +4 % His-B (Hi-sA : 25 mM NaP pH 7,8, 2 M NaCl, 0,1 mM PMSF ; His-B : His-A + 0,5 M Imidazole) soit 20mM Imidazole. La colonne est ensuite lavée par 10 CV du tampon d'équilibration puis on élue par un gradient linéaire de 4 % à 100 % His-B. Les fractions d'élution sont analysées par SDS-PAGE et celles contenant la protéine fusion sans HU sauvage sont rassemblées (« AF1 »).

Le tampon de la fraction AF1 est échangé par Centricon Amicon (seuil10kD) avec le tampon entérokinase (20 mM NaP pH7,8 ; 150 mM NaCl ; 2mM CaCl₂). La concentration en protéine est ajustée à 1 mg/mL, puis on supplémente par de l'enterokinase à un ratio de entérokinase/HU de 10⁻⁵ (m/m). Le clivage est réalisé à 16°C durant 12h et est arrêté par ajout de 0,1mM final PMSF (« EK »)

Le protéolysat (EK) est chargé sur une colonne Ni-Nta préalablement équilibrée dans un tampon His-C (25 mM NaP pH7,8, 150 mM NaCl, 0,1 mM PMSF, 2,5 mM β-mercaptoéthanol). Les protéines non retenues (« Flow-through »), dont la protéine HU, sont collectées (« AF2 »).

La fraction AF2 est ajustée par dilution à 50mM NaCl final et chargée sur une colonne HS-20 préalablement équilibré dans du tampon HS-A +10 % HS-B (HS-A : 25 mM Tris-HCl pH7,5, 1 mM EDTA, 0,1 mM PMSF, 2,5 mM β-mercaptoéthanol ; HS-B : HS-A + 1M NaCl). Après un lavage de 10CV avec le tampon d'équilibration, on élue par un gradient linéaire de 10 à 100 % HS-

B sur 80CV. Les fractions collectées sont analysées sur SDS-PAGE et les fraction contenant protéine HU pure sont rassemblée (« HS20 »).

Ce protocole à montré un rendement très faible, du fait du faible niveau de d'expression de cette protéine de fusion, avec un rendement de **1mg de protéine pour 10g de culot bactérien**.

2.3.3. Purification des protéines fusionnées avec la MBP

Remarque : des chromatogrammes et gels SDS-PAGE correspondant à cette purification sont présentés dans le chapitre 3 (partie 3.3.4 p 102 à 108).

De manières analogues au protocole de purification de la fusion 8-His, les culots bactérien sont repris à raison de 10 mL par gramme de cellule dans le tampon de lyse (25 mM Tris/HCl pH 7,5 ; 800mM NaCl ; 1 mM EDTA ; 0,1 mM PMSF). La suspension est supplémentée par du lysosyme (1 mg/mL final) et incubée 15min à 30°C sous agitation. On réalise ensuite 6 à 10 cycles de sonication (1min : 5"pulse : 5"repos ; 40 % puissance sur sonificateur 750 Watt équipé d'une sonde 13 mm). La suspension doit passer par une phase visqueuse (libération de l'ADN génomique) puis redevenir fluide lorsque l'ADN est suffisamment fractionné. Le lysat est ensuite centrifugé 1 h à 4°C et à 8000 rpm (rotor SS-34). Le surnageant est filtré à 0,2µm et constitue l'extrait brut (« EB »)

L'extrait brut est ensuite chargé sur une colonne amylose préalablement équilibrée dans le tampon Amy-A (25 mM Tris/Hcl pH7,5 ; 800mM NaCl). La colonne est lavée avec 10 CV du tampon Amy-A. On élue ensuite la protéine fusion par le tampon Amy-B (25 mM Tris/Hcl pH7,5 ; 800 mM NaCl ; 100mM Maltose) (les 10mM ne sont pas suffisant, probablement à cause de la dimérisation *via* HU). L'éluat est collecté et constitue la fraction « Amy ».

Le tampon de la fraction Amy est échangé par centricon™ Amicon™ (30 kDa) contre le tampon compatible avec l'enterokinase (25 mM Tris/HCl pH 7,5 ; 150 mM NaCl ; 2 mM CaCl₂). La concentration en fusion ne doit pas excéder 10 mg/mL (~11 UA à 280 nm) afin d'éviter des problèmes de précipitation. On se placera à cette concentration maximale pour le clivage protéolytique. On ajoute l'entérokinase à un rapport 10⁻⁷ protéase sur protéine (m/m) et on incubé 15h à 16°C. Le clivage est stoppé par ajout de PMSF (0,1 mM final) et passage sur une colonne benzamidine (fixe spécifiquement les protéases à sérine). Le « Flow-through » de la benzamidine constitue la fraction « EK ».

Après ajustement à 50mM NaCl par dilution dans du tampon SP-A (25 mM Tris/HCl pH 7,5 ; 0,1mM PMSF ; 1mM EDTA), la fraction EK est chargée sur une colonne SPFF préalablement équilibrée dans un tampon SP-A +5 % SP-B (SP-B : SP-A +1M NaCl). La colonne est ensuite lavée par 10 CV du tampon d'équilibration. La fusion non clivée est éluée au cours du lavage et peut être réservée pour une purification ultérieure. On élue ensuite par un gradient linéaire de 10 à 60 % SP-B sur 10 CV. Les fractions d'élution sont analysées par SDS-PAGE et les fractions contenant HU sont rassemblées et constituent la fraction « FF ».

La fraction FF est concentrée sur centricon™ Amicon™ (10 kD) jusqu'à un volume de 1 mL et est chargée sur une colonne S200 préalablement équilibrée dans du tampon SP-B. Les fractions d'élution sont analysées par SDS-PAGE et les fractions contenant HU sont rassemblées et constituent la fraction « S200 ».

La concentration en NaCl de la fraction S200 est ajustée à 50 mM par dilution et est ensuite chargée sur une colonne 20CM équilibrée dans un tampon SP-A+5%B. Après lavage par 10 CV de tampon d'équilibration, on élue par un gradient linéaire de 5 % à 75 % SP-B sur 80 CV. Les fractions d'élution sont analysées par SDS-PAGE et celles contenant la protéine HU pure sont rassemblées (« 20CM »).

2.4. Cristallographie des rayons X

2.4.1. Cristallogénèse

Tous les cristaux ont été obtenus par diffusion en phase vapeur, soit en goutte suspendue soit en goutte assise (**Figure 27**). Les échantillons de protéine sont placés dans un tampon minimal (10mM HEPES pH 7,5 ; 200mM NaCl). La concentration relativement faible de tampon permet d'une part, de stabiliser le pH de la solution mère de protéine et d'autre part d'autoriser des variations lors de mélanges avec les solutions précipitantes des kits de cristallisation (avec des tampons 10 à 20 fois plus concentrés). Le NaCl permet de s'assurer que les protéines HU, fortement polaires, restent stables avant le processus de cristallogénèse.

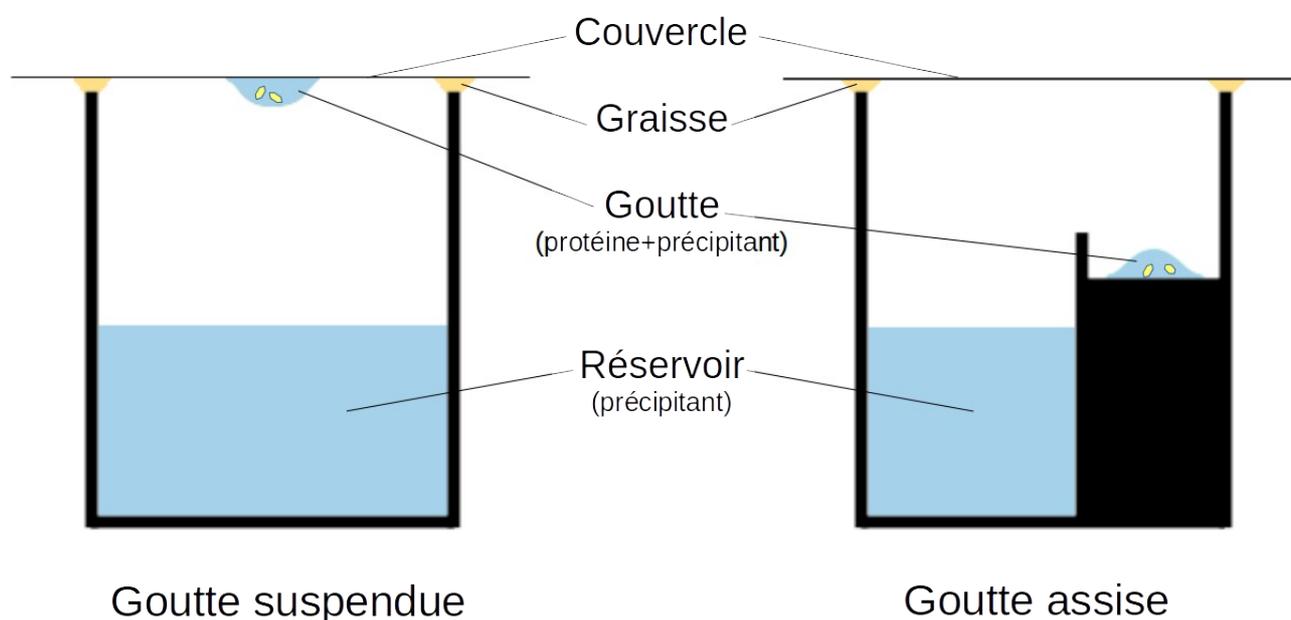


Figure 27: Techniques de cristallogénèse par diffusion en phase vapeur

Pour la protéine HU de *L. lactis*, un échantillon de protéine fraîchement purifiée et concentrée à 20mg/mL a été criblé en utilisant les 50 solutions du kit basic de cristallography de Sigma-aldrich (composition des solutions dans le tableau 23 en annexe).

2.4.2. Collecte des données de diffraction des rayons X

Les cristaux de Ec-HU β_2 et Ll-HU ont été plongés dans des solutions cryoprotectantes. Du paratone N (Hampton Research) a été utilisé dans le cas de Ec-HU β_2 . La solution mère supplémentée par du glycérol (15 % final) a été utilisée dans le cas de Ll-HU.

Les expériences de diffraction des rayons X sur les cristaux congelés ont été menées à l'European Synchrotron Radiation Facility (ESRF Grenoble) sur la ligne de lumière ID23-1.

Les images de diffraction ont été collectées à 100K jusqu'à une résolution de 1,25Å pour HU β_2 et 2,19Å pour LHU.

Les réflexions de chaque image des jeux de données ont été indexées par le logiciel XDS¹⁶⁹ et mises à l'échelle par le logiciel SCALA de la suite logiciel CCP4¹⁷⁰.

2.4.3. Remplacement moléculaire et affinement

Les structures ont été résolues par remplacement moléculaire avec le logiciel PHASER¹⁷¹ en utilisant la structure cristallographique de HU α_2 de *E.coli*¹⁶⁶ (code PDB : 1MUL) qui possède 69 % et 54 % d'identité de séquence pour HU β_2 de *E.coli* et HU de *L.lactis*, respectivement. L'affinement des modèles a été réalisé en utilisant la suite logiciel PHENIX¹⁷² le logiciel COOT¹⁷³. La qualité des modèles structuraux a été évaluée par le logiciel PROCHECK¹⁷⁴.

2.5. Résonance Magnétique Nucléaire

Les expériences RMN ont été enregistrées d'abord sur un spectromètre VARIAN INOVA 600MHz avec une sonde chaude puis sur des spectromètres BRUKER ASCEND 700MHz ou BRUKER ASCEND 950MHz (TGI RMN de Gif-sur-Yvette) équipés de sondes cryogéniques.

Les différents dimères de HU sont analysés dans le tampon de RMN (20mM NaP, 200mM NaCl, 0,05mM DTT, 5%-10 % D₂O) soit dans des tubes SHIGEMI 5mm (Echantillons analysés sur le 600MHz) soit dans des tubes 3mm standard (700 et 900MHz équipés de sondes cryogéniques)

2.5.1. Attribution des résonances

Les résonances des atomes du squelette peptidique des dimères de HU ont été obtenus par attribution séquentielle avec des expériences triples-résonances¹⁷⁵ (HNCO, HNCA, HNCOA, HNCACB, CBCACONH, HBHACONH, HBHANH) en utilisant les versions BEST^{176,177} lorsqu'elles étaient disponibles. Les résonances des atomes provenant des chaînes latérales ont été attribuées en utilisant des expériences HCcCONH, CcCONH et HCCH-TOCSY.

La majorité des expériences ont été enregistrées en utilisant un échantillonnage non uniforme (« non

uniform sampling » NUS) pour optimiser leur durée d'acquisition et la résolution des spectres.

Les spectres ont été transformés soit directement dans TopSpin™ (Bruker) soit en utilisant les logiciels NMRPIPE¹⁷⁸ et mddnmr^{179,180} (en « compressed sensing »). Ils ont ensuite été attribués en utilisant le logiciel CCPNMR¹⁸¹.

2.5.2. Calculs des structures

Des expériences ¹³C-HSQC-NOESY, ¹⁵N-HSQC-NOESY et/ou H.H-NOESY (temps de mélange 100 ms) ont été attribuées manuellement (CCPNMR¹⁸¹) puis semi-automatiquement en utilisant le logiciel ARIA2¹⁸² basé sur le programme CNS¹⁸³. Les contraintes des NOEs et des angles dièdres prédis à partir des déplacements chimiques (en utilisant le logiciel DANGLE¹⁸⁴) ont été utilisés pour le calcul itératif de structures dans ARIA2.

2.5.3. Relaxation

Les expériences de relaxation T1 ont été enregistrées en faisant le délai de relaxation : 10, 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1250, 1500, 2000 ms. Celles de relaxation T2 ont été enregistrées en faisant varier le délai de relaxation T2 : 10, 30, 50 70, 90, 110, 130 150, 170, 190, 210, 230, 250, 310, 350, 410 et 490 ms. Un délais de recyclage D1 de 3s est utilisé pour les deux types d'expériences.

Les erreurs sur R1 et R2 sont estimées par une répétition des ajustements 1000 fois sur une distribution aléatoire dans la gamme de $I \pm \Delta I$ en utilisant le logiciel Sparky (<https://www.cgl.ucsf.edu/home/sparky/>). Les spectres ¹⁵N NOE ont été enregistrés en utilisant une période de présaturation de 3s et un délai de relaxation de 3s également. Les spectres de référence sont enregistrés avec un délai équivalent de 6s sans présaturation. Les NOE ¹H-¹⁵N sont extraits du rapport des intensités des pics provenant des expériences avec et en absence de présaturation (moyennés sur 2 séries d'expériences). L'erreur sur les NOE est calculée à partir des rapports signaux sur bruits.

2.5.4. ZZ-¹⁵N-HSQC : extraction des constantes de vitesse d'échange

Les expériences ZZ-¹⁵NHSQC ont été enregistrées en faisant varier le délai T (50, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000 ms). L'intensité des pics autos (NN et II) et croisés (NI et IN) suivent les équations suivantes :¹⁸⁵

$$I_{NN}(T) = I_{NN}(0) \times A_N \times \frac{(-(\lambda_2 - a_{11})e^{-\lambda_1 T} + (\lambda_1 - a_{11})e^{-\lambda_2 T})}{(\lambda_1 - \lambda_2)} \quad (1a)$$

$$I_{II}(T) = I_{II}(0) \times A_I \times \frac{(-(\lambda_2 - a_{22})e^{-\lambda_1 T} + (\lambda_1 - a_{22})e^{-\lambda_2 T})}{(\lambda_1 - \lambda_2)} \quad (1b)$$

$$I_{NI}(T) = I_{NN}(0) \times A_I \times \frac{(a_{21}e^{-\lambda_1 T} - a_{21}e^{-\lambda_2 T})}{(\lambda_1 - \lambda_2)} \quad (1c)$$

$$I_{IN}(T) = I_{II}(0) \times A_N \times \frac{(a_{12}e^{-\lambda_1 T} - a_{12}e^{-\lambda_2 T})}{(\lambda_1 - \lambda_2)} \quad (1d)$$

Avec :

$$\lambda_{1,2} = \frac{1}{2}((a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + 4k_{NI}k_{IN}})$$

$$a_{11} = R_{1N} + k_{NI}$$

$$a_{12} = -k_{IN}$$

$$a_{22} = R_{1D} + k_{IN}$$

$$a_{21} = -k_{NI}$$

$I_{NN}(0)$ et $I_{II}(0)$ représentent l'aimantation des états N_2 et I_2 (respectivement) au départ du temps de mélange T. A_N et A_I représentent l'efficacité de transfert de cohérence, $^{15}N \rightarrow ^1H_N$, pendant l'étape INEPT retour, pour l'aimantation associée aux état N_2 et I_2 , respectivement.

Le calcul des facteur A_N et A_I pour chaque résidu est réalisé comme publié par Tollinger et collaborateurs.¹⁸⁶ Pour le calcul des constantes de vitesse d'échange de $HU\alpha_2$ ($k_1 \leftrightarrow k_{NI}$ et $k_{-1} \leftrightarrow k_{IN}$), les volumes des pics du résidu K90 en fonction du délai T ont été ajustés à l'aide du logiciel Qtiplot (<http://www.qtiplot.com/>) avec les équations 1a-d.

2.5.5. Perturbations de déplacements chimiques (CSP)

Les perturbations de déplacements chimiques entre les conformations N_2 et I_2 ont été calculées en utilisant l'équation suivante :

$$CSP = \sqrt{(\delta H_I - \delta H_N)^2 + \left(\frac{\delta N_I - \delta N_N}{6}\right)^2}$$

Où δH et δN correspondent aux déplacements chimiques proton et azote, respectivement. L'indice I et N correspondent aux états I_2 et N_2 , respectivement. Le facteur 6 permet de corriger la différence de largeur spectrale entre les deux atomes (de l'ordre de 5 ppm pour H et 35 ppm pou N).

2.6. Bio-informatique et logiciels

2.6.1. Visualisation des structures et génération des figures

Toutes les figures montrant un modèle atomique ont été réalisées en utilisant les logiciels Pymol¹⁸⁷ ou UCSF Chimera¹⁸⁸, en utilisant le moteur de rendu POV-RAY. Les fichiers de structures ont été téléchargés depuis la PDB (Protein Data Bank; <http://www.rcsb.org/pdb/home/home.do>).

2.6.2. Graphiques et ajustement des courbes expérimentales

Les graphiques et les ajustements des courbes expérimentales présentés dans cette thèse ont été réalisés en utilisant les logiciels QtiPlot (<http://www.qtiplot.com/>) ou kaleidagraph (http://www.synergy.com/wordpress_650164087/).

2.6.3. Criblage, préparation et analyse des structures

Pour identifier l'ensemble des structures de HU et de ses homologues structuraux, le logiciel DALI a été utilisé en utilisant la structure de HU α_2 de *E. coli* comme structure de référence en fixant un seuil d'homologie à un z-score supérieur ou égal à 6. Pour comparer efficacement les structures résolues par différentes méthodes, les structures cristallographiques ont été protonées et minimisées à l'aide du logiciel SybylTM(Tripos).

Le logiciel PISA a été utilisé pour analyser l'interaction des chaînes de HU et de ses homologues pour former des dimères. Cet outil permet d'évaluer la probabilité qu'un oligomère décrit dans une structure cristallographique soit physiologique ou artificiellement maintenu par l'empilement cristallin. Pour cela, il analyse automatiquement les surfaces d'interactions entre les monomères, la présence de pont salins et de liaisons hydrogène. Nous avons utilisé ce logiciel pour comparer les propriétés d'interaction des chaînes protéiques au sein des dimères de HU et de ses homologues.

Les poches hydrophobes des dimères de HU ont été analysés en utilisant le logiciel CastP¹⁸⁹ (<http://sts-fw.bioengr.uic.edu/castp/calculation.php>) en utilisant une sonde de 1,2 Å de rayon sur les structures protonées et minimisées.

2.6.4. Alignement de séquences primaires

Les séquences protéiques obtenues sur le site de la PDB (<http://www.rcsb.org/pdb/home/home.do>)

ou Uniprot (<http://www.uniprot.org/>) ont été alignées en utilisant le logiciel Clustal Omega.¹⁹⁰ Les alignements obtenus sont ensuite visualisés et analysés en utilisant Jalview¹⁹¹.

2.7. Échange de chaînes entre les dimères

Les expériences d'échange de chaînes ont été réalisées par mélange volume à volume d'homodimères purs (10 μ M HU, 350 mM Acétate d'ammonium pH 5) préincubés à la température désirée (20°C, 27°C ou 37°C). Le mélange réactionnel est également thermostaté et des échantillons sont prélevés au cours du temps et congelés immédiatement dans l'azote liquide puis stockée à -20°C avant analyse.

2.8. Biochimie

2.8.1. Pontage chimique par le glutaraldehyde

Les protéines HU (1 mg/mL, soit \sim 50 μ M) ont été pontées dans leur tampon de stockage (20mM NaP pH7,5 ; 200mM NaCl ; 0,05mM DTT) avec du glutaraldéhyde (0,01% final m/v, sauf lorsque indiqué différemment). On incube à 37°C (sauf si indiqué différemment) pendant 30 min. Puis, la réaction est arrêtée par ajout de tampon Tris/HCl pH 7,5 (1 M final). Les échantillons sont conservés à -20°C avant analyse en SDS-PAGE ou NAT-PAGE.

2.8.2. Oxydation et purification de HU α_2 F79C en conditions natives

Le tampon de la protéine HU α_2 F79C (pure et réduite) est échangé par chromatographie d'exclusion sur une colonne Biospin 6 (exclusion à 6 kDa, Biorad) équilibrée dans le tampon d'oxydation (200mM Acetate d'ammonium pH 5 dans la condition retenue ; 20mM NaP pH8, 200mM NaCl ou 200mM Bicarbonate d'amonium pH8 dans les tests de mise au point). La concentration de la protéine est ajustée à 1 mg/mL (\sim 50 μ M) puis on ajoute 1,1 équivalent de CuCl₂. Après d'incubation (66 h à 4°C ou 6h à 37°C dans les conditions retenues) la solution protéique est filtrée à 0,2 μ m puis purifiée sur chromatographie échangeuse de cation faible (20CM, cf p 214).

2.8.3. Electrophorèses en gel de polyacrylamide et d'agarose

Les gels **SDS-PAGE** sont composés d'un gel de concentration (4,8 % m/v acrylamide :bis (29:1), 125 mM Tris/HCl pH 6,8, 0,1 % SDS) et d'un gel de séparation (12 % ou 16 % acrylamide :bis (29:1), 375 mM Tris/HCl pH 8,8, 0,1% SDS). Les échantillons sont supplémentés par un tampon de

charge (125mM Tris/HCl pH6,8 ; 20 % glycérol ; 0,2 % bleu de bromophénol ; 2 % SDS ; ±280mM β-mercaptoethanol final) avant d'être chargés. La migration est réalisée dans le tampon de migration (25mM Tris ; 50mM glycine ; 1% SDS) à 180V pendant environ 50 min, puis les gels sont incubés 20 min dans une solution de coloration (10 % acide acétique, 50% éthanol et 0,25 % bleu de coomassie R250) sous agitation. Les gels sont ensuite décolorés dans une solution de décoloration (10 % éthanol et 7 % acide acétique) sous agitation avant d'être photographiés.

Les gels **NAT-PAGE** ont été réalisés en adaptant le protocole décrit par Pellegrini et collaborateurs.¹⁹² Les échantillons sont supplémentés par un tampon de charge (750mM Acide acétique pH5,5, 20 % Glycérol et 0,01 % pyronine Y) puis chargé sur un gel de polyacrylamide (15% acrylamide:bis (29:1)). La migration se fait de l'anode vers la cathode, avec une solution 113 mM acide ε-6-amino-n-caproïc 5mM Acide acétique pH 5,8 du côté de l'anode et une solution de 30mM d'acide acétique pH 3,6 du côté de la cathode, durant 3 h à 30 mA constant.

Les **gels d'agarose** ont été réalisés à 0,8 ou 2 % m/v. Les échantillons sont supplémentés par un tampon de charge (10% glycérol, ±0,025 % bleu de bromophénol, ±0,025 % xylène cyanol) avant d'être chargés sur le gel. La migration est réalisée dans un tampon de migration (40 mM Tris/Acétate, 2 mM EDTA) durant 20 à 30min à 100V. Les gels sont incubés dans une solution de bromure d'éthidium (BET) sous agitation pendant 20 min. L'excédent de BET à la surface du gel est rincé à l'eau avec précaution.

Les photographies des gels SDS-PAGE, NAT-PAGE et d'agaroses (UV) ont été obtenus par un système Geldoc™ XR+ (Biorad).

2.9. Simulation de dynamique moléculaire

Les simulations de dynamique moléculaire ont été réalisées avec un cluster Beowulf linux multiprocesseur, à l'aide de Gromacs¹⁹³ (version 4.5.4) en utilisant le champs de force MARTINI^{194,195}. Les topologies et structures « gros grains » de HUα₂ et HUβ₂ ont été obtenues à partir des modèles atomistiques en utilisant le script « martinize » développé par Djurre de Jong, Jaakko J. Uusitalo et Tsjerk A. Wassenaar. Les structures « gros grains » ont été solvatées dans une boîte d'eau supplémentée par du NaCl (100 mM avec un défaut de sodium pour neutraliser le système), puis minimisées. La température a été équilibrée par des pas de 50K durant chacun 100 ns (50K, 100K, 150K, 200K, 250K, 300K, 350K, 370K). Enfin les simulations dites « productives »

ont été menées sur des durée de 10 μ s. L'analyse des trajectoires a été réalisée à l'aide de gromacs-tool et les représentations graphiques faites via le logiciel Grace (<http://plasma-gate.weizmann.ac.il/Grace/>).

2.10. Spectrométrie de masse

Les analyses de spectrométrie de masse en conditions natives (Suivi de l'échange de chaînes et détection du tétramère) ont été réalisées sur des appareils à source électrospray : ESI-ion trap model Esquire HCT ou maXis ESI-UHR-Qq-TOF (Bruker) en adaptant et optimisant les paramètres publiés par Vis et collaborateurs.¹⁹⁶ Les échantillons ont été maintenus à froid jusqu'à la source d'ionisation par un système de refroidissement à eau glacée au niveau du capillaire et par des sacs de glace au niveau de la seringue Hamilton permettant l'injection de l'échantillon. Ceci afin de s'assurer que l'échange ne se poursuive pas au cours de la manipulation des échantillons précédant l'analyse.

Les analyses de spectrométrie de masse en conditions dénaturantes (Suivi de l'oxydation du pont disulfure de HU α_2 F79C et analyse de l'homogénéité des protéines purifiées) ont été réalisées en mode positif sur un appareil MALDI-TOF (Bruker). Les échantillons sont déposés selon la méthode de dépôt en couche ultra-mince (UTL) publiée par Gabant G. et Cadene M..¹⁹⁷ Les spectres ont été calibrés et attribués en utilisant le logiciel Flex (Bruker). L'attribution des produits de dégradation lors des oxydations prolongées a été réalisée à l'aide du logiciel mMass.^{198,199}

Chapitre 3 : Stratégie de clonage pour la production et la purification des variants de HU

3.1. Contraintes à résoudre

Pour mettre en évidence les résidus clés nécessaires à l'échange de chaînes, nous avons prévu de réaliser de la mutagenèse dirigée sur les protéines HU d'*E. coli*. Nous avons également prévu de produire la protéine HU d'autres organismes (*L. lactis*). L'obstacle majeur pour ces productions dans *E. coli* réside dans la purification à homogénéité de HU hétérologues sans contamination par les protéines HU endogènes de type sauvage d' *E. coli*. Pour contourner ce problème, deux stratégies ont été envisagées :

(i) la première aurait consisté à produire ces protéines dans des souches $\text{hup}^{\text{A-}/\text{B-}}$, c'est à dire dépourvues des gènes codant pour les protéines HU d'*E. coli*. Ainsi, on pourrait exprimer spécifiquement les variants de la protéine en absence de HU de type sauvage. Cependant, les souches doublement mutantes $\text{hup}^{\text{A-}/\text{B-}}$ présentent un phénotype fortement altéré avec des perturbations de la croissance et de la division cellulaire.^{141,200} Dans le cas où le gène de HU hétérologue ou muté ne compenseraient pas ce phénotype, on risquerait d'avoir des difficultés pour la culture, et *a fortiori* pour la sur-expression de protéines dans cette souche.

(ii) la deuxième stratégie envisagée et retenue, est la production des protéines hétérologues fusionnées avec une étiquette d'affinité. Ce type de stratégie est largement employé et a fait l'objet d'une étude comparative.²⁰¹ Parmi les étiquettes testées par les auteurs (HIS, CBP, CYD, strep H, FLAG, HPC, GST et MBP), les étiquettes HIS (histidine) et MBP (Maltose Binding Protein) présentent des coûts réduits et d'excellents taux d'expression/purification, et ont donc été envisagés pour notre système. De plus, la surproduction efficace des protéines HU fusionnées avec une étiquette HIS a déjà été mise en évidence par Pellegrini et collaborateurs¹⁹².

La séquence ajoutée doit également comporter un site de clivage protéolytique permettant d'obtenir une protéine sans étiquette. Différentes protéases site-spécifiques sont vendues par différents fournisseurs et sont listées dans le tableau 3. Certaines sont extrêmement spécifiques comme la protéase TEV (Tobacco Etch Virus) ou la protéase HRV3C (Human rhinovirus 3C) qui possèdent des sites à 7 et 8 acides aminés, respectivement. Malheureusement, leur clivage laisserait des acides

aminés surnuméraires du côté N-terminal de HU, ce qui peut affecter sa structuration et/ou sa stabilité. Trois protéases, la thrombine, le facteur Xa et l'entérokinase, permettent un clivage sans acide aminé résiduel. Cependant, la thrombine et le facteur Xa possèdent des sites secondaires de type Gly-Arg que l'on retrouve dans le bras dynamique et hautement accessible de HU. Ces protéases sont donc à proscrire dans notre cas.

Tableau 3: Sites de clivage de différentes protéases disponibles Le symbole « || » indique le point de coupure de la protéase.

Protéase	Site de clivage principal
Thrombine*	A-B-Pro-Arg- -X-Y
Facteur Xa	Ile-(Glu ou Asp)-Gly-Arg-
Entérokinase	Asp-Asp-Asp-Asp-lys-
Protéase TEV	Glu-Asn-Leu-Tyr-Phe-Gln- -Gly
Protéase HRV3C	Leu-Glu-Val-Leu-Phe-Gln- -Gly-Pro

*site de clivage de la thrombine : A et B sont des acides aminés hydrophobes et X et Y des acides aminés non acides.

Nous avons d'opté pour un site entérokinase pour notre construction fusion. Une séquence plus efficacement clivée, Asp-Asp-Asp-Asp-Arg, a été mise en évidence par Gasparian et collaborateurs²⁰² avec une efficacité 3 à 6 fois supérieure au site standard et permet d'encore améliorer la spécificité de cette protéase. C'est la séquence que nous avons utilisée dans notre construction.

3.2. HU fusionnée à une étiquette Histidines

3.2.1. Construction des vecteurs pET28a-8HE-hupA/B

Le plasmide pET-28a+ (Novagen™), disponible au laboratoire, a été utilisé comme vecteur initial. Son multi-site de clonage est précédé par une séquence codante pour une étiquette 6-histidines (6-HIS) suivie d'un site de clivage protéolytique pour la thrombine (**Figure 27**). On souhaitait étendre l'étiquette 6 histidines à 8 histidines, pour améliorer son affinité pour le nickel, et remplacer le site de clivage à la thrombine par un site de clivage à l'entérokinase. Dans un premier temps, nous avons modifié le pET-28a+ en ajoutant 2 codons histidines et deux sites de restriction NdeI et EarI, par Quickchange™ en utilisant les oligonucléotides 1 et 2 (**Tableau 24 en annexe**). Le site NdeI est utilisé pour le clonage tandis que le site EarI nous permet de contrôler la modification du plasmide.

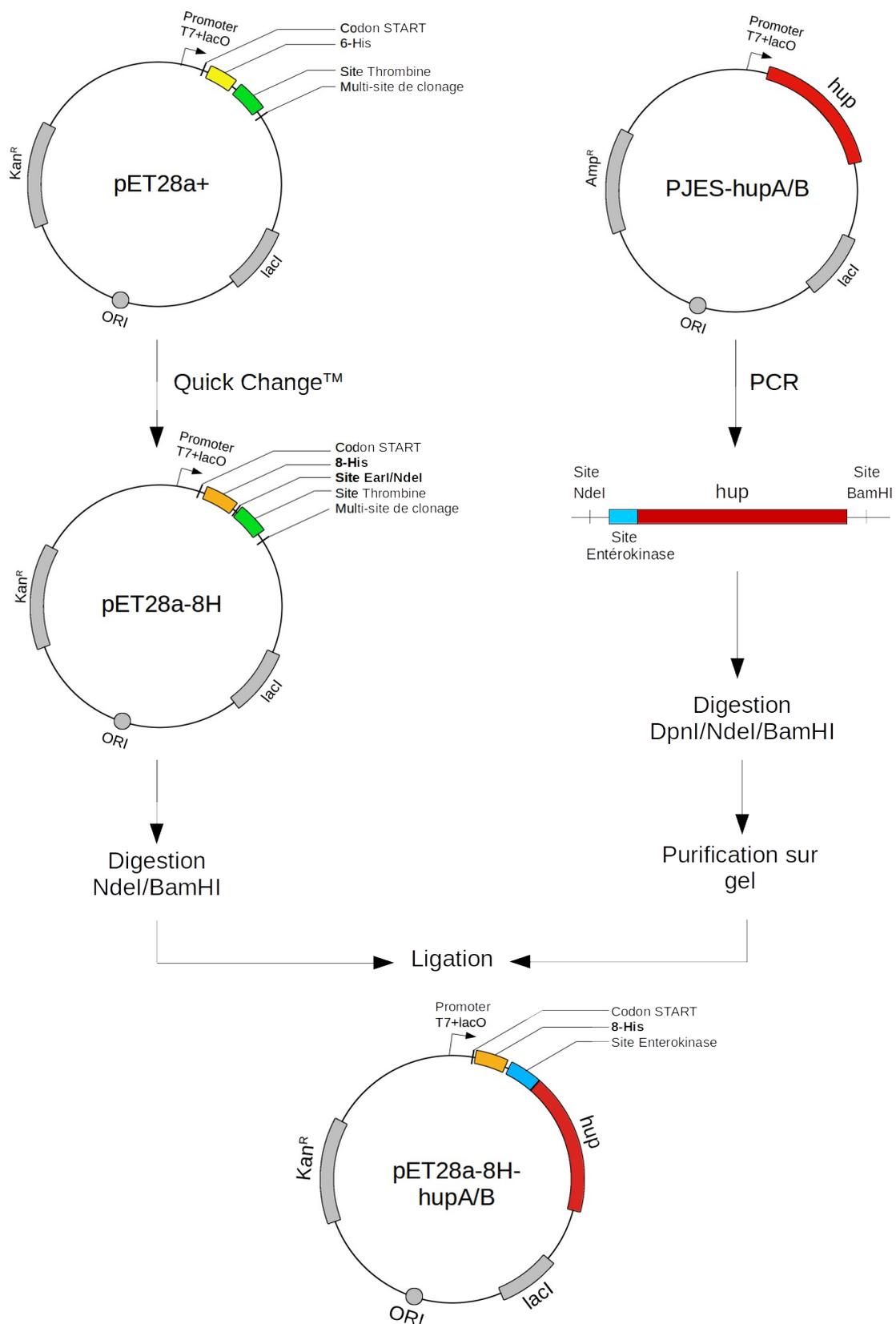


Figure 27: Stratégie de construction du vecteur pET28a-8HE-hupA/B

Il existe 3 sites *EarI* dans pET28a+ alors que le plasmide modifié pET28a-8H en possède 4 (**Figure 28A**). Après Quick Change™, on réalise des minipreps sur différents clones et on digère les plasmides obtenus par l'enzyme de restriction *EarI* (1h à 37°C). On sépare ensuite les produits de digestion par électrophorèse sur gel d'agarose (0,8%). Le gel correspondant est présenté **Figure 28B**. Les plasmides non digérés (ND) sont chargés comme témoin. Les plasmides digérés (D) présentent 2 types de profil :

(i) le premier correspond à la piste 1(D) pour laquelle on observe 4 bandes. Les 3 bandes qui ont migré le plus loin correspondent en taille aux fragments de digestion de pET28a+ non modifié. La 4^e bande correspond soit à une digestion incomplète (1871 pb+2367 pb) soit au plasmide non digéré encore circulaire et sur-enroulé. Dans le cas de ce plasmide, on suppose que la digestion *DpnI* du protocole de Quick Change™ a été incomplète et il restait donc du plasmide matrice qui a été amplifié par ce clone.

(ii) le second profil est observé dans les pistes 2(D) et 3(D). 5 bandes sont visibles sur ces pistes. Les 4 bandes qui ont migré le plus loin sont compatibles avec la carte de restriction théorique du pET28a-8H. Le Quick Change™ a fonctionné pour ces 2 plasmides. Après confirmation par séquençage, on utilise le plasmide 2 pour le clonage.

En parallèle, on amplifie les gènes *hupA* et *hupB* des plasmides pJES-*hupA* et pJES-*hupB* par PCR. Les amorces comprennent des extrémités cohésives du côté 5' permettant d'ajouter les sites de restriction *NdeI* et *BamHI* pour le clonage ainsi que la séquence codant pour le site de clivage protéolytique de l'entérokinase. Le produit de PCR théorique est présenté **Figure 27**. Comme les amorces ont des tailles significativement différentes, on a testé deux températures d'hybridation lors de la PCR. On analyse ensuite une fraction de ces produits de PCR par électrophorèse sur gel d'agarose (2%) (**Figure 28 C**). La taille apparente des produits de PCR est compatible avec la taille attendue (345 pb) et la PCR semble efficace aux deux températures d'hybridation testées. On conservera la température la plus stringente (65°C) pour la suite. Ce produit de PCR a été digéré par *DpnI* (élimination des plasmides pJES-*hupA/B*) ainsi que *NdeI* et *BamHI* (37°C 2h) pour créer les extrémités cohésives nécessaires au clonage. Il a ensuite été purifié sur gel d'agarose 2 % (basse température de fusion) à l'aide du kit GFX™(GE Healthcare™).

Le plasmide pET28a-8H (n°2) sélectionné est également digéré par *NdeI* et *BamHI*. Puis, le produit est purifié à l'aide du kit GFX™.

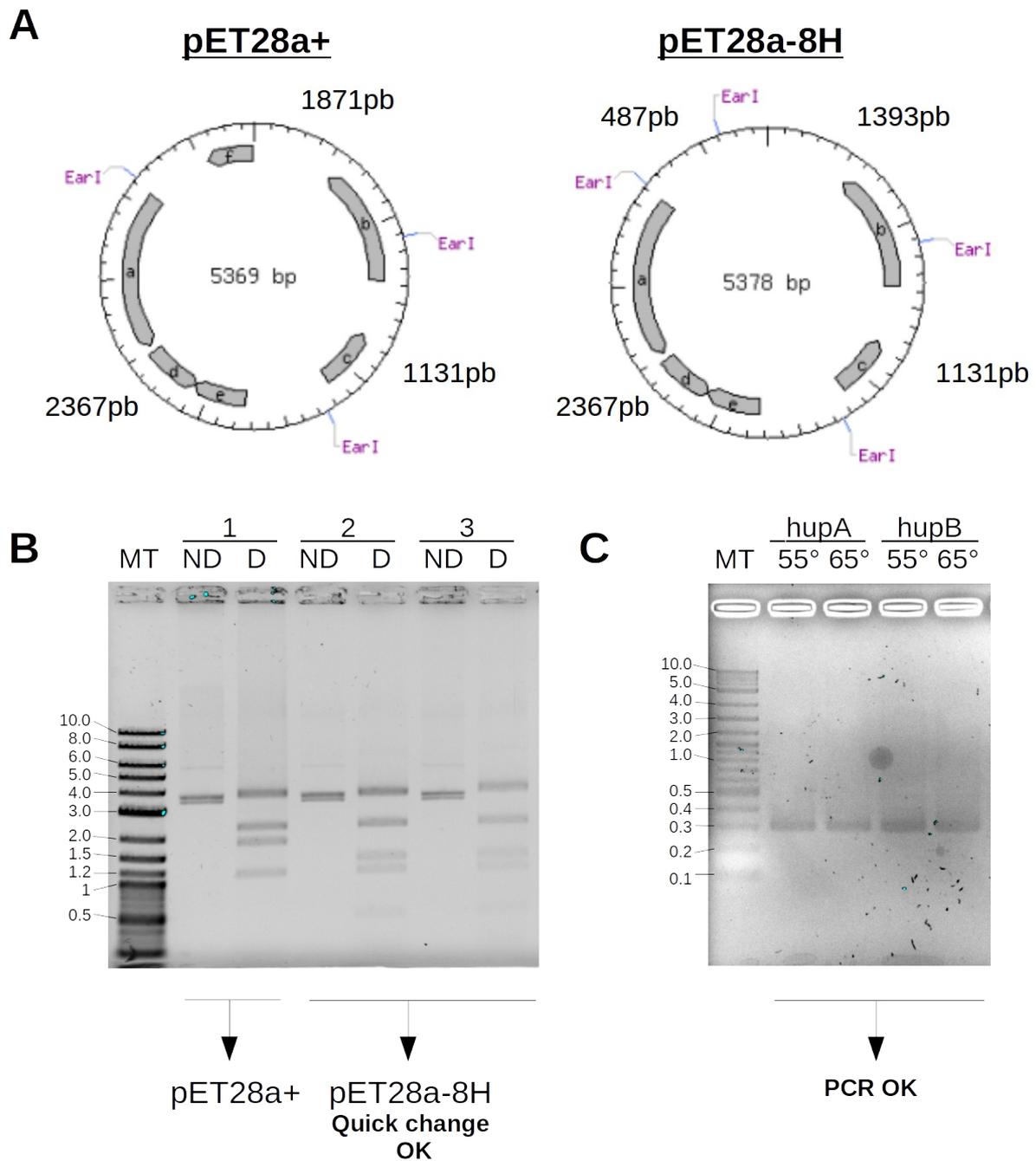


Figure 28: Contrôle des partenaires de ligation A- Carte de restriction théorique du plasmide de départ (pET28a+) et du plasmide modifié par Quick Change™ (pET28a-8H) (généré par l'outil NEBcutter™ v2.0 <http://tools.neb.com/NEBcutter2/index.php>) B- Analyse par électrophorèse sur gel d'agarose (0,8%) de 3 mini-preps de plasmide issues du Quick Change™ sur pET28a+, après digestion (D) ou non digéré (ND) par EarI. C- Analyse par électrophorèse sur gel d'agarose (2%) des produits de PCR sur pJES-hupA/B. Deux températures d'hybridation des amorces ont été testées. La condition de succès la plus stringente (65°C) a été retenue.

On réalise une ligation enzymatique en utilisant le Quick Ligation Kit™ (New England Biolabs™) avec les partenaires dans un ratio 2:1 insert:plasmide. Puis on transforme des bactéries NEB-10β™ compétentes (New England Biolabs™) avec le produit de ligation. Des minipreps sont réalisées sur les clones obtenus et la présence de l'insert dans les plasmides est vérifiée par séquençage.

Ces expériences ont été répétées de nombreuses fois avant d'obtenir des clones comportant les plasmides avec leur insert. Très peu de clones ont été obtenus en fin de clonage et une grande partie possédait le plasmide modifié seul. On peut suspecter une médiocre efficacité de la digestion enzymatique du produit de PCR. La longueur importante des amorces due à l'introduction du site de clivage entérokinase en *sus* du site de restriction, ne nous a permis d'ajouter que seulement 6 paires de bases en amont de ce site. On peut suspecter que la digestion a été peu efficace du fait de la position du site en bordure d'oligonucléotide. De plus, les étapes de purification du plasmide et de l'insert font perdre beaucoup de matériel avant l'étape ultime de ligation. De ce fait, des méthodes alternatives seront envisagées pour les prochains clonages.

3.2.2. Tests d'expressions du vecteur pET28a-8HE-hupA/B

Le protocole d'expression standard (vecteur pJES) des protéines HU de type sauvage a été utilisé pour exprimer la protéine fusion. Des échantillons sont prélevés de la culture à différents temps d'induction et on réalise des extraits totaux. Ces extraits sont analysés par SDS-PAGE (**Figure 29**). On observe que le niveau d'expression de cette protéine fusion est faible comparé à celui de la protéine de type sauvage. Nous avons d'abord suspecté un défaut de l'IPTG ou de sélection par l'antibiotique. Cependant le niveau d'expression n'a pas pu être amélioré, même en re préparant fraîchement toutes les solutions utilisées pour la culture.

Différentes pistes ont alors été explorées, sans succès, pour essayer d'expliquer cette baisse de niveau d'expression. La séquence ajoutée du côté N-terminal du gène a été mise au point en évitant les codons rares chez *E.coli* (<10 % d'utilisation, cf **Tableau 25 en annexe**). Aucune formation de structure secondaire de taille significative (type tige-boucle, qui pourrait gêner l'expression) n'a pu être mise en évidence dans cette région.

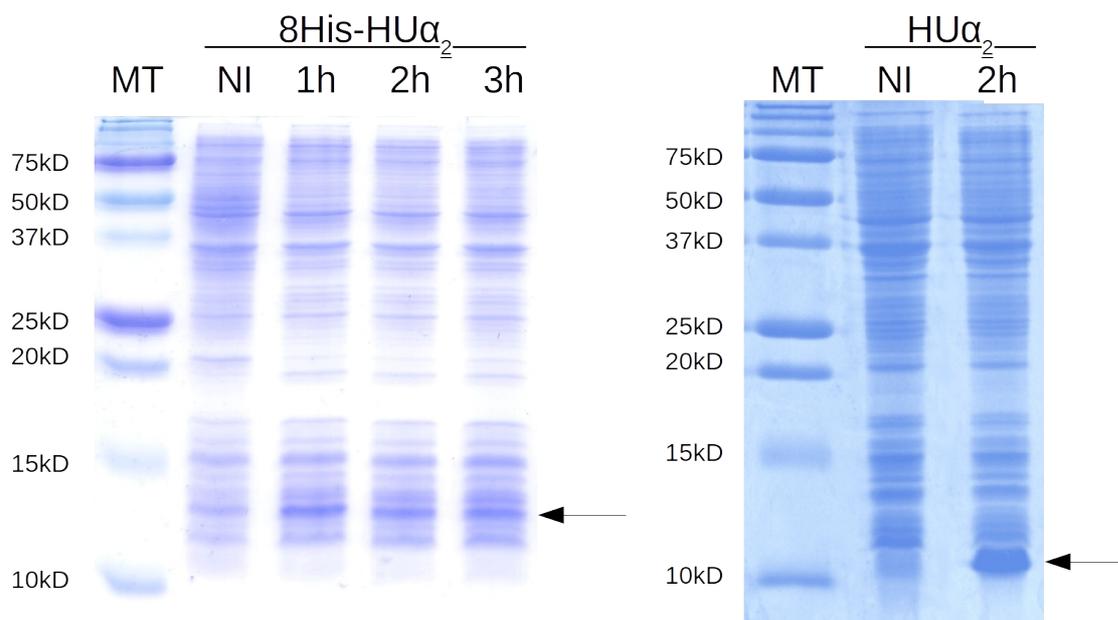


Figure 29: Expression de $HU\alpha_2$ fusionnée avec ou sans étiquette 8His SDS-PAGE sur des extraits totaux à différent temps d'induction (NI :non induit). A gauche la protéine fusion (8 His- $HU\alpha_2$ ~12,5kD) exprimée par le vecteur pET28-8HE-hupA. A droite la protéine sauvage ($HU\alpha_2$ ~9kD) exprimée par le vecteur pJES-hupA. Les bandes correspondant aux protéines sur-exprimées sont indiquées par des flèches. (MT : Marqueurs de tailles)

3.2.3. Reclonage des gènes *hupA/B* étiquetés 8His dans pJES

Nous faisons l'hypothèse qu'il pourrait y avoir une mutation dans notre vecteur pET28a+ de départ, ce qui expliquerait le faible niveau d'expression de la protéine fusion de pET28-8HE-hupA/B. Pour vérifier cette hypothèse, et tenter d'obtenir une expression du niveau de pJES-hupA (Figure 29, panneau de droite), nous avons reclone les gènes *hupA* et *hupB* fusionnés avec leur étiquette 8His dans le vecteur pJES de départ.

Un site de restriction *NcoI* est présent au démarrage de l'étiquette d'affinité dans pET28-8HE-hupA/B. Les sites *NcoI* et *BamHI* ont été utilisés pour le clonage. Le site de restriction *NcoI* (absent dans pJES) est introduit par Quick Change™ à partir des couples d'amorces 7/8 et 9/10 pour pJES-hupA et pJES-hupB, respectivement (Tableau 24 en annexe). La modification des plasmides est vérifiée par séquençage. Les vecteurs pJES-hupA/B modifiés ainsi que pET28-8HE-hupA/B sont digérés par *NcoI/BamHI* en parallèle (37°C/2h).

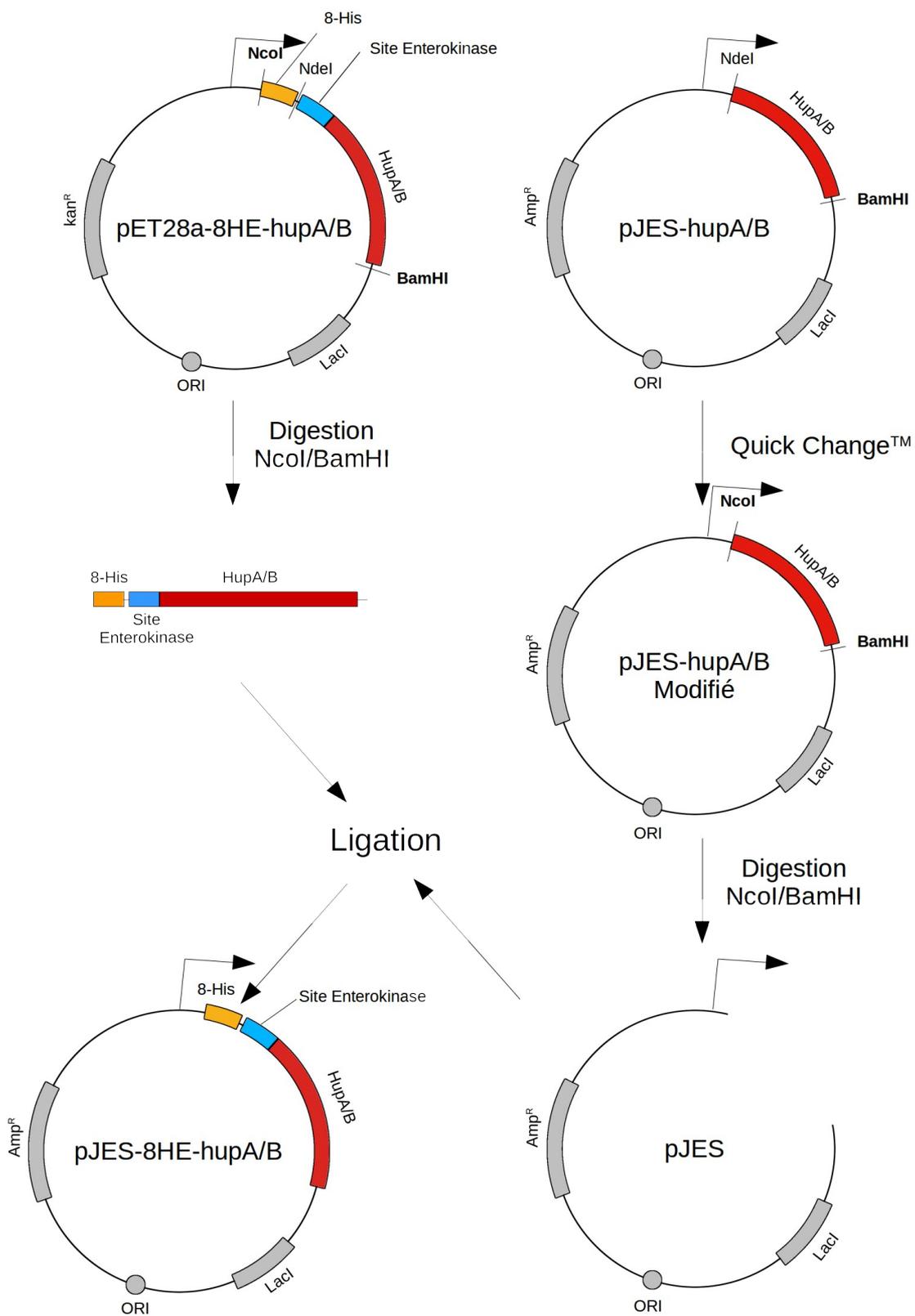


Figure 30: Procédure de reclonage des gènes *hup* étiquetés 8His dans pJES

Les produits de digestion (pJES sans *hup* et 8HE-*hupA/B* sans pET28) sont purifiés sur gel d'agarose à basse température de fusion (0,8 % pour les vecteurs et 2 % pour les inserts) à l'aide du kit GFX™ (GE Healthcare™). Les inserts et les vecteurs (en proportion 2:1) sont ligués en utilisant le Quick Ligation Kit™ (New England Biolabs™) et le produit de ligation est utilisé pour transformer des bactéries NEB-10β™ calcium-compétentes (New England Biolabs™). Des minipreps sont réalisées sur les clones obtenus et la présence de l'insert dans les plasmides est vérifiée par séquençage. Les vecteurs pJES-8HE-*hupA/B* obtenus sont ensuite soumis à des tests d'expression.

3.2.4. Tests d'expression du vecteur pJES-8HE-*hupA/B*

On utilise le protocole standard d'expression de la protéine HU de type sauvage pour exprimer les protéines fusions provenant des 2 types de vecteurs élaborés : pJES-*hupA/B* et pET28a-*hupA/B*. Les analyses des SDS-PAGE des extraits totaux prélevés à différents temps d'incubation révèlent qu'il n'y a pas d'effet de vecteur pour l'expression de la protéine fusion (**Figure 31**). Nous avons donc conservé le système pET28a pour la suite. La purification plus aisée, en moins d'étapes grâce à l'étiquette d'affinité, peut permettre de compenser la faible expression par un meilleur rendement de purification.

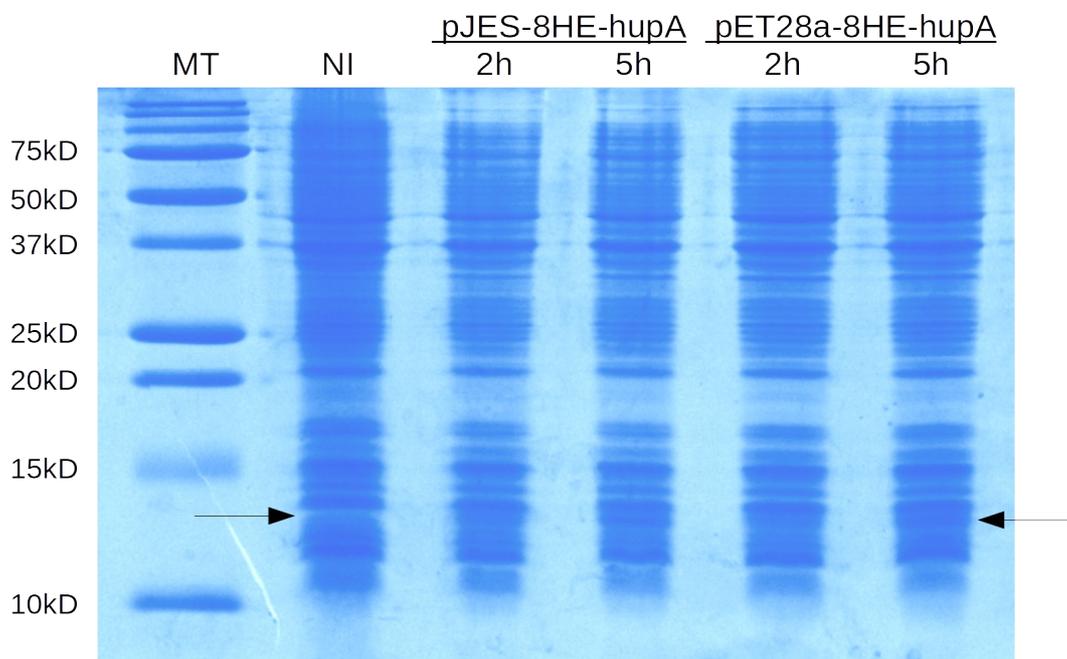


Figure 31: Test d'expressions des protéines fusions (8his) clonées dans deux vecteurs différents Les flèches indiquent la migration de la bande correspondant à la protéine fusion.

3.2.5. Purification de la protéine HU fusionnée avec une étiquette 8His

L'étiquette d'affinité présente deux atouts majeurs dans la purification de la protéine HU : (i) elle permet d'avoir un nouveau mode de séparation, la chromatographie d'affinité, et ainsi de séparer les protéines HU endogènes de la protéine exogène apportée par le plasmide, (ii) elle permet de contourner une difficulté majeure de cette purification, à savoir l'élimination de l'ADN bactérien. En effet, on ne peut pas éliminer l'ADN bactérien par chromatographie échangeuse d'ion. Un chargement à basse force ionique sur un tel support entraînerait le maintien de l'interaction ADN-HU et on co-éliminerait les deux espèces. A l'inverse un chargement à haute force ionique dissocierait le complexe ADN-HU mais ne permettrait la fixation ni de l'un ni de l'autre sur le support.

Dans le cas de la protéine de type sauvage, on réalise une lyse basée sur des cycles de congélation-décongélation supplémentés par du lysozyme. Il s'en suit une étape de précipitation au sulfate d'ammonium et une ultra centrifugation pour éliminer l'ADN génomique bactérien. Cependant, cette élimination d'ADN est imparfaite, comme en témoigne la filtration difficile de l'extrait précédent les étapes de purification par chromatographie.

Dans le cas de la protéine de fusion, la lyse est effectuée par sonification pour fragmenter l'ADN et la première étape de purification de l'extrait obtenu est une chromatographie d'affinité réalisée à haute force ionique pour éliminer l'ADN sans effet entraînant sur la protéine fusion.

Lyse des bactéries (EB)

Le culot bactérien est repris dans le tampon de lyse (25 mM NaP pH 7,8 ; 2 M NaCl ; 0,1 mM PMSF ; 20 mM Imidazole) à raison de 10 mL de tampon de lyse par gramme de cellules. Le phosphate de sodium a été préféré au Tris pour sa stabilité de pH en fonction de la température. En effet, une variation de pH pourrait inhiber la fixation à l'étape de chromatographie d'affinité. On exclut également l'EDTA pour s'assurer de la stabilité du nickel fixé sur le support d'affinité.

La suspension est lysée par ajout de 1 mg/mL final de lysosyme (incubation 15 min à 30°C) et par des cycles de sonification (40 % de puissance sur un sonificateur 750 Watt équipé d'une sonde 13 mm). Le tout est placé dans la glace pour limiter l'échauffement lié à la sonification. Chaque cycle, d'une durée de 1 minute, contient des impulsions de 5 secondes espacés par des temps de

repos de 5 secondes également. Au cours des cycles de sonification, la suspension bactérienne passe d'un état fluide à un état visqueux qui correspond à la libération de la chromatine bactérienne dans le milieu de lyse. En continuant la sonification, on retrouve un aspect fluide, qui correspond au morcellement de l'ADN bactérien. Ce stade est généralement atteint au bout 6 à 10 cycles de sonification. Le lysat est ensuite centrifugé deux heures à 4°C et à 8000 rpm (rotor SS-34). Les débris membranaires se retrouvent dans le culot alors que les protéines et acides nucléiques solubles sont récupérés dans le surnageant. Ce surnageant est filtré à 0,2 µm. Le filtrat correspond à notre extrait brut (EB).

1^{ère} Chromatographie d'affinité (AF1)

L'extrait brut est chargé sur une colonne His-trapTM 1 mL (GE HealthcareTM) préalablement équilibrée dans le tampon His 4% B (25 mM NaP pH 7,8 ; 2 M NaCl ; 0,1 mM PMSF ; 20 mM Imidazole) équivalent au tampon de lyse. On lave ensuite avec 10 volumes de colonne de ce tampon puis on élue par un gradient linéaire sur 30 volumes de colonnes vers 100 % de tampon His B (25 mM NaP pH 7,8 ; 2 M NaCl ; 0,1 mM PMSF ; 500 mM Imidazole). L'élution est suivie par l'absorbance à 230 nm (courbe magenta **Figure 32**) et les fractions d'élution sont analysées par SDS-PAGE (**figure 32 C**).

Malgré la contribution importante de l'imidazole à l'absorbance à 230 nm, on distingue 3 pics lors de l'élution. L'analyse SDS-PAGE révèle que les fractions appartenant au premier pic correspondent à des contaminants et sont de fait élués à faible concentration en imidazole (100 mM, 20% B). On a ensuite un deuxième pic élué à environ 150 mM imidazole (30% B) dont les fractions comportent deux bandes majoritaires en SDS-PAGE. Leurs poids moléculaires apparents sont compatibles avec un mélange de HU de type sauvage et de HU fusionnée avec l'étiquette 8-His. Il semblerait que la protéine sauvage puisse échanger avec la protéine fusion pour former un héli-duplexe malgré l'étiquette d'affinité. Enfin un troisième pic est élué à haute concentration en imidazole (200 mM Imidazole, 40% B) et correspond au dimère de protéine fusion avec une étiquette 8 histidines par chaîne. Seules les fractions contenant la protéine HU fusionnée à l'étiquette 8-His sur les deux chaînes (fractions 23-32) sont rassemblées pour les étapes de purification suivantes et constituent la fraction « AF1 ».

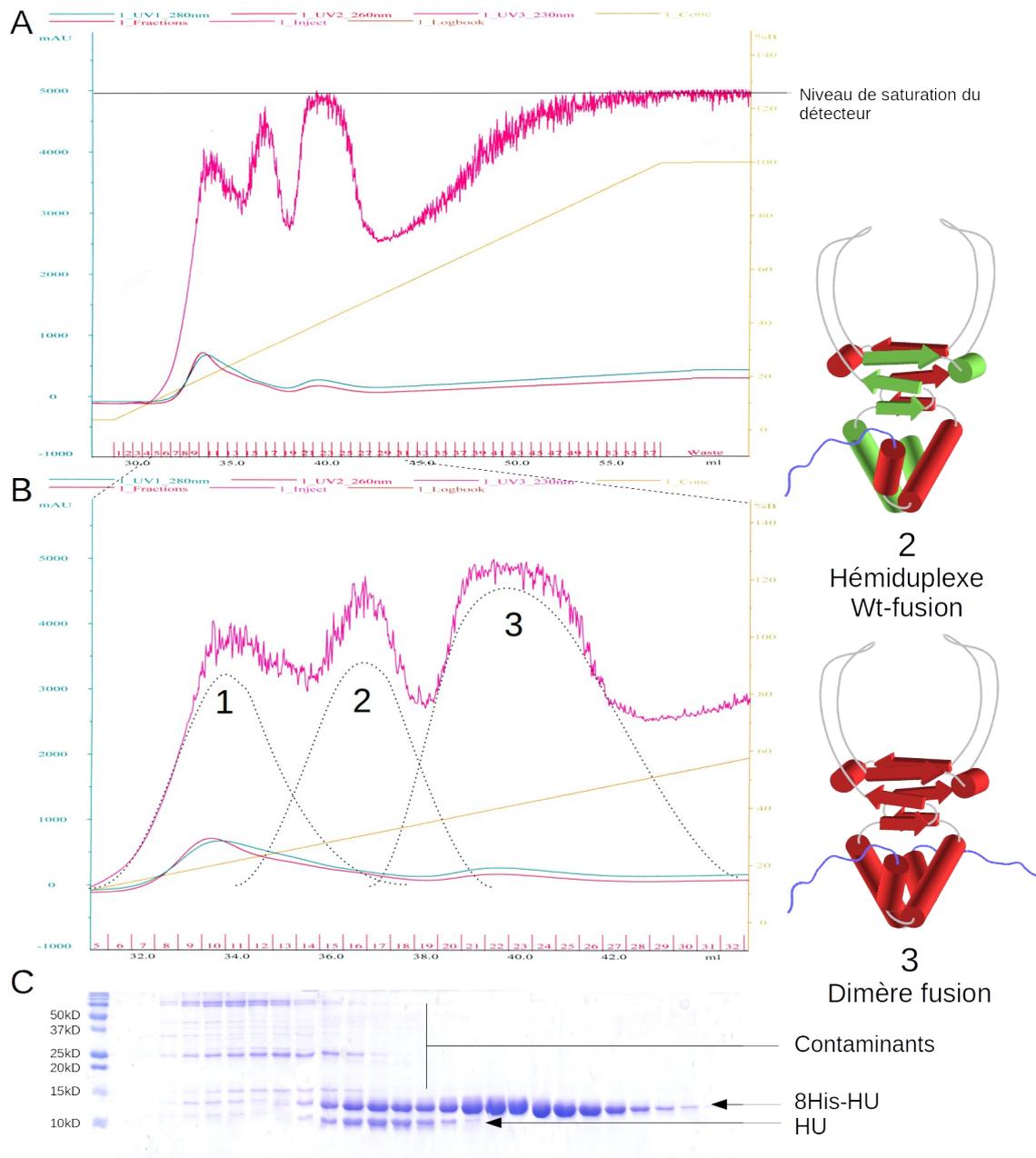


Figure 32: Première chromatographie d'affinité sur His-Trap A- Chromatogramme complet avec l'absorbance à 230, 260 et 280nm en magenta rouge et bleu, respectivement. La proportion en tampon B est indiquée en orange et les fractions d'élution sont numérotées en bas du chromatogramme. B- Agrandissement du chromatogramme sur la région d'élution des protéines. Une déconvolution est proposée en pointillé (pic 1 à 3) sur la base de l'analyse SDS-PAGE des fractions. Le pic 1 correspond a des contaminants élués à faible concentration en imidazole, le pic 2 correspond à un hémiduplexe entre la protéine sauvage naturelle (en vert) et la protéine fusion (en rouge) élué à une concentration intermédiaire en imidazole. Le pic 3, élué à haute concentration en imidazole, correspond à l'homodimère fusion (deux chaînes rouges). C- Analyse SDS-PAGE des fractions d'élution.

Clivage de l'étiquette 8His par l'entérokinase (EK)

Le tampon d'éluion de la fraction AF1 est échangé par le tampon entérokinase (20 mM NaP pH 7,8 ; 50 mM NaCl ; 2 mM CaCl₂) par centricon™ Amicon™ (10 kD). La concentration de la protéine est ajustée à 1 mg/ml, en faisant l'approximation qu'elle seule contribue à l'absorbance de l'échantillon. On teste la protéolyse de l'étiquette sur des petits volumes (10 µL) à différents rapports Enterokinase/HU et à deux forces ioniques. La digestion est opérée sur la nuit soit à température ambiante (~25°C) soit à 16°C. Les produits de digestion sont ensuite analysés par SDS-PAGE (**figure 33**).

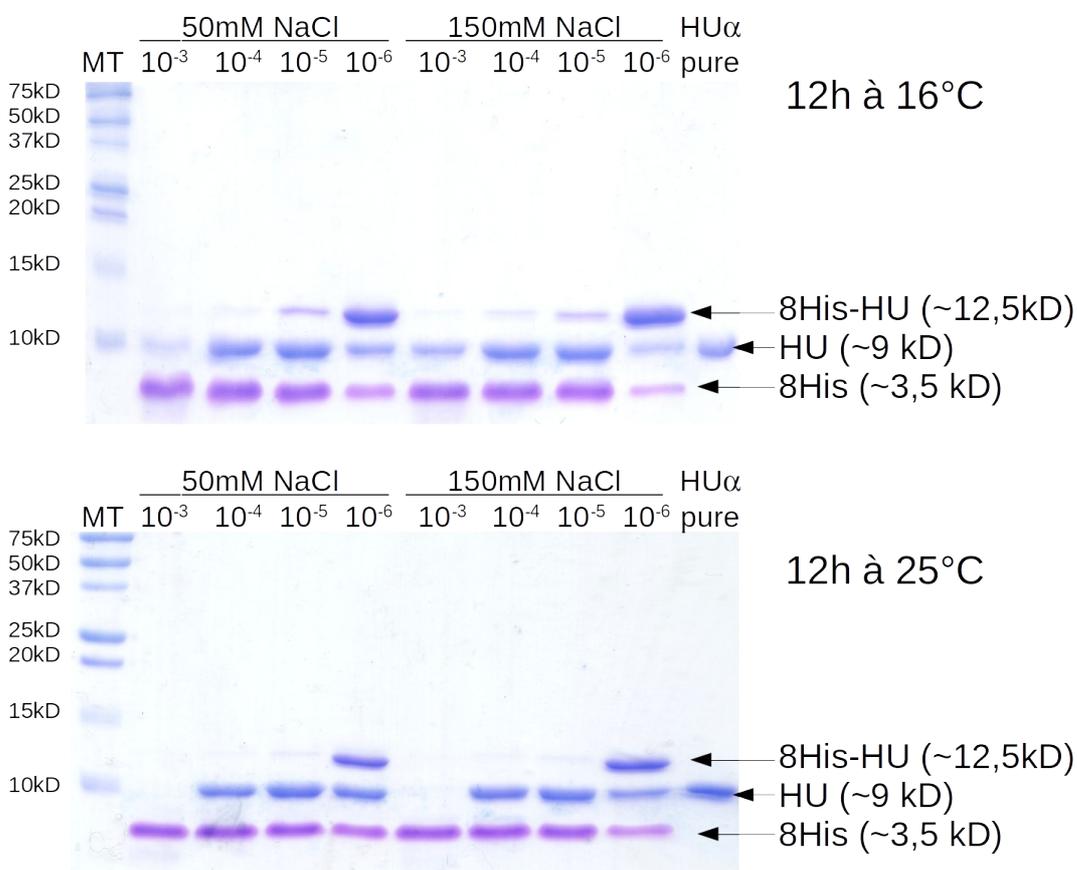


Figure 33: Analyse SDS-PAGE du clivage protéolytique de l'étiquette 8His 4 rapports de concentrations molaires Enterokinase/HU (10⁻³ à 10⁻⁶), deux forces ioniques (50 mM à gauche et 150 mM NaCl à droite) et deux températures d'incubation (16°C en haut et 25°C en bas) sont testés sur une incubation de 12 h.

On observe de la dégradation non spécifique de HU lorsque le rapport entérokinase/HU est trop élevé (10^{-3}) en particulier à 25°C. Les bandes de HU sont d'ailleurs un peu plus intenses sur le gel des produits de clivage obtenus à 16°C. On préfère donc se placer à 16°C pour préserver la protéine. La concentration en NaCl ne semble pas influencer sur le clivage ni en terme d'activité ni en terme de spécificité de coupure. On préfère donc se placer à 150mM NaCl pour conserver la protéine dans son état le plus stable. Le facteur 10^{-5} entérokinase/HU semble un bon compromis entre efficacité de clivage et consommation en protéase.

En conclusion, les tests analytiques nous ont permis de retenir les conditions de clivage suivantes : 16°C, 12 h, rapport 10^{-5} protéase/HU, 1 mg/ml pour la concentration de HU. La fraction AF1 est donc clivée dans ces conditions. La protéolyse est arrêtée par ajout de 0,1 mM final de fluorure de phénylméthylsulfonyle (PMSF) qui est un inhibiteur de protéase à sérine dont fait partie l'entérokinase.

Remarque : la bande de l'étiquette 8His présente une coloration différente au bleu de coomassie par rapport aux autres protéines. C'est assez inattendu mais ce type de coloration du bleu de coomassie a été observé pour des pH très acides (coloration rougeâtre) ou très basiques (coloration plutôt rosée).²⁰³ Dans notre cas, on peut faire l'hypothèse que cette coloration est liée à la séquence de l'étiquette libre qui est très amphiphile. 4 acides aspartiques du côté carboxy-terminal forment un cluster acide (avec 5 fonctions acide carboxyliques) alors que les 8 histines forment un cluster plutôt basique côté N-terminal. Il est possible que l'état de protonation du coomassie soit influencé par ces environnements locaux et donne cette couleur violette sur les gels SDS-PAGE.

2^{ème} chromatographie d'affinité (AF2)

Une fois l'étiquette histidine clivée par l'enterokinase, l'étiquette libre ainsi que d'autres contaminants co-élués lors de la première chromatographie d'affinité sont éliminés en passant le protéolysat une seconde fois sur cette colonne. La protéine HU libre n'est pas retenue sur la colonne. Le « flow through » (FT) de la colonne est collecté et constitue la fraction « AF2 ». Un lavage à haute concentration en imidazole (+imi **figure 34**) permet d'éliminer l'étiquette 8His qui était restée fixée sur la résine d'affinité.

Chromatographie échangeuse d'ion (20CM)

Les derniers contaminants difficilement distinguables sur SDS-PAGE sont éliminés par une étape de chromatographie échangeuse d'ions. On utilise une colonne POROS™ HS-20 (échangeur fort de cations avec des billes de type chromatographie par perfusion). Cette résine supporte de très hautes pressions (14 MPa) et permet une purification à haut débit (10 mL/min). La fraction AF2 est fixée à basse force ionique (50 mM NaCl) et la colonne est ensuite éluée par un gradient linéaire en NaCl. Le chromatogramme correspondant est présenté **Figure 35**. HU se fixe assez fortement sur ce support et le premier gradient, prévu initialement jusqu'à 600mM NaCl, n'était pas assez élevé. Il a été poursuivi manuellement ce qui a généré l'artefact que l'on observe sur le chromatogramme (**Figure 35**).

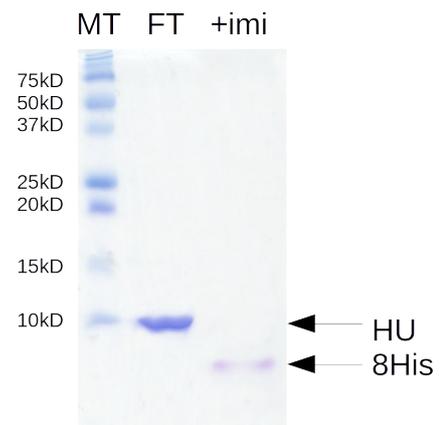


Figure 34: Seconde chromatographie d'affinité Analyse SDS PAGE avec un marqueur de taille (MT) la fraction non fixée sur la colonne (FT) et la fraction éluée par un lavage 0,5M imidazole (+imi).

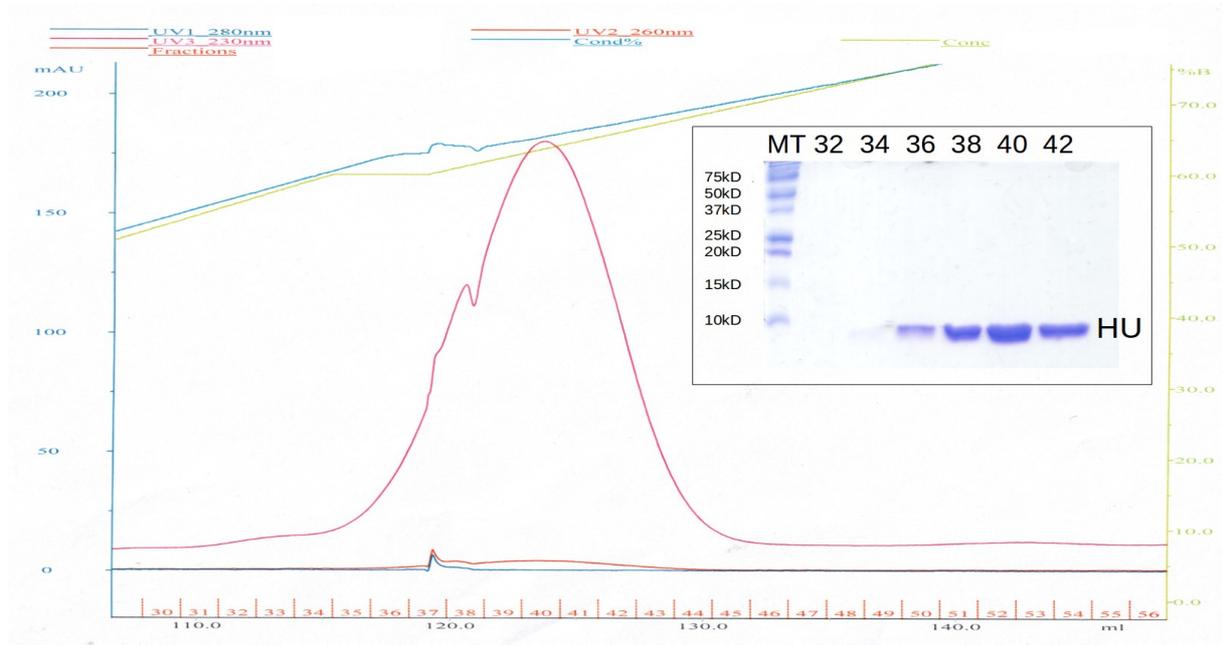


Figure 35: Chromatographie Echangeuse d'ions HS-20 Avec l'absorbance à 230, 260 et 280nm respectivement en magenta, rouge et bleu. La proportion de tampon SP-B au cours du gradient est indiqué en vert et la conductimétrie en bleu clair. Les fractions d'élution de 1mL sont numérotées en bas de chromatogramme et le gel SDS-PAGE correspondant est présenté dans l'encart.

La protéine est purifiée à homogénéité comme en témoigne la faible absorption à 280 nm par rapport à 230 nm. En effet, HU est dépourvue de résidus tyrosines et tryptophanes et n'absorbe quasiment pas à 280 nm. La protéine est cependant très faiblement concentrée et une précipitation à l'acétone a été nécessaire pour obtenir une concentration en protéine suffisante pour être observable sur SDS-PAGE.

3.2.6. Contrôle de repliement par RMN

La structure d'une protéine étant intimement liée à sa séquence primaire, l'utilisation de systèmes fusion pour la purification de protéine est toujours sujet à caution. En effet, une protéine produite avec une étiquette d'affinité clivée *a posteriori* peut avoir une structuration altérée par rapport à une protéine directement produite et purifiée sans étiquette. Pour contrôler la structuration de la protéine HU après clivage de son étiquette d'affinité, la protéine HU α_2 marquée ^{15}N a été purifiée selon le protocole détaillé ci-avant. Moins d'un milligramme de protéine a été obtenu à partir des 5 g de culot bactérien et nous a permis d'obtenir un échantillon à une concentration de seulement 100 μM (300 μL en tube Shigemi). Un spectre ^{15}N -HSQC a été enregistré avec cet échantillon (**Figure 36 en bleu**) et comparé avec un spectre de référence sur un échantillon issu d'une production sans étiquette d'affinité (**Figure 36 en rouge**) acquis dans les mêmes conditions à l'exception de la concentration (1 mM). Malgré le faible niveau de signal sur bruit du spectre de la protéine issue de la fusion, la grande majorité des pics sont visibles et superposables avec ceux de la protéine native. La production avec l'étiquette 8-Histidines n'altère donc pas le bon repliement du dimère de HU. Cette observation est cohérente avec le travail de Pelligrini et collaborateurs¹⁹² qui ont montré que la protéine HU fusionnée avec une étiquette 6-Histidines N-terminale conservait son activité de fixation sur l'ADN. Ce qui implique que la structure tridimensionnelle n'a pas été altérée par la présence de l'étiquette.

3.2.7. Conclusion

On obtient environ 1 mg de protéine pour 10 g de cellules ce qui est très insuffisant pour de futures études structurales des variants de HU. Le niveau d'expression de la protéine fusion est trop faible, au point où la contamination par la protéine de type sauvage est significative. L'optimisation du protocole ou la mise en place d'un protocole alternatif est nécessaire pour atteindre cet objectif.

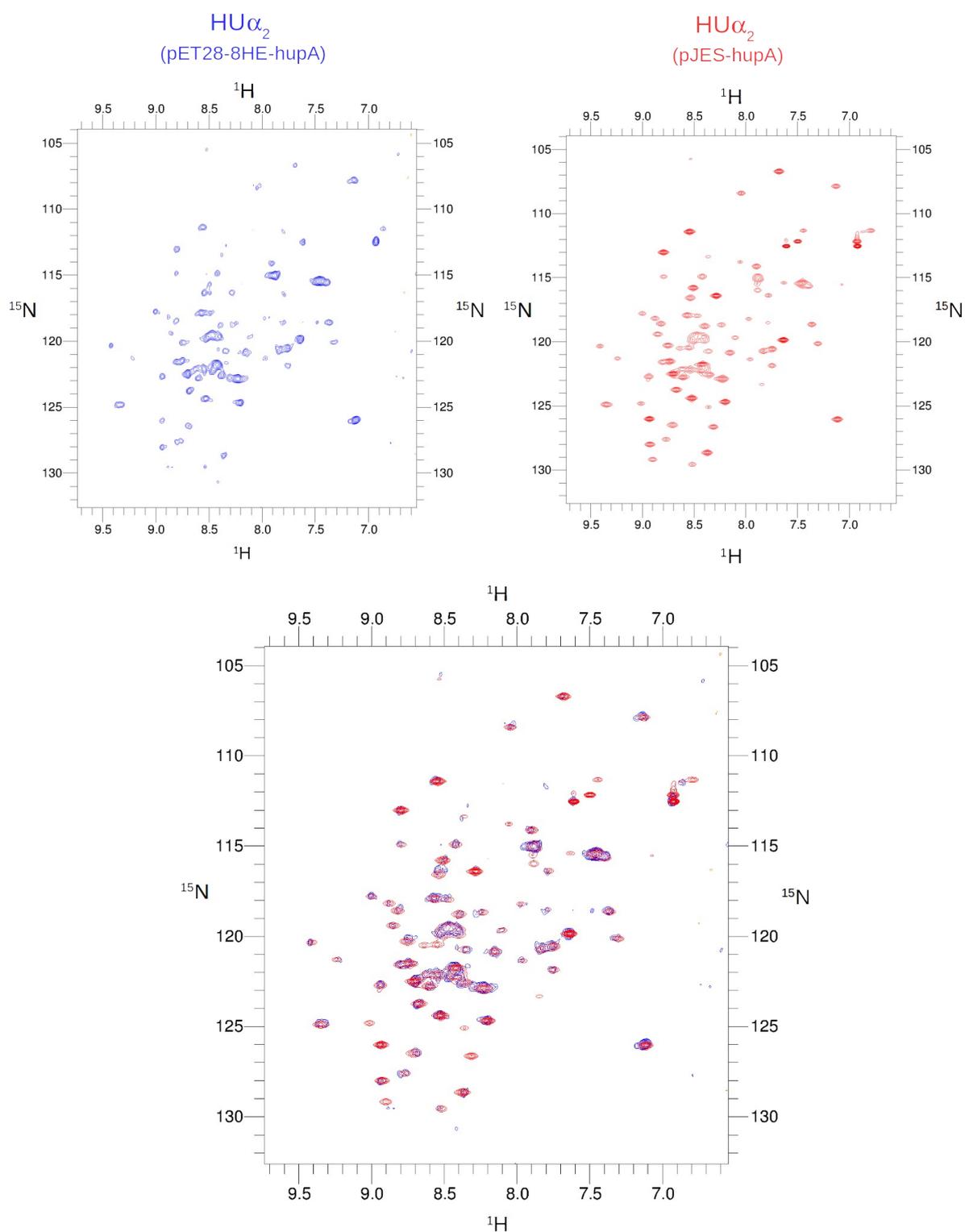


Figure 36: Spectres ^{15}N -HSQC comparatifs entre la protéine HU α_2 fusionnée 8-His clivée (en bleu), et la protéine HU α_2 de type sauvage (en rouge) Enregistrée à 600Mhz avec un échantillon à 100 μM pour la protéine issue de pET28-8HE-hupA et avec un échantillon à 1mM pour la protéine issue de pJES-hupA.

3.3. HU fusionnée avec la Maltose Binding Protein (MBP)

Face aux problèmes d'expression rencontrés avec l'étiquette 8-Histidines, nous avons décidé de tester une autre étiquette d'affinité, la Maltose Binding protein (MBP). Cette étiquette a été employée avec succès pour d'autres protéines fusions étudiées au laboratoire.²⁰⁴ L'étiquette MBP est aussi connue pour améliorer la solubilité de la protéine sur-exprimée.

3.3.1. Stratégies alternatives pour le clonage de gènes

Parmi les méthodes de clonage disponibles, la plus couramment employée repose sur l'utilisation d'enzyme de restrictions sur un plasmide et un insert comprenant le gène d'intérêt. Ces enzymes génèrent des extrémités cohésives ou franches sur les deux partenaires qui peuvent être ligués par voie enzymatique (ligase) et ainsi générer la construction désirée. C'est la méthode que nous avons utilisée pour le clonage des gènes *hupA* et *hupB* dans le vecteur recombinant codant pour 8His-HU. Bien que robuste, cette méthode présente la contrainte de nécessiter des sites de restrictions adaptés sur les deux partenaires. De plus, le succès de chaque étape n'est pas systématiquement vérifiable et certaines étapes peuvent présenter des rendements faibles (purification sur gel, digestion à l'extrémité de l'insert, ligation) n'aboutissant parfois à aucun clone comportant la construction désirée à la fin de l'expérience.

Pour gagner du temps sur les expériences de clonage, des méthodes alternatives ont été envisagées pour le clonage des gènes *hupA* et *hupB* dans un vecteur conçu pour l'expression de protéines fusionnées à la MBP (pMAL-c5e). Les méthodes alternatives de clonage de gènes sont nombreuses et possèdent chacune leurs limites. On décrira succinctement quelques méthodes avec leurs avantages et leurs limites. Un schéma bilan est présenté **Figure 37**.

Le clonage TA²⁰⁵ exploite l'absence d'activité exonucléase 3' → 5' d'ADN polymérases telles que la Taq polymérase qui peut ajouter une Adénine sur les extrémités 3' de l'insert. Le vecteur est parallèlement digéré par une enzyme de restriction (ou amplifié par PCR) générant des extrémités franches et une thymine est ajoutée sur les extrémités 3' par une terminale transférase. On ligue ensuite les deux partenaires par une ligase. Cette méthode couple les avantages de la génération de partenaires à bouts francs sans digestion de l'insert et une ligation plus efficace sur des partenaires à bouts cohésifs (bien que sur une seule paire de base). Néanmoins, cette méthode nécessite un traitement supplémentaire du vecteur et n'est pas directionnelle. De plus, elle ajoute une paire AT

aux points de ligation.

La méthode Gateway^{206,207} exploite le phénomène de recombinaison site-spécifique pour l'intégration de gènes dans un vecteur. Cette méthode, adaptée au haut débit, présente une haute spécificité et efficacité. Cependant, elle nécessite la présence de séquences spécifiques, des sites de recombinaison, à la fois sur les inserts et les vecteurs. Un clonage initial est réalisé pour constituer le vecteur dit d'entrée (avec des sites de recombinaison flanquant l'ORF du gène). Ce dernier pourra ensuite être utilisé pour le sous-clonage par recombinaison dans différents vecteurs de destination (pour l'expression avec différents tags ou promoteurs par exemple) qui comportent les sites de recombinaison compatibles. Bien que permettant le sous clonage rapide et en parallèle dans de multiples vecteurs à tester, la nécessité d'un clonage initial dans un vecteur d'entrée ajouté à la nécessité de la présence de séquences de recombinaison sur chaque vecteur (non disponibles dans la plupart des vecteurs commerciaux) rendent cette méthode trop contraignante dans notre cas.

Au contraire du Gateway, **le clonage par recombinaison homologue** ne nécessite pas de séquences spécifiques en bordure de gène pour le clonage. Les séquences bordant l'insert et le site de clonage doivent cependant être homologues. Il existe 3 mécanismes de recombinaison homologue : RecA-dépendant, RecA-indépendant et Red/ET-dépendant. Le mécanisme RecA-dépendant nécessite des séquences homologues entre les partenaires de l'ordre de 1 kbp et des souches RecA⁺.²⁰⁸ Elle est plutôt utilisée pour la reconstitution de génomes recombinants.²⁰⁹ La recombinaison homologue RecA-indépendante fonctionne dans des souches RecA et est réalisée par co-transformation du vecteur et de l'insert linéaire ayant plus de 10 pb homologues à leurs extrémités.²¹⁰ Bien qu'aisé à mettre en pratique, ce mécanisme de recombinaison n'est pas bien connu et présente une faible efficacité, sauf dans le cas où l'origine de répllication est apportée par l'insert. La recombinaison Red/ET-dépendant nécessite la co-transformation du vecteur et de l'insert linéaire comportants 30 à 50 pb homologues à leurs extrémités dans une souche adaptée (comme recBC sbcA) exprimant RedE/RedT mais pas RecBCD.²⁰⁸ Ces trois mécanismes présentent des méthodologies similaires mais nécessitent des souches particulières ou présentent de faibles efficacités.



Figure 37: Méthodes alternatives de clonage

* : 1 kbp, 10 bp ou 30 bp pour recombinaison RecA-dépendante, RecA-indépendante ou Red-ET-dépendante respectivement.** : Souche RecA⁺ (recombinaison RecA-dépendante) ou RecBC sbcA (Red-ET-dépendante).

La méthode SLIC (Sequence and Ligation Independent Cloning) détourne le système de réparation de l'ADN par recombinaison homologue pour le clonage de gène *in vitro*.²¹¹ Un intermédiaire de réparation est généré en plusieurs étapes. Le gène d'intérêt est amplifié par PCR en ajoutant des régions de séquences identiques à celles bordant le site de clonage du vecteur cible. Le vecteur linéarisé et l'insert sont ensuite traités par une exonucléase (comme l'ADN polymérase T4) afin de générer des extrémités simples brin (phénomène de résection) d'environ 30 pb. Les partenaires sont hybridés en présence de RecA (nécessaire en cas de faibles concentrations uniquement, <1 ng/μL) puis transformés dans des bactéries. Le système de réparation bactérien permet d'obtenir le vecteur, contenant le gène d'intérêt et double brin. Cette méthode présente l'avantage de ne pas nécessiter de site spécifique d'endonucléase ou de recombinaison, mais nécessite des étapes supplémentaires pour la résection (qui doit être bien calibrée pour obtenir des extrémités cohésives de taille souhaitée) et pour l'hybridation dans certains cas.

La dernière méthode discutée ici, et celle que nous avons retenue, est la méthode de **clonage par PCR à « méga-amorce »** (« mega priming PCR » ou « Restriction Free cloning »).²¹² Cette méthode est comparable au QuickChange™ à la différence qu'un gène est inséré à la place d'une mutation. Les amorces assez longues (~40-50pb) comportent deux régions : une région permettant l'amplification de l'insert du côté 3' et une région homologue au plasmide côté 5' (**Figure 38**). L'ADN source et le vecteur cible sont présents dans le même milieu réactionnel. Durant les premiers cycles de PCR, le vecteur cible n'est pas amplifié car l'amorce ne présente pas d'extrémité 3' hybridée contrairement au gène d'intérêt (**Figure 38**). À partir du 3^{ème} cycle, les inserts générés par PCR présentent des extrémités 3' compatibles avec le vecteur cible et permettent son amplification en intégrant le gène d'intérêt. Le produit de PCR est traité par l'endonucléase DpnI pour éliminer le vecteur cible vide puis est utilisé pour transformer des bactéries.

Cette dernière méthode de clonage par PCR à méga-amorce présente l'avantage d'être réalisable en pratiquement une seule étape. Elle ne nécessite aucun traitement supplémentaire pour la préparation de l'insert ou du vecteur avant clonage, ni d'enzyme ou de vecteur spécifique pour être réalisée. L'insertion peut être contrôlée directement par électrophorèse sur gel d'agarose.

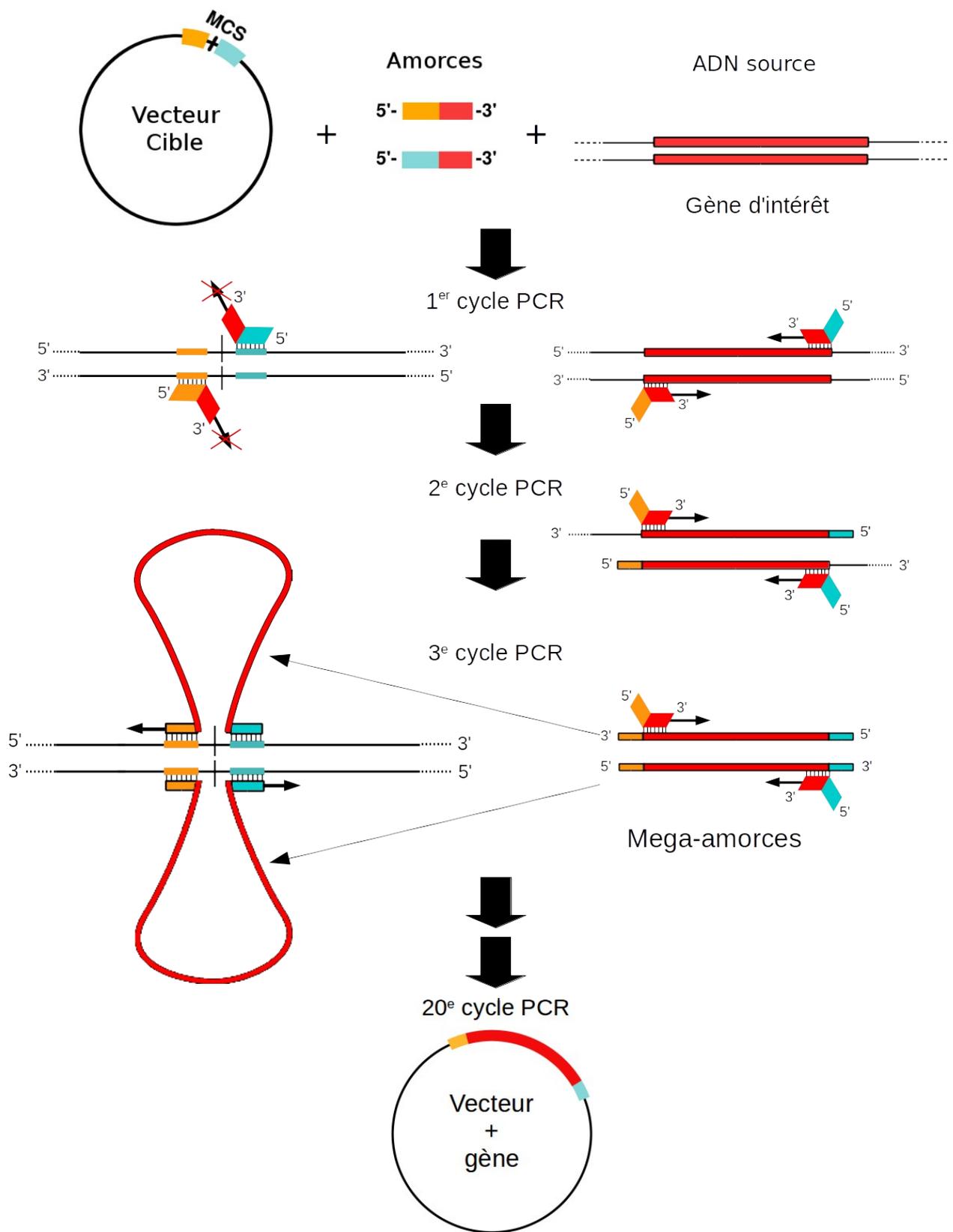


Figure 38: Schéma de clonage sans site de restriction par PCR à méga-amorces

3.3.2. Construction des vecteurs pMAL-c5E-hupA/B

Les gènes *hupA* et *hupB* ont été clonés dans le vecteur pMAL-c5e par PCR à méga-amorces en utilisant les couples d'amorces 11/12 et 13/14, respectivement (**Tableau 24 en annexe**). Différentes concentrations en amorces ont été testées (**Figure 39**). Moins la concentration en amorces est importante, moins l'hybridation des méga-amorces néosynthétisées sera concurrencée. C'est ce qu'on observe sur le gel d'agarose (**Figure 39**). 50 nM d'amorces sont utilisés pour produire le vecteur d'intérêt par PCR. Après digestion DpnI et transformation dans une souche NEB-10 β , on réalise une mini-prep sur des clones isolés. La clonage correct des gènes *hupA* et *hupB* a été contrôlé par séquençage.

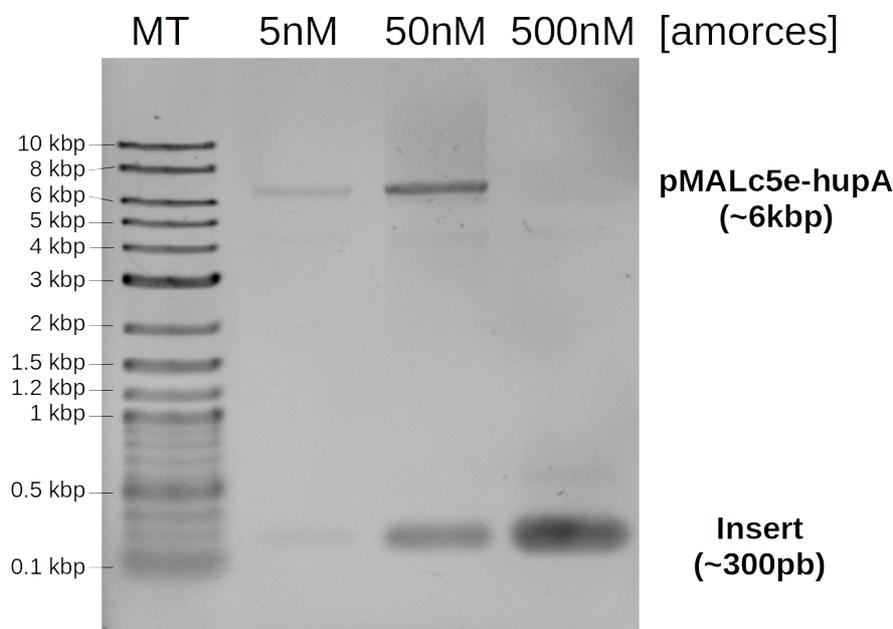


Figure 39: Contrôle de clonage du gène *hup* dans pMALc5e par PCR à méga-amorces Electrophorèse sur gel d'agarose (0,8%) des produits de PCR.

3.3.3. Tests d'expressions

Pour exprimer la protéine fusion, on suit le protocole standard d'expression de HU en ajoutant du glucose (2 % final) dans le milieu LB. Le glucose en excès permet d'inhiber la production d'amylases qui pourraient dégrader plus rapidement le support de chromatographie d'affinité (colonne amylose).

Des extraits solubles et totaux ont été réalisés à différents temps d'induction et analysés par SDS-PAGE. On observe une forte sur-expression dans les extraits totaux en comparaison avec le niveau d'expression de la fusion 8His. Bien que ce niveau soit élevé dès 1h d'incubation, on observe une augmentation de la quantité de protéine soluble à des temps d'incubation plus long. On a donc fixé la durée d'induction à 3h pour les prochaines cultures.

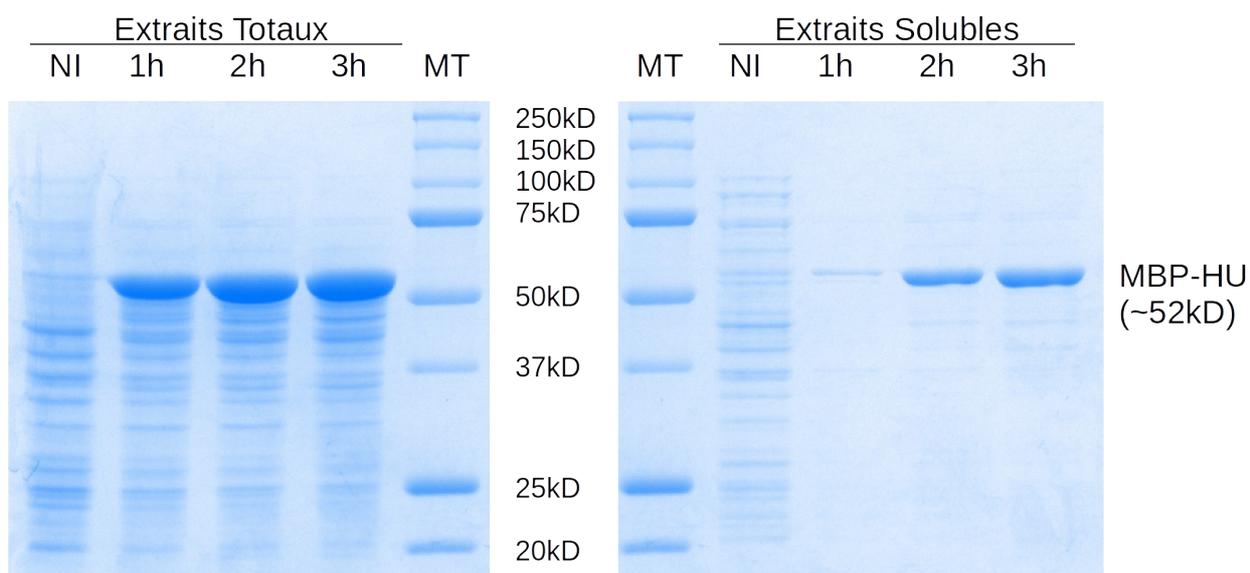


Figure 40: Analyse SDS-PAGE de l'expression de la protéine HU α_2 fusionnée à la MBP

Ce niveau de sur-expression est très encourageant. On peut espérer obtenir, à l'issue de la purification, des quantités de protéines suffisantes pour des études structurales.

3.3.4. Purification de la protéine

Lyse des bactéries (EB)

Nous avons opté pour une stratégie analogue à la méthode de purification de la fusion 8His. On resuspend les culots cellulaires à raison de 10 mL de tampon de lyse (25 mM Tris/HCl pH 7,5 ; 800 mM NaCl ; 1 mM EDTA ; 0,1 mM PMSF) par gramme de cellules. La lyse est réalisée conjointement par sonication et par le lysosyme (1 mg/mL final). La suspension est centrifugée 2 h à 4°C et 8000 rpm. On récupère le surnageant et on filtre à 0,2 μ m. Ce filtrat constitue notre extrait brut (EB **figure 45**).

Chromatographie d'affinité (Amy)

La fraction EB est chargée sur une colonne d'amylose HF (New England Biolabs™) équilibrée dans le tampon AmyA (25 mM Tris/HCl pH 7,5 ; 800 mM NaCl). L'extrait brut est ensuite chargé sur la colonne, puis on lave avec du tampon Amy A jusqu'au retour de l'absorbance à 280 nm au niveau de la ligne de base (**Figure 41 à gauche**). On élue ensuite la protéine fusion avec la solution AmyB (25 mM Tris/HCl pH 7,5 ; 800 mM NaCl ; 100 mM Maltose). Des échantillons sont prélevés au cours de l'élution et analysés par SDS-PAGE (**Figure 41 panneau de droite**). Les fractions d'élution sont rassemblées et constitue la fraction « Amy ».

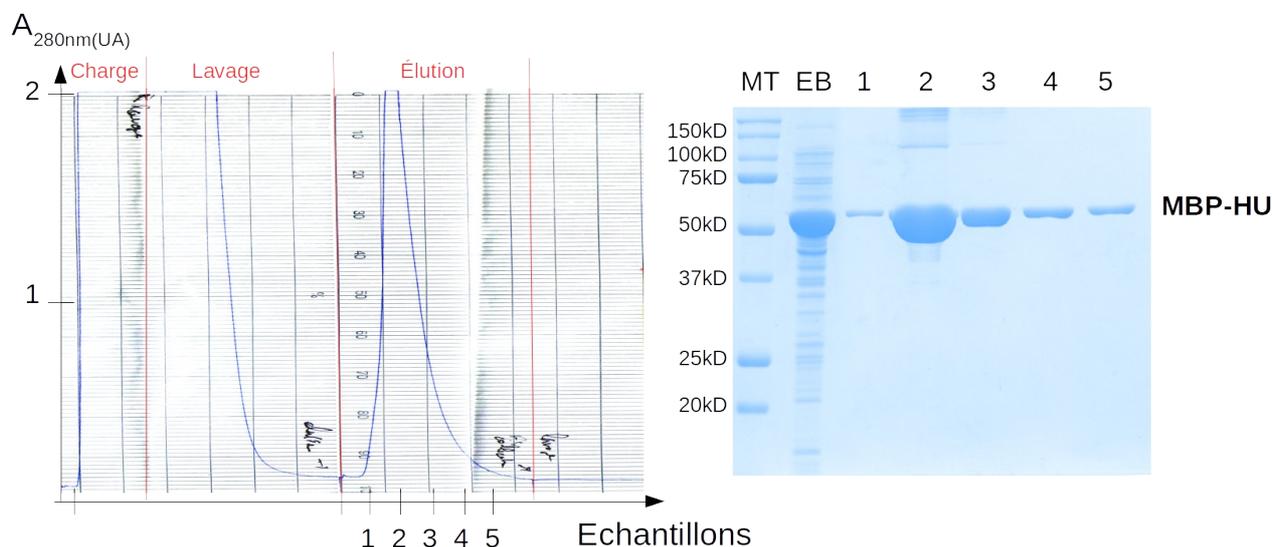


Figure 41: Chromatographie d'affinité sur Amylose HF A gauche le chromatogramme avec en bleu l'absorbance à 280nm. Les étapes de purification sont indiquées en rouges au dessus du chromatogramme et les fractions prélevées pour analyse SDS-PAGE sont indiquées en noir en dessous. A droite l' analyse SDS-PAGE de 5 échantillons prélevé au cours de l'élution.

Clivage de l'étiquette MBP par l'entérokinase (EK)

On échange le tampon d'élution dans lequel se trouve la fraction Amy contre un tampon compatible avec l'entérokinase (25 mM Tris/HCl pH7,5 ; 150 mM NaCl ; 2 mM CaCl₂) par centricon™ Amicon™ (30 kDa). La concentration de fusion ne doit pas dépasser 10mg/mL (11UA à 280nm si la fusion était pure d'après son coefficient d'extinction molaire). Au dessus de 10mg/mL, la protéine forme un précipité jaunâtre dans le centricon.

Malgré les tests de clivage réalisés sur la fusion 8-His-HU, on ne sait pas *a priori* si l'étiquette MBP, de grande taille (~42 kDa), modifie la cinétique de clivage par l'entérokinase. Des tests dans des petits volumes (10 μ L) ont été réalisés pour déterminer les meilleures conditions de clivage. On considère que la majorité de l'absorbance à 280 nm provient de la protéine fusion et on calcule les rapports protéase/protéine en utilisant la concentration en MBP-HU maximum équivalente. La concentration en protéine fusion est ajustée à 10 mg/mL, soit environ 11 UA_{280nm}, et on réalise la protéolyse à différents rapports de concentrations molaires protéase/protéine (**Figure 27**). Du fait de la grande différence de taille entre la protéine HU fusionnée et la protéine HU libre, deux analyses SDS-PAGE sont présentées. Sur le panneau de droite, l'analyse par SDS-PAGE 12 % permet une bonne résolution des protéines de tailles élevées telles que la protéine fusion MBP-HU (~52 kDa) et l'étiquette MBP libre (~43 kDa). A gauche, un gel à 16 % acrylamide et une quantité supérieure d'échantillon chargé permet l'observation de la protéine HU libérée. On observe sur ces gels que le rapport protéase/protéine 10^{-7} semble le plus adapté pour le clivage de l'étiquette MBP. Ce rapport est meilleur que celui utilisé pour la fusion His, probablement parce que la concentration en substrat est plus élevée (~10mg/mL au lieu de 1mg/mL) et l'incubation est un peu plus longue (15h au lieu de 12h). La taille de l'étiquette ne semble pas représenter un obstacle dans pour le clivage.

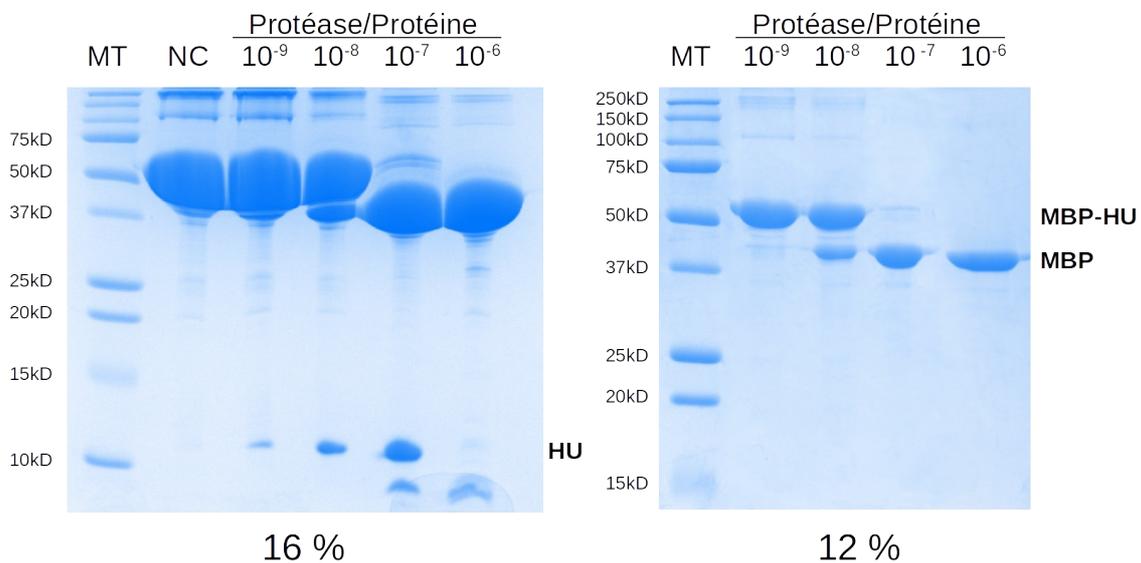


Figure 42: Clivage de l'étiquette MBP par l'entérokinase Analyse SDS-PAGE après une incubation de 15h à 16°C. A gauche, l'analyse SDS-PAGE 16 % acrylamide permet l'observation de la protéine HU libérée de l'étiquette MBP. A droite, l'analyse SDS-PAGE 12 % acrylamide avec moins de matériel chargé et une migration plus longue, permet de séparer la fusion de la MBP seule.

Le clivage est stoppé par ajout de PMSF (0,1 mM final). La protéase est ensuite éliminée par un passage du protéolysat sur une colonne Benzamidine-Sepharose (GEHealthcare™) qui fixe spécifiquement les protéases à sérine telles que l'entérokinase. La fraction de purification résultante est nommée « EK ».

Chromatographie échangeuse d'ion SP-FF (FF)

Remarque : Du fait de la grande quantité de protéine et pour maintenir une bonne séparation, cette étape a été réalisée en deux fois.

Après ajustement de la fraction EK à 50 mM NaCl (par dilution) elle est chargée sur une colonne SP-FF équilibrée dans un tampon 95 % SP-A (25 mM Tris-HCl pH7,5 ; 0,1 mM PMSF ; 1 mM EDTA) 5% SP-B (25 mM Tris-HCl pH 7,5 ; 1 M NaCl ; 0,1 mM PMSF ; 1 mM EDTA). Après chargement de l'échantillon, on observe un pic d'élution au cours du lavage à 50 mM NaCl. Ce pic correspond à la fusion non clivée et peut être mis de côté pour une nouvelle purification (**Figure 43 B et C panneau de gauche**). Les protéines fixées sur la colonne sont ensuite éluées par un gradient linéaire en NaCl. Bien que l'analyse SDS-PAGE des fractions d'élution semble indiquer un échantillon homogène (**Figure 43 C panneau de droite**) un épaulement sur le chromatogramme montre le contraire (**Figure 43 B panneau de droite**). De plus, l'analyse SDS-PAGE sur l'ensemble des fractions collectées à l'issue de cette étape (**Figure 45 FF**) présente des traces de fusion non clivée dans cet échantillon (présence probable de traces d'hétérodimères HU/MBP-HU).

D'autres étapes sont donc nécessaires pour purifier cet échantillon à homogénéité. La grande différence de taille entre la protéine fusion (~52 kDa) et la protéine libre (~19 kDa) permet de les séparer par tamisage moléculaire (S200).

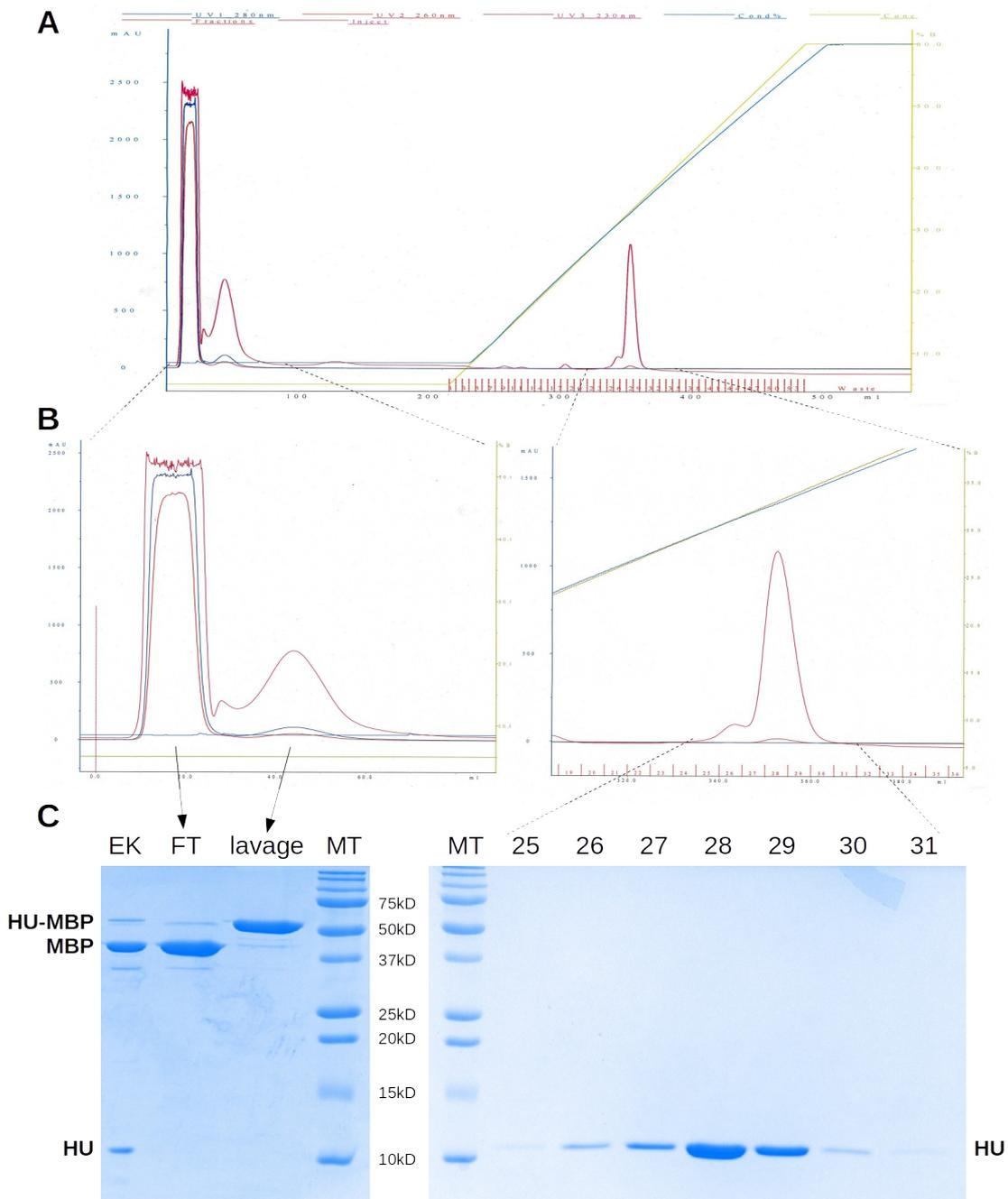


Figure 43: Chromatographie échangeuse de cations SP-FF A- Chromatogramme global avec bleu, rouge et magenta l'absorbance à 280, 260 et 230 nm, respectivement. La proportion en solution B au cours du gradient est représentée en vert et la conductimétrie en bleu clair. Les fractions collectées automatiquement sont numérotées en rouge en bas du chromatogramme. B- Agrandissement du chromatogramme sur les régions d'intérêt. C- Analyse SDS-PAGE (avec MT : marqueur de taille ; EK produit de protéolyse à l'entérokinase ; FT : protéines non retenues ; « lavage » pic élué au cours du lavage).

Tamisage Moléculaire (S200)

La fraction FF est chargée sur une colonne S200 (GE Healthcare™) préalablement équilibrée dans le tampon SP-B. Le chromatogramme correspondant est présenté sur le panneau de gauche de la **Figure 44**. Des contaminants de hauts poids moléculaires ont pu être éliminés dans le volume d'exclusion (V(Exclusion)). On voit qu'il subsiste une espèce contaminante qui absorbe légèrement à 280nm et qui contamine encore l'échantillon. Les fractions correspondant à l'ensemble du pic HU α 2 (28 à 44) sont rassemblées et constituent la fraction « S200 ».

Chromatographie échangeuse de cation (20CM)

La séparation d'espèces proches sur échangeur d'ion (comme les différents dimères de HU) peuvent être séparés sur une résine échangeuse faible de cations (20CM). Après ajustement de la fraction S200 à 50mM NaCl, on la charge sur une colonne 20CM préalablement équilibrée dans le tampon 95 % SP- 5% B (25 mM Tris-HCl pH 7,5 ; 50mM NaCl ; 0,1 mM PMSF ; 1 mM EDTA). Après lavage par 10 volumes de ce tampon, la protéine est éluée par un gradient linéaire en NaCl. Une très

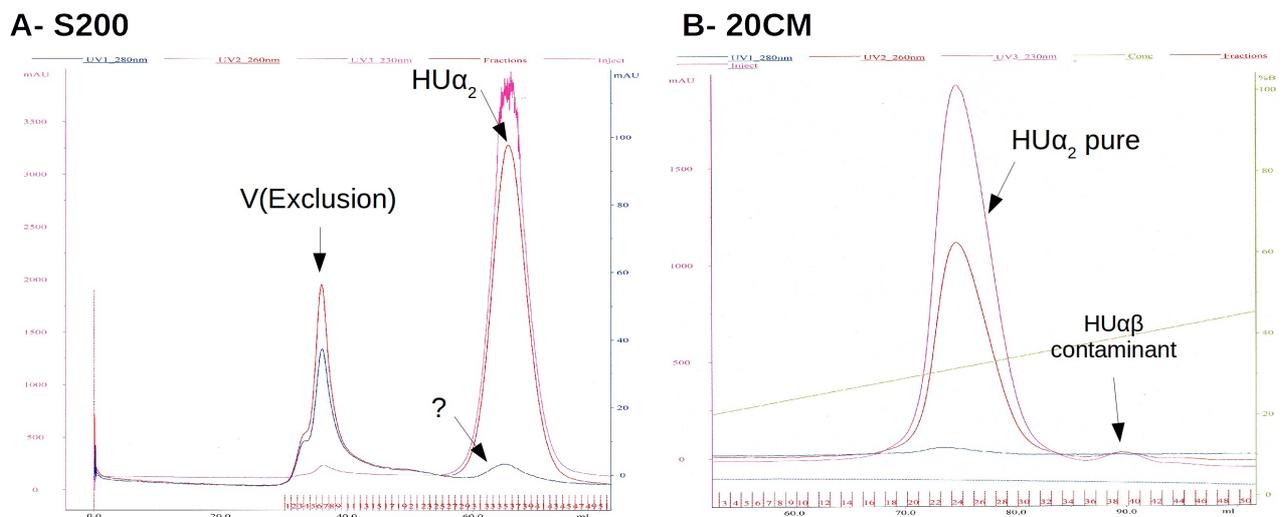


Figure 44: Tamisage moléculaire (A- S200) et Chromatographie échangeuse de cation (B- 20CM)

faible quantité d'hétérodimère est visible sur le chromatogramme et correspond à une contamination par la protéine sauvage. Cependant cette quantité est négligeable et la protéine recombinante est aisément isolée. On rassemble les fractions d'éluion du pic de HU α 2 pure (20-30) qui constitue la fraction 20 CM (**Figure 45**).

3.3.5. Conclusion

Le gel résumé de la purification de HU α_2 avec une étiquette MBP est présenté **Figure 45**. L'échantillon final est purifié à homogénéité, sans aucun contaminant visible sur SDS-PAGE.

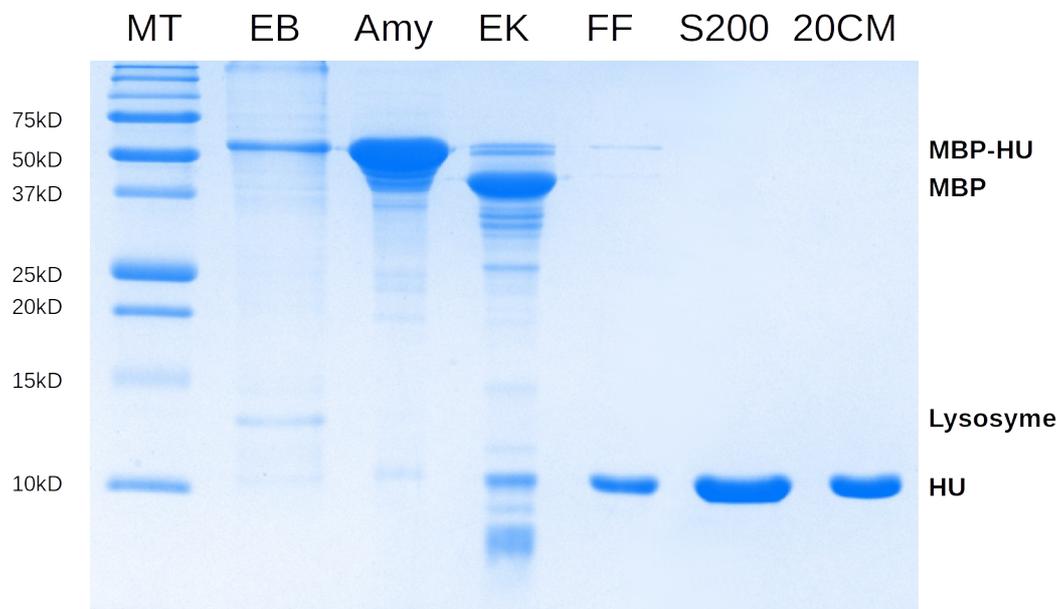


Figure 45: SDS-PAGE: Bilan de purification de la fusion MBP-HU

Les quantités (15 mg de protéine pour 10 g de cellule) et le niveau de pureté atteint en utilisant la fusion avec la MBP sont satisfaisants pour des études structurales. Nous avons opté pour ce protocole pour la production et la purification des variants de HU. Le bon repliement de la protéine purifiée en suivant cette méthode a été vérifiée par RMN en superposant les spectres ^{15}N -HSQC des protéines marquées issues de la purification standard avec pJES ou celle avec pMAL. Aucune différence significative n'a pu être observée, ce qui témoigne du bon repliement de la protéine HU, même après coupure de l'étiquette d'affinité.

3.4. À retenir

Objectif : Etablir une stratégie de production/purification des variants de HU	
Optimisation des étapes de clonage	
Etiquette 8His (Plasmides pET28a ou pJES)	Etiquette MBP (Plasmide pMAL-C5e)
Clonage par excision/ligation	Clonage par PCR à mega amorce
Site de clivage à l'enterokinase	Site de clivage à l'enterokinase
Optimisation des étapes de purification	
Chromatographie d'affinité	Chromatographie d'affinité (amylose)
Clivage étiquette 8His à l'enterokinase	Clivage étiquette MBP à l'enterokinase
Chromatographie d'affinité	Chromatographie échangeuse de cations (SP-FF)
Chromatographie échangeuse d'ions	Tamisaie moléculaire
	Chromatographie échangeuse de cations (20CM)
Pureté	
A homogénéité	A homogénéité
Rendement	
1mg HU / 10g cellules : rendement trop faible pour études structurales	15mg HU / 10g cellules : Bon rendement
Vérification du repliement	
Repliement confirmé par RMN	Repliement confirmé RMN
<p>Conclusion : Le protocole de production/purification optimisé avec la protéine de fusion HU-MBP (colonne de droite) est retenu pour la production des variants de HU.</p>	

Chapitre 4 : Caractérisation des structures natives

Le point de départ de la caractérisation du mécanisme d'échange est l'analyse structurale des dimères natifs de HU de *E. coli* (**Figure 45**) en comparaison avec d'autres structures de HU ou de ses homologues provenant d'autres organismes. On sait que certains homologues de HU sont capables d'échanger activement leurs chaînes (protéine TF1 du bactériophage λ),²¹³ alors que certaines protéines HU en sont incapables (HU de *Geobacillus stearothermophilus*).¹⁹⁶ On sait également que certaines protéines HU peuvent adopter une conformation intermédiaire (I_2) (HU de *Thermotoga maritima*)²¹⁴ alors que d'autres HU en sont incapables (HU de *Geobacillus stearothermophilus*).^{215,216} Le but est d'identifier les acides aminés essentiels pour la transition vers la conformation dite I_2 et pour l'échange de chaînes.

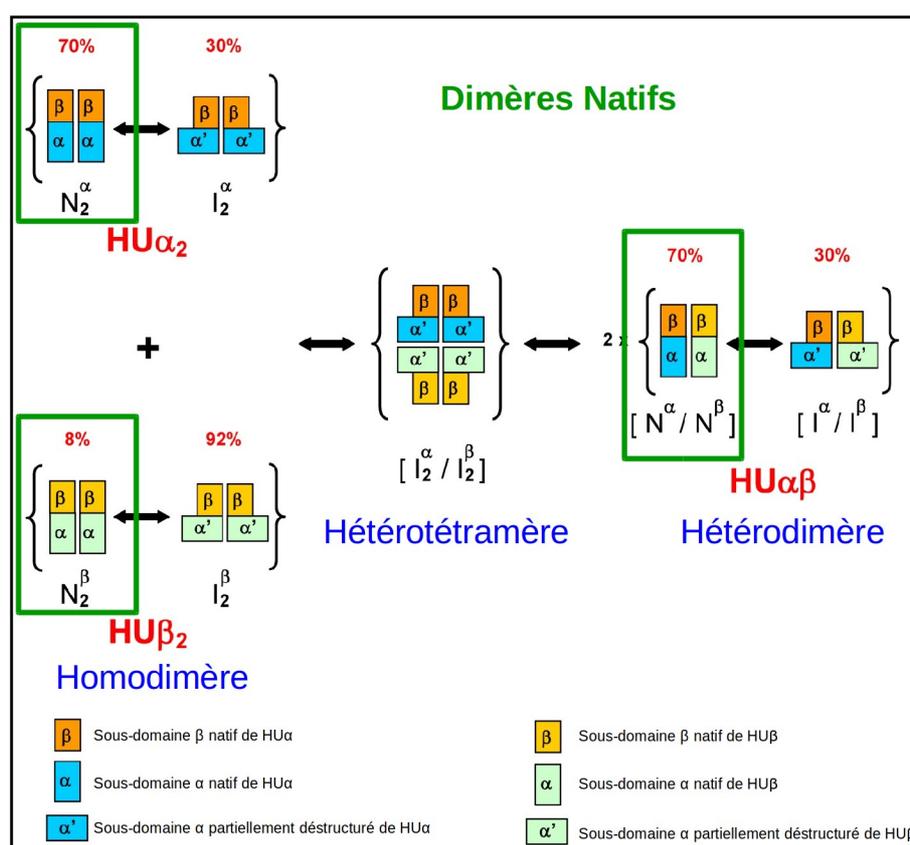


Figure 46: Dimères natifs (N_2) dans le mécanisme putatif d'échange de chaînes des dimères de HU

Pour cette analyse, j'ai exploité les données structurales acquises au cours de ma thèse (RMN et cristallographie) et les ai confronté avec celles disponibles dans la PDB (« Protein Data Bank » <http://www.rcsb.org/pdb/home/home.do>). Outre l'analyse des structures 3D, j'ai utilisé la RMN pour analyser la dynamique et la stabilité des homodimères afin d'identifier les régions qui auraient une propension à se déstructurer et à permettre de transiter vers la conformation I₂.

Chez *E. coli*, les structures natives de l'homodimère HU α_2 (code pdb : 1MUL, résolution 2,30Å) et de l'hétérodimère HU $\alpha\beta$ (code pdb : 2O97, résolution 2,45Å) ont déjà été résolues par diffraction des rayons X.^{2,3} La structure de l'homodimère HU β_2 est manquante. Hors cette structure est essentielle puisqu'il s'agit du dimère qui a le Tm₁ de transition de N₂ vers I₂ le plus bas (27°C). La résolution de sa structure permettra de compléter l'analyse structurale des dimères de HU de *E. coli* dans leur conformation native.

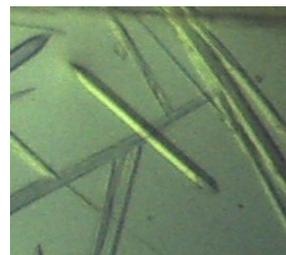
4.1. Structure cristallographique de HU β_2

Des tests de cristallogénèse ont été entrepris afin d'obtenir la structure cristallographique de HU β_2 sous sa conformation native et ainsi de compléter la collection des structures cristallographiques des HU d'*E. coli*. La cristallogénèse d'une protéine aussi dynamique que HU β_2 ,¹⁶⁷ une protéine liée aux chocs froids et la moins stable des HU d'*E. coli*, constituait un réel défi. Néanmoins, des cristaux ont été obtenus à température ambiante et ont diffracté à la plus haute résolution rapportée pour les cristaux de protéine de cette famille.

4.1.1. Cristallogénèse

La protéine concentrée à 20 mg/mL dans le tampon de cristallisation (10 mM HEPES pH 7,5 ; 200 mM NaCl) a été criblée contre les 96 conditions du kit JCSG-*plus*TM (Molecular Dimensions) (Voir composition dans les tableaux 21 et 22 en annexe p 307-308). La technique de cristallisation utilisée ici est la diffusion en phase vapeur comme décrit dans le chapitre 2 (matériels et méthodes p 70). Des gouttes assises de 200 nL ou 300 nL (ratio 1:1 et 2:1 protéine:précipitant) ont été réalisées à l'aide d'un robot de cristallisation mosquito[®] (ttplabtech) et incubées à 21°C. Les meilleurs cristaux ont été obtenus dans la condition 1.40 (0,2M Sulfate de lithium ; 0,1M Acétate de sodium pH4,5 ; 30 % PEG 8000) au ratio 1:1. Ces cristaux ont été extrêmement difficiles à reproduire après le premier criblage. On peut mettre en cause la faible stabilité de la protéine ou le changement de cinétique de diffusion de vapeur dans les essais en plus gros volumes (**Figure 47**).

Les cristaux du criblage initial ont été utilisés pour les expériences de diffraction des rayons X et ont permis d'atteindre une excellente résolution (1,25 Å). Cette résolution est la meilleure obtenue parmi les structures de HU déposées dans la PDB (1,60 à 3,02 Å).



4.1.2. Diffraction des rayons X et résolution de la structure cristallographique

Les données de diffraction ont été collectées jusqu'à 1,25 Å à l'ESRF (European Synchrotron Radiation Facility, Grenoble) sur la ligne de lumière ID23-1. Ces données nous ont permis de résoudre la structure cristallographique de HU β_2 par remplacement moléculaire en utilisant la structure de HU α_2 (code PDB 1MUL) comme modèle de recherche. HU β_2 a cristallisé dans le groupe d'espace P6 $_5$ 22 avec 1 monomère par unité asymétrique. Il est intéressant de noter que dans ce groupe d'espace, les protéines s'assemblent dans le cristal en formant un axe hélicoïdal. Dans le cas de HU β_2 , au centre de cet axe se forme un pore au niveau de la selle et des bras de la protéine d'environ 25 Å de diamètre (**Figure 48**). Ce pore pourrait être exploité pour réaliser du trempage de cristaux avec des ligands ciblant cette zone de la protéine comme ceux utilisés pour cibler la protéine HU de *M. tuberculosis*.²¹⁸

Figure 47: Cristaux de HU β_2

Obtenu dans les conditions 1.40 du kit JCSG+ (0,2 M Sulfate de lithium ; 0,1 M Acétate de sodium pH 4,5 ; 30 % PEG 8000). Dimensions approximatives : 0,2 x 0,01 x 0,01 mm

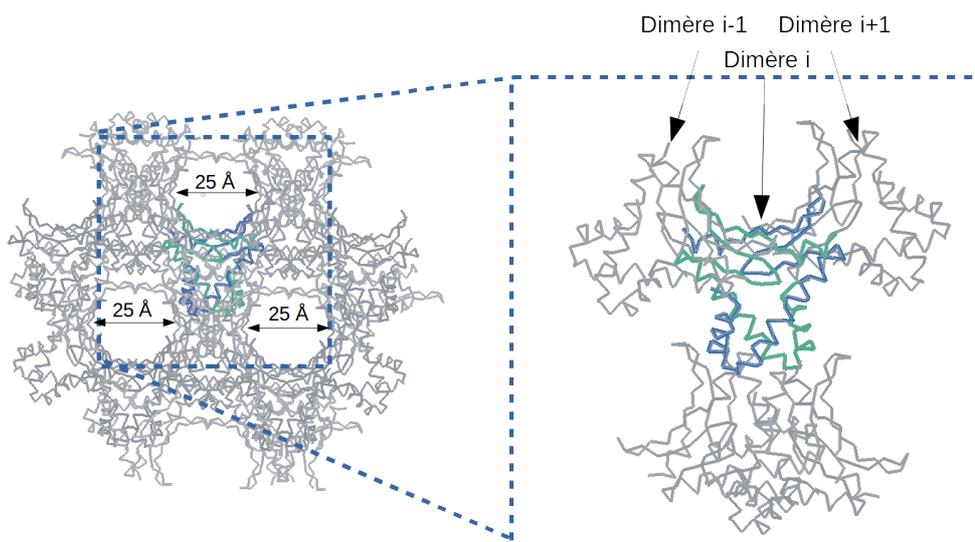


Figure 48: Empilement cristallin de HU β_2

Le modèle structural a été affiné jusqu'à des facteurs R de 15,6 % et R_{free} de 17,3 % dans la gamme de résolution de 44,26-1,25 Å (**Tableau 4**). A cette résolution, le placement des chaînes latérales est non ambigu en terme de rotamères et la contribution des hydrogènes dans la densité électronique est visible (**Figure 49 A**). Seules quelques chaînes latérales longues non contraintes à la surface de la protéine n'ont pas pu être construites entièrement (K53, E54, R55, K75 et R80). Par ailleurs, les résidus 56 à 73 sont manquants à cause de la flexibilité des bras. Le dimère biologique est obtenu par une symétrie de type C2 et est représenté **Figure 49 B**.

La structure de $\text{HU}\beta_2$ suit le repliement standard des protéines de la famille et est bien sûr très proches des autres structures de dimères de HU de *E. coli* avec un RMSD de 0,45 et 0,40 Å (sur les $\text{C}\alpha$) avec les structures de $\text{HU}\alpha_2$ et $\text{HU}\alpha\beta$, respectivement. Une analyse plus détaillée de la structure est faite en parallèle des autres structures disponibles plus loin dans ce chapitre.

Les coordonnées de $\text{HU}\beta_2$ ont été déposées dans la **PDB** sous le code **4P3V**.

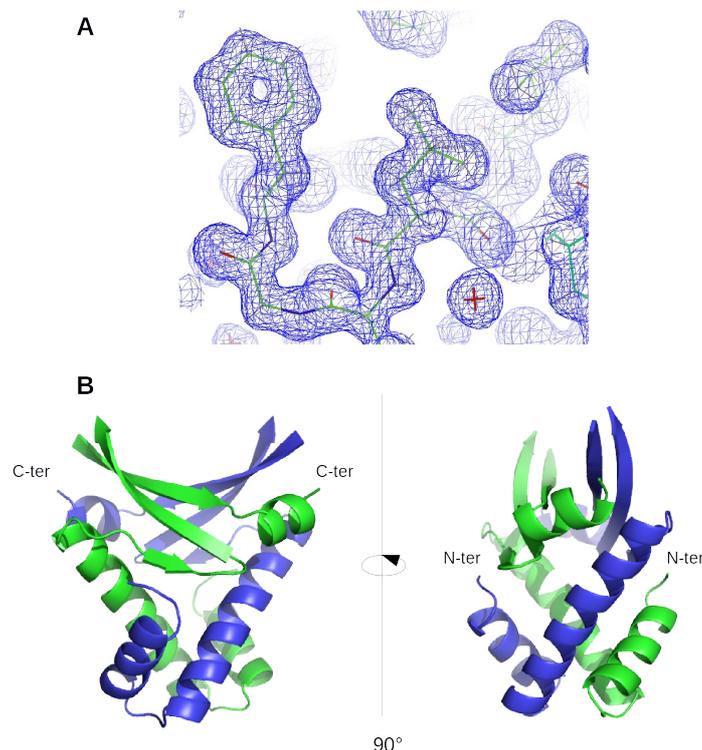


Figure 49: Carte de densité électronique (A) et structure cristallographique (B) de $\text{HU}\beta_2$. La carte de densité électronique ($2Fo - Fc$) est représentée avec un niveau de contour à 2σ dans la région du tour α entre le brin 1 et 2 du feuillet β .

Tableau 4: Diffraction et statistiques d'affinement de HU β ₂

Statistiques de la collecte des données	
Source de rayonnement	ESRF ID23-1
Longueur d'onde (Å)	0.98400
Groupe d'espace	<i>P</i> 6 ₅ 22
Dimensions de la maille: <i>a</i> = <i>b</i> , <i>c</i> (Å)	51.11, 110.50
Gamme de résolution (Å)	44.26-1.25 (1.27-1.25)
Observations totales	227333 (9489)
Réflexions uniques	24255 (1137)
Completude (%)	99.4 (97.9)
Multiplicité	9.4 (8.3)
<i>R</i> _{merge} ^a	0.045 (0.892)
<i>I</i> / σ moyen (<i>I</i>)	20.9 (2.1)
Affinement et statistiques du modèle	
Gamme de résolution (Å)	44.26-1.25
Nombre de réflexions utilisées	24213
<i>R</i> _{work} ^b / <i>R</i> _{free} ^c	0.156/0.173
Facteurs B moyen (Å ²)	
Tous les atomes	23.97
Atomes de la protéine	22.74
Atomes Li ⁺	19.21
Atomes H ₂ O	35.17
RMSD par rapport à l'idéal	
Longueur des liaisons (Å)	0.005
Angles des liaisons (°)	0.888
Angles dièdres (°)	8.959
Analyse Ramachandran (% of residues)	
Régions favorables / autorisés / interdites	98.6/1.4/0.0
Nombre d'atomes	
Protéine	1026
Li ⁺	1
eau	113

^a $R_{merge} = \frac{\sum_h \sum_i |I_{h,i} - \langle I \rangle_h|}{\sum_h \sum_i I_{h,i}}$ où $\langle I \rangle_h$ représente l'intensité moyenne des réflexions symétriques équivalentes.

^b $R_{work} = \frac{\sum_h ||F_o| - |F_c||}{\sum_h |F_o|}$, où F_o et F_c représentent les amplitudes des facteurs de structures observés et calculés respectivement pour les réflexions *h*.

^c R_{free} représente la valeur *R* extraite de 5 % des données de réflexion qui ne sont pas incluses dans le processus d'affinement.

4.2. La protéine HU de *Lactococcus lactis* (LIHU)

Deux raisons ont motivé notre choix d'étudier la protéine HU de *L. lactis* :

(i) Cette protéine constitue potentiellement un contrôle négatif d'échange de chaînes pour les dimères de *E. coli*. *L. lactis* étant une bactérie mésophile, ses propriétés thermodynamiques peuvent être proches de celles des HU de *E. coli*. Cependant, *L. lactis* ne possède qu'un seul gène codant pour un seul homodimère de HU. La structure primaire de la protéine HU de *L. lactis* est assez proche de celles des isoformes d'*E. coli* avec 54 % d'identité de séquence. Mais, comme nous le verrons dans l'analyse comparative des HU plus loin dans ce chapitre, elle possède certains acides aminés clefs retrouvés chez les HU de *Bacillus subtilis* et de *Geobacillus stearothermophilus* (qui ne possèdent pas de forme intermédiaire I₂), et pourrait permettre de vérifier l'importance de la nature de ces acides aminés dans l'échange de chaîne nécessaire à la formation de l'hétérodimère HU $\alpha\beta$ d'*E. coli*.

(ii) Le second intérêt de la protéine HU de cet organisme concerne l'étude de la stimulation de l'activité de la Formamidopyrimidine ADN glycosylase (Fpg) du système de réparation par excision de base (BER) par HU (Chapitre connexe). Les protéines Fpg d'*E. coli* et de *L. lactis* étaient disponibles au laboratoire et nous avons voulu comparer l'effet de HU sur cette enzyme dans les différents organismes.

4.2.1. Culture de la souche MG1363 de *Lactococcus lactis*

La souche MG1363 de L. lactis nous a été généreusement offerte par Philippe Gaudu (INRA Jouy-en-josas) et Anne Bravard (CEA Fontenay-aux-roses).

Le premier objectif consistait à obtenir des colonies isolées de *L. lactis*. Une attention particulière fut nécessaire afin d'éviter toute contamination puisque l'on travaille en absence de sélection par antibiotique. La manipulation est réalisée en parallèle d'un contrôle négatif sans bactérie pour vérifier une éventuelle contamination des milieux de culture.

Des bactéries stockées en gélose profonde ont permis d'ensemencer une culture de 10 mL de milieu M17 + 0,5% glucose (favorable à la culture des lactocoques). *L. lactis* étant une bactérie anaérobie aérotoleante, on incube la culture sur la nuit sans agitation (tube fermé) à 30°C (température optimale de croissance).

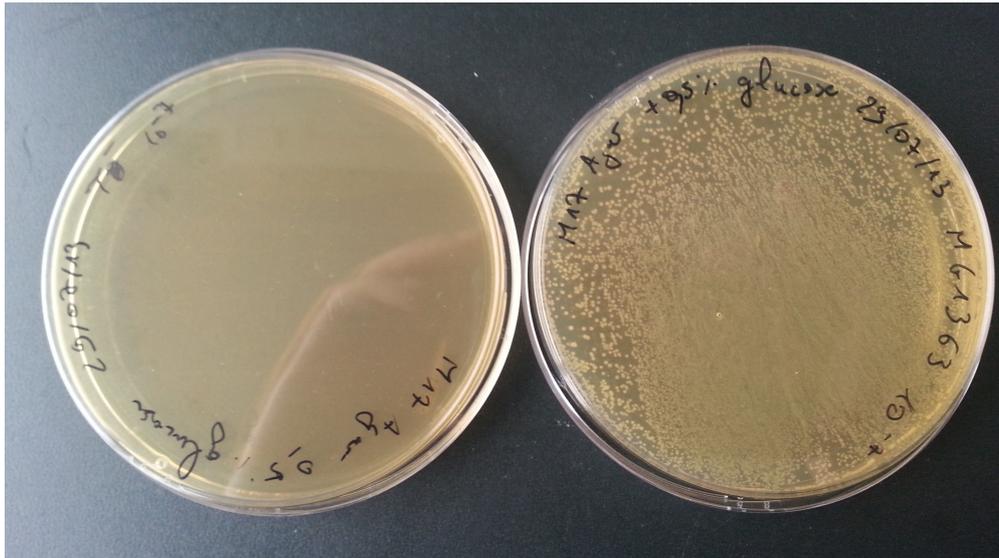


Figure 50: Colonies de *L.lactis* sur gélose Agar-M17-glucose-(0,5%) dépourvue de sélection A gauche, témoin négatif. A droite, colonies de *L.lactis*.

Les bactéries sédimentées sont resuspendues par une brève agitation puis 200 μL (dilué à 10^{-7}) de suspension bactérienne sont étalés sur une boîte de Pétri contenant une gélose Agar-M17-glucose (0,5%) incubée sur la nuit à 30°C. Le résultat est présenté **Figure 50**. Aucune contamination n'a été observée dans la boîte témoin et les colonies observées proviennent donc bien de la souche présente dans la gélose de départ. On peut passer ensuite à l'étape de clonage.

4.2.2. Clonage du gène *hup* de *L. lactis* dans le vecteur pMAL-c5e

Le gène *hup* (ou *hsla*) de *L. lactis* a été cloné dans le vecteur pMAL-c5e en couplant la méthode de clonage par PCR à méga amorce (**détaillée Figure 38 p 100**) et la PCR sur colonie. Le clonage est donc réalisé en une seule étape en resuspendant une colonie de *L.lactis* dans 50 μL de milieu LB. 5 μL de cette suspension seront utilisées dans une réaction PCR de 50 μL final. Les amorces n°15 et 16 (tableau en annexe) ont été utilisé pour ce clonage. Une première étape de dénaturation thermique longue (5min à 98°C) permet la lyse des bactéries et la libération de l'ADN génomique de *L.lactis* qui sert de

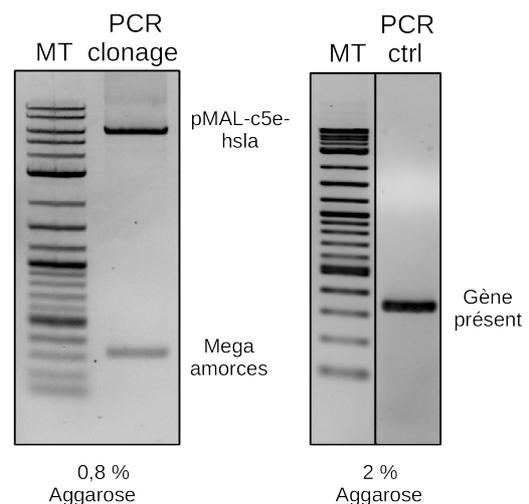


Figure 51: Contrôle de clonage par électrophorèses sur gels d'agarose

matrice. La suite de la réaction PCR se déroule normalement comme détaillé **Figure 38**. On peut contrôler l'efficacité de clonage par électrophorèse sur gel d'agarose (**Figure 51 à gauche**). La faible quantité de vecteur de départ (10 ng) ne permet pas de le détecter. En revanche, on observe que des méga-amorces sont générées ainsi que le vecteur cloné (p_{mal-c5e-hslA}). Après digestion DPNI pour éliminer le relicat de plasmide de départ, le produit de PCR est utilisée pour transformer des Bactéries NEB-10 β (New England Biolabs™) que l'on utilise pour réaliser des minipreps. La présence de l'insert dans le vecteur purifié est vérifiée par PCR (**Figure 51 à droite**) et par séquençage.

4.2.3. Production et purification de la protéine HU de *L.lactis*

La production et la purification ont été réalisées selon le protocole mis au point dans le **chapitre 3**. La protéine HU de *Lactococcus lactis* s'exprime et se purifie de manière analogue à celles de *E. coli* (**Figure 52**).

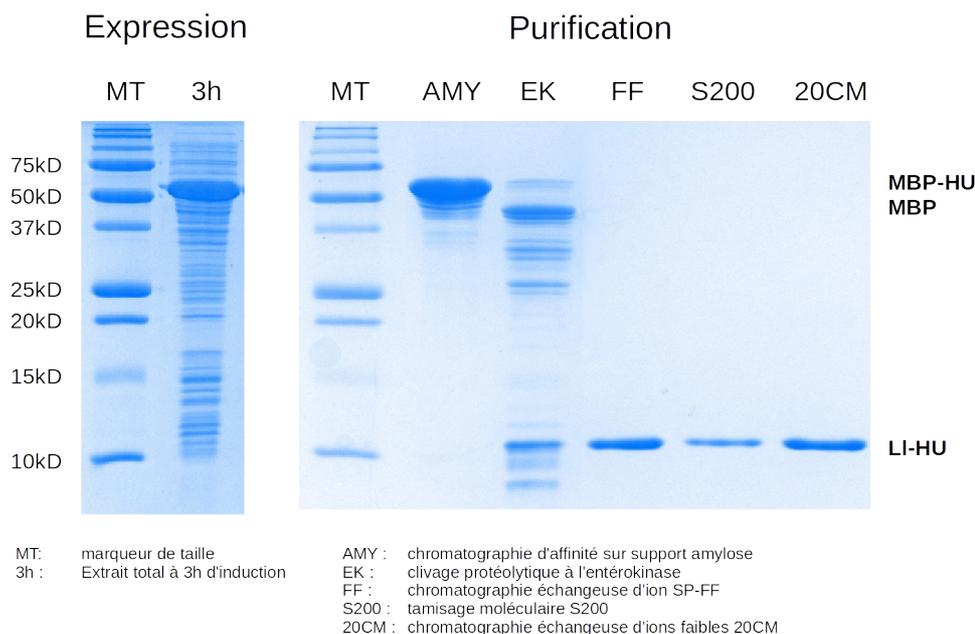


Figure 52: Contrôle de l'expression et de la purification de HU de *L.lactis* par SDS-PAGE

On peut cependant noter une différence par rapport à la purification des isoformes de *E. coli* au niveau de la dernière étape de purification (20CM). Cette chromatographie échangeuse d'ion faibles est très efficace dans la séparation des protéines HU entre elles. On peut séparer les trois dimères de *E.coli* sur ce support qui élue à environ 290, 370 et 410 mM NaCl (pour HU α_2 , HU $\alpha\beta$ et HU β_2 ,

respectivement). Dans le cas de HU *L.lactis*, la protéine est éluée à 450mM NaCl. Bien que les points isoélectriques soient très similaires entre ces quatre protéines (**Tableau 5**), cette différence d'éluion à haute force ionique est imputable à une densité de charge plus importante chez LIHU. Il est possible que cette différence de charge se traduise par une différence de fixation sur l'ADN de cette protéine.

Tableau 5: Comparatif du nombre de charges et du point isoélectrique des HU de *E.coli* et *L.lactis*

Protéine	Point isoélectrique	Nombre d'acides aminés basiques (Arg + Lys)	Nombre d'acides aminés acides (Asp + Glu)
EcHU α_2	9,57	28	20
EcHU β_2	9,69	28	20
EcHU $\alpha\beta$	9,63	28	20
LIHU	9,52	32	24

4.2.4. Cristallogénèse

La protéine purifiée est concentrée à 20 mg/mL dans une solution tampon minimale (10 mM HEPES pH 7,5 ; 200 mM NaCl). Dans un premier temps, on crible les conditions de cristallisation du kit Sigma™ 1 (Crystallization Basic Kit) en mélange 1 μ L/ 1 μ L (protéine / Agent précipitant). Sur les 50 conditions testées, 4 ont permis la formation de cristaux dans les 48h (conditions 6, 15, 20, et 40, **tableau 23 p 309 en annexe**). Les cristaux obtenus formaient des aiguilles extrêmement ramifiées et ont diffracté à des résolutions supérieures à 8 Å. Une 5^{ème} condition (condition 47) a permis d'obtenir des cristaux de grandes tailles sous forme de plaques hexagonales au bout de 5 jours (**Figure 53**). Contrairement à HU β_2 de *E. coli*, ces conditions étaient très reproductibles et témoignent d'une assez bonne stabilité de la protéine dans le temps.



Figure 53: Cristal de HU de *L.lactis* obtenu dans la condition 47 du kit de cristallogénèse Sigma 1 (0,1M Acetate de sodium pH4,6 ; 2M Sulfate d'ammonium).

Dimensions approximatives : 0,6 x 0,4 x 0,01 mm

4.2.5. Diffraction des rayons X et structure cristallographique préliminaire de LIHU

Les données de diffraction ont été collectées à 2,19 Å à l'ESRF (Grenoble) sur la ligne ID23-1. La protéine HU de *L.lactis* a cristallisé dans le groupe d'espace c222(1) avec 4 monomères par unité asymétrique (**Figure 54**). Comme pour HU β_2 , la structure a été résolue par remplacement moléculaire en utilisant la structure de HU α_2 de *E.coli* (code PDB : 1MUL) comme modèle de recherche. Les deux dimères présents dans l'unité asymétrique interagissent par leurs hélices α (**Figure 54**). Les bras sont stabilisés par des interactions inter-dimères dans la maille du cristal et ils sont observables pour chaque chaîne de l'unité asymétrique. Les chaînes latérales des résidus suivant sont incomplets : K20, E31, R56, R59, E60, K71, K81 et K90 sur la chaîne A ; Q5, K20, R62, Q65, K71, K84 et K91 sur la chaîne B ; K20, K54, E55, R62, Q65 et K71 sur la chaîne C ; E55, Q65, E68, K71, K81 et K91 sur la chaîne D.

Le modèle de LIHU est qualifié de préliminaire car les statistiques d'affinement ne sont pas satisfaisantes (notamment l'écart entre les valeurs de R et de $R_{free} > 5\%$). Cette structure est toutefois assez proche des structures de HU d'*E. coli* avec un RMSD de 0,62, 0,52 et 0,48 avec HU α_2 , HU $\alpha\beta$ et HU β_2 , respectivement.

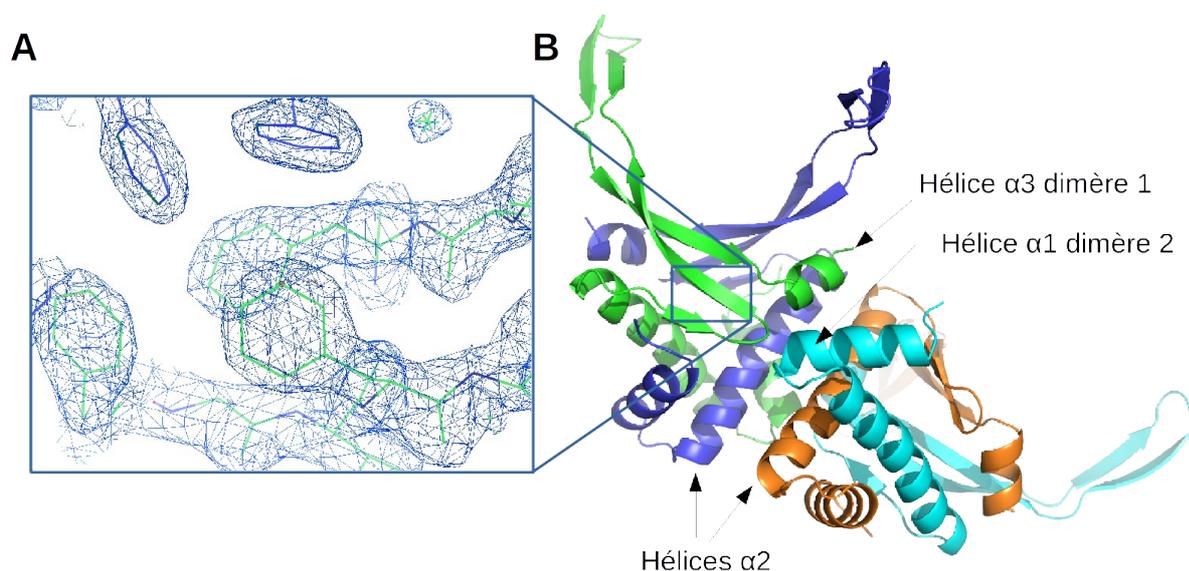


Figure 54: Carte de densité électronique (A) et structure cristallographique préliminaire (B) de LIHU A- Région sélectionnée de la carte de densité électronique ($2Fo - Fc$) représentée avec un niveau de contour à 2σ (région du cœur hydrophobe). B- Modèle en cours d'affinement de LIHU.

Tableau 6: Diffraction et statistiques d'affinement préliminaire de HU de *L.lactis*

Statistiques de la collecte des données	
Source de rayonnement	ESRF ID23-1
Longueur d'onde (Å)	0.98400
Groupe d'espace	C222 ₁
Dimensions de la maille: <i>a</i> , <i>b</i> , <i>c</i> (Å)	47.78, 99.00, 168.30
Gamme de résolution (Å)	47.49-2.19 (2.31-2.19)
Observations totales	133676 (18870)
Reflections uniques	20860 (2928)
Complétude (%)	99.6 (97.9)
multiplicité	6.4 (6.4)
R_{merge}^a	0.090 (0.858)
I/σ moyen (<i>I</i>)	13.5 (2.2)
Affinement préliminaire et statistiques des modèles	
Gamme de résolution (Å)	47.49-2.19
Nombre de réflexions utilisées	20822
$R_{\text{work}}^b / R_{\text{free}}^c$	0.1876/0.2469
Facteurs B moyens (Å ²)	
Tous atomes	53.97
Atomes de la chaîne A	50.45
Atomes de la chaîne B	53.25
Atomes de la chaîne C	52.91
Atomes de la chaîne D	53.48
Atomes de SO ₄ ²⁻	113.95
Atomes de H ₂ O	52.47
RMSD par rapport à l'idéal	
Longueur des liaisons (Å)	0.008
Angles des liaisons (°)	0.929
Angles dièdres (°)	14.674
Analyse Ramachandran (% de résidus)	
Régions favorables / autorisées / interdites	97.2/2.8/0.0
Nombre d'atomes	
Chaîne protéique A	645
Chaîne protéique B	645
Chaîne protéique C	641
Chaîne protéique D	661
Molécules de SO ₄ ²⁻	65
Molécules d'H ₂ O	103

^a $R_{\text{merge}} = \frac{\sum_h \sum_i |I_{h,i} - \langle I \rangle_h|}{\sum_h \sum_i I_{h,i}}$ où $\langle I \rangle_h$ représente l'intensité moyenne des réflexions symétriques équivalentes.

^b $R_{\text{work}} = \frac{\sum_h ||F_o| - |F_c||}{\sum_h |F_o|}$, où F_o et F_c représentent les amplitudes des facteurs de structures observés et calculés respectivement pour les réflexions *h*.

^c R_{free} représente la valeur R extraite de 5 % des données de réflexion qui ne sont pas inclus dans le processus d'affinement.

4.3. Analyse comparative des structures natives de HU

4.3.1. Criblage de la PDB

Pour collecter les structures de HU et de ses homologues dans la pdb, on utilise le serveur Dali²¹⁹(http://ekhidna.biocenter.helsinki.fi/dali_server/). Ce service permet de rechercher automatiquement des structures dans la PDB (Protein Data Bank) qui présentent un repliement proche de la structure de référence, ce qui est bien plus efficace qu'une recherche par mots clefs. En utilisant la structure de HU α_2 de *E.coli* (code pdb : 1MUL) comme référence, 23 structures ont été identifiées avec une similarité significative (dans notre cas, Z-score > 6) (**Tableau 7**).

Parmi ces 23 structures, 12 appartiennent à des protéines HU provenant de 6 organismes bactériens différents. On couvre cinq phyla différents : *Actinobacteria* (*M.tuberculosis*), *Cyanobacteria* (*Anabaena sp*), *Thermotogae* (*T.maritima*), *Enterobacteria* (*E.coli*) et *Firmicutes* (*G.stearothermophilus* et *B.anthraxis*). 7 structures correspondent à IHF chez *E.coli* uniquement et 2 à TF1 du bacteriophage SPO1. La multiplicité des structures de protéines provenant du même organisme est souvent due à des formes mutantes ou des complexes avec différents ADN. L'homologie structurale entre HU, IHF et TF1 a déjà été décrite dans le chapitre introduction et il n'est pas étonnant de sortir ces structures lors de notre criblage.

Deux autres protéines présentant une homologie structurale significative sont sorties de ce criblage : Hbb de *B.burgdorferi* (bactérie du phylum spirochaetes responsable de la maladie de Lyme) et BVU_2165 de *B.vulgatus* (bactérie du phylum Bacteroidetes). Hbb est la seule protéine de la famille structurale de HU/IHF identifiée dans le génome de *B.burgdorferi*. Elle est capable de compléter les souches d'*E.coli* dépourvues des gènes codant pour IHF.²²⁰ Cependant, contrairement à IHF qui est obligatoirement hétérodimérique, Hbb est homodimérique comme la plupart des HU. Elle présente une identité de séquence assez similaire pour HU (α : 26,7 % ; β : 33,3 %) et pour IHF (α : 28,3 % ; β : 28,6 %) d'*E.coli*, HU β étant la plus proche. La structure de Hbb reste très similaire aux HU et IHF, à la différence près que la première hélice est un peu plus longue et s'écarte davantage du corps de la protéine (**Figure 55**). La courbure de l'ADN en U induite par la protéine est similaire à celles observées avec HU et IHF (**Figure 55**). Il est difficile de catégoriser Hbb comme une IHF ou une HU. Il est possible que la toute petite taille du génome de *B.burgdorferi* (~1000 kbp) le contraigne à limiter le nombre de gènes redondants ou similaires, et que Hbb assure des fonctions multiples.

Tableau 7: Structures de HU et de ses homologues disponibles dans la pdb M : méthode, R : résolution, RX : Diffraction des rayons X.

Protéine	Organisme	Variant	Code PDB	Z-score	RMSD 1MUL	Commentaire	M	R (Å)
HU	<i>E.coli</i>	HU α_2	1MUL	14.4	0.0		RX	2,30
		HU $\alpha\beta$	2O97	11.7	0.7		RX	2,45
	<i>G. stearothermophilus</i>	HU	1HUU	12.5	0.8		RX	2,00
			1HUE	9.3	2.1		RMN	
	<i>B. anthracis</i>	HU	3RHI	11.6	0.8		RX	2,48
	<i>T. maritima</i>	HU	1RIY	11.5	0.8	Sauvage Mutant V42I	RX	1,80
1B8Z			10.7	1.0	RX		1,60	
<i>Anabaena sp</i>	HU	1P71	11.0	1.0	Complexes avec 3 ADN différents	RX	1,90	
		1P51	10.9	1.0		RX	2,50	
		1P78	10.8	1.0		RX	2,25	
<i>M. tuberculosis</i>	HU	4PT4	10.4	1.0	PH 8.2 PH 6.7 + MnCl ₂	RX	2,04	
		4DKY	10.4	1.1		RX	2,48	
IHF	<i>E. coli</i>	IHF $\alpha\beta$	2HT0	9.9	1.6	ADN double nick	RX	2,00
			1IHF	9.9	1.6	ADN simple nick	RX	2,50
			1OWF	9.7	1.4	β -E44A+ADN	RX	1,95
			1OWG	9.7	1.4	Wt+ADN muté	RX	2,10
			1OUZ	9.6	1.4	β -E44A+ADN muté	RX	2,41
			2IIE	7.3	1.4	chimère $\alpha\beta$	RX	2,41
			2IIF	7.2	1.4	chimère $\alpha\beta$ α K45E	RX	2,72
TF1	Bacteriophage SPO1	TF1	1WTU	6.2	3.5	Sauvage mutant E15G T32I	RMN	
			1EXE	6.9	2.7		RMN	
Hbb	<i>B. burgdorferi</i>	Hbb	2NP2	9.4	1.1	IHF ou HU ?	RX	3,02
BVU_2 165	<i>B. vulgatus</i>	BVU_2 165	4FMR	7.3	2.4	Fonction inconnue	RX	2,25

La deuxième protéine non classée, BVU_2165, provenant de l'organisme *B.vulgatus*, n'a pas de rôle identifié. La structure a été résolue par la plateforme à haut débit JCSG (« Joint Group for Structural Genomics »). Cette structure homodimérique de 56 kDa présente un motif HU qui semble servir de domaine de dimérisation (**Figure 55 en bleu et vert**). Des extensions N- et C-terminales de grandes tailles forment des domaines supplémentaires Ig-like (**Figure 55 en cyan et vert clair**). Ces domaines masquent partiellement les côtés de la « selle » de fixation de l'ADN du domaine HU. Si cette protéine est capable de se lier à l'ADN, le mode de fixation doit être différent. La torsion en U de l'ADN induite par la protéine devrait être prohibée par la présence de ces domaines Ig-like ou nécessiterait un changement de conformation important.

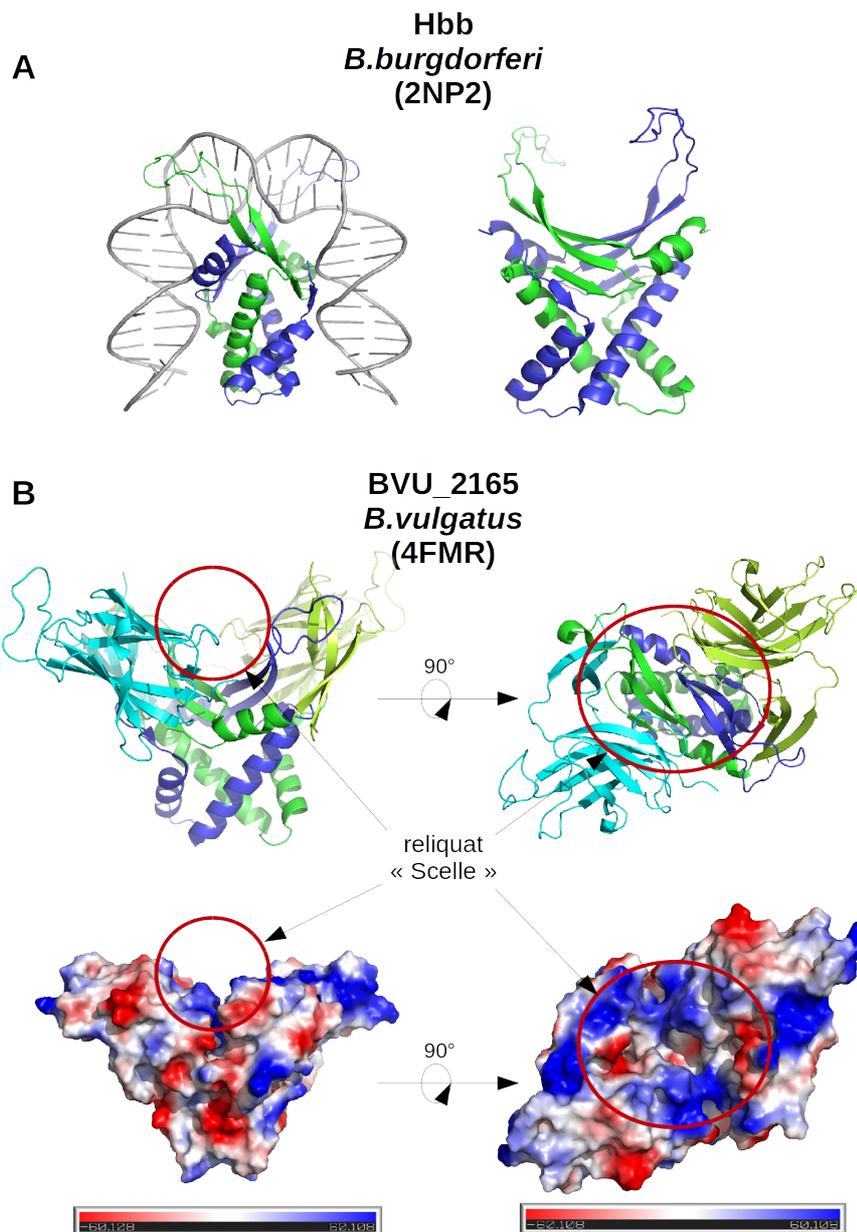


Figure 55: Structures d'homologues de HU A- Hbb de *B. burgdorferi*. B- BVU_2165 de *B. vulgatus*. Représentation « cartoon » (en haut) et représentation des surfaces électrostatiques (en bas)

D'autre part, les prolines d'intercalation des bras en β -hairpin ne sont pas conservées dans cette protéine. Toutefois, la région comportant le reliquat de la scelle flanquée des deux sandwichs β présente une surface globalement basique qui pourrait constituer un site de fixation à l'ADN (**Figure 55**). Une telle structure montre que malgré des extensions de grandes tailles, une dimérisation semble toujours possible. C'est ce que nous avons également pu observer avec les protéines fusionnées aux étiquettes d'affinité (MBP, 8HIS).

Plusieurs structures sont disponibles pour la plupart des protéines de chaque organisme issues de ce criblage. Ces redondances sont dues à des mutations, à des complexes avec différentes séquences d'ADN, des conditions différentes ou encore des méthodes de résolution de structure différentes pour les mêmes protéines. Pour simplifier l'analyse à suivre, nous avons choisi une structure de référence pour chaque dimère de HU de chaque espèce en privilégiant les protéines libres dépourvues de mutations.

4.3.2. Analyse des surfaces d'interactions des dimères

L'outil PISA (Protein Interfaces Surfaces and Assemblies)²²¹ a initialement été développé pour l'identification des oligomères dans les structures cristallographiques. Ce logiciel propose une analyse approfondie et automatique des paramètres d'interaction protéine-protéine tels que la surface d'interaction, la différence d'énergie libre de solvation entre le dimère et les monomères, ou encore le nombre de liaisons hydrogènes et de ponts salins que nous avons utilisés pour comparer les structures de HU et de ses homologues dans les différentes espèces.

Tableau 8: Comparaison des surfaces d'interaction des dimères de HU, IHF et TF1 Δ^iG représente la différence d'énergie libre de solvation entre les monomères et le dimère. Cette énergie ne tient pas compte des liaisons hydrogène et ponts salins à l'interface. $n(LH)_i$ et $n(PS)_i$ représentent le nombre de liaisons hydrogène inter-chaînes et de ponts salins inter-chaînes, respectivement.

Protéine	Organisme	PDB	Surface d'interaction (Å ²)	Δ^iG (kcal/mol)	$n(LH)_i$	$n(PS)_i$
HU α_2	<i>E.coli</i>	1MUL	1745,1	-36,3	8	2
HU $\alpha\beta$	<i>E.coli</i>	2O97	1787,6	-36,1	7	3
HU β_2	<i>E.coli</i>	4P3V	1803,9	-31,6	8	2
HU	<i>L.lactis</i>	-	2037,6	-45,5	4	1
	<i>G.stearothermophilus</i>	1HUU	1778,6	-40,7	4	0
	<i>M.tuberculosis</i>	4PT4	2050,6	-38	6	2
	<i>B.anthraxis</i>	3RHI	1784,5	-35,6	7	1
	<i>Anabaena sp</i>	1P71	1978,3	-40	7	2
	<i>T.maritima</i>	1RIY	1885,6	-40,6	4	0
IHF	<i>E.coli</i>	1IHF	2421,1	-36,4	16	6
TF1	SPO1	1WTU	2796,6	-45,3	6	0

L'analyse des interfaces entre les monomères dans les dimères de HU, IHF et TF1 révèle des différences assez subtiles. Les HU ont toutes une surface d'interaction assez importante, de l'ordre de 2000 Å², qui représente environ un tiers de la surface totale d'un monomère (**Tableau 8**). La différence d'énergie libre de solvatation des dimères par rapport aux monomères (Δ^iG) est assez cohérente pour les dimères de HU d'*E.coli* avec HU β_2 la moins stable et HU α_2 et HU $\alpha\beta$ un peu plus stables. Ces résultats sont à mettre en parallèle avec les différences de Tm₁ présentées dans l'introduction (**Tableau 2 p 53**). Le Δ^iG présenté ici ne tient pas compte des liaisons hydrogène ou des ponts salins et est plutôt révélateur de l'affinité entre des surfaces hydrophobes des chaînes du dimère. La plupart des autres HU présentent des Δ^iG plus bas que ceux des HU d'*E.coli*. C'est le cas notamment du HU de *G.stearothermophilus* qui ne présente pas de forme I₂ ni de mécanisme actif d'échange comme pour les dimères de *E. coli*.¹⁷⁷⁻¹⁷⁹

Du point de vue des interactions polaires, le nombre de liaisons hydrogène et de ponts salins semble avoir un impact mineur sur la stabilité des dimères, que ce soit pour la forme I₂ (Tm₁) ou la stabilité globale des dimères (Tm₂). La protéine HU de *G.stearothermophilus*, qui représente la HU la plus stable de notre comparatif, possède le plus petit nombre de liaisons hydrogène inter-chaînes, et aucun pont salin inter-chaîne. Ces quatre liaisons hydrogène, entre la partie N-terminale et le brin 1 du feuillet β ne sont pas spécifiques à HU de *G.stearothermophilus* et sont retrouvés chez la plupart de ses homologues (**Figure 56**). Par conséquent, on ne peut établir de corrélations directes entre le nombre ou la position de liaison hydrogène et la présence d'une forme intermédiaire avec un Tm plus ou moins élevé. C'est cohérent avec les observations sur l'effet du NaCl sur les Tm des protéines HU d'*E. coli* par microcalorimétrie (**Tableau 2 p 53**). En effet, l'augmentation de la concentration en NaCl, bien qu'il puisse constituer un compétiteur de ponts salins, stabilise les dimères de HU en augmentant à la fois les Tm₁ et Tm₂. Ceci peut s'expliquer par une augmentation de la polarité du solvant par le NaCl, exacerbant les effets hydrophobes et stabilisant le dimère en inhibant la dissociation et la formation de l'intermédiaire. On peut en déduire que la nature des interactions stabilisant la forme intermédiaire et la forme native possède une composante hydrophobe majoritaire. On peut toutefois signaler que l'abondance des interactions polaires inter-chaîne peut être à l'origine des préférences entre les chaînes des dimères. Dans le cas de HU $\alpha\beta$, qui est le dimère de HU majoritaire chez *E.coli*, il y a un pont salin supplémentaire qui n'existe pas chez les homodimères et qui doit participer à la reconnaissance spécifique d'une chaîne avec l'autre. C'est encore plus prononcé pour IHF qui présente le nombre le plus important de liaisons polaires inter-chaînes avec 16 liaisons hydrogène et 6 ponts salins.

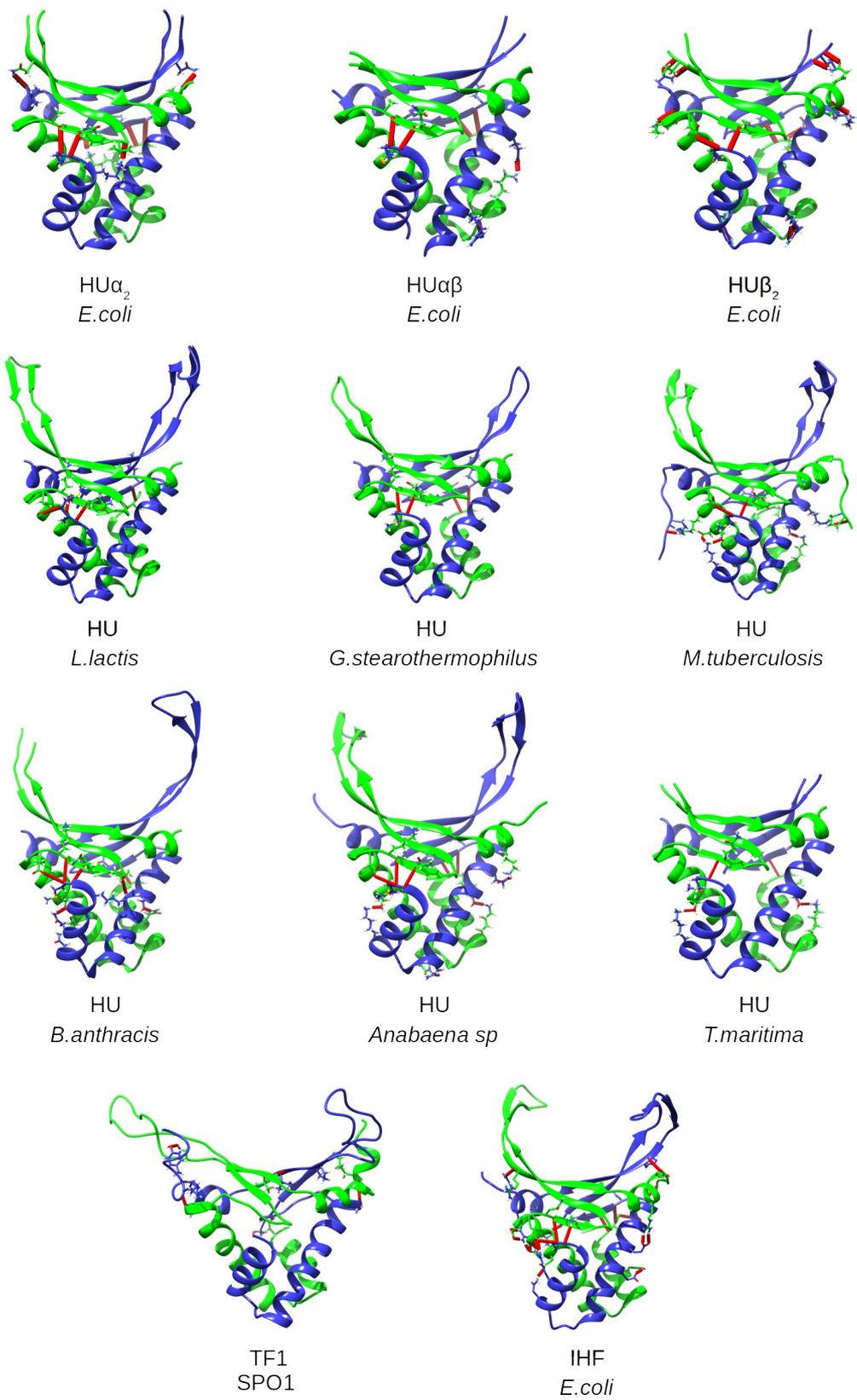


Figure 56: Interactions polaires entre les chaînes des dimères de HU, IHF et TF1

Cette reconnaissance extrêmement spécifique peut expliquer le caractère hétérodimérique obligatoire de IHF. TF1 paraît un peu atypique comparé aux autres dimères. En effet, sachant que TF1 est capable d'échanger ses chaînes, on attendait un profil similaire aux HU d'*E. coli*. Toutefois sa valeur très basse de Δ^iG doit être rapportée à une surface d'interaction plus étendue ($\sim 2800 \text{ \AA}^2$) et n'exclut pas de déstabilisation locale permettant la formation d'un état intermédiaire et donc l'échange de chaînes.

4.3.3. Analyses des cavités hydrophobes des dimères

Le rôle prépondérant des interactions hydrophobes dans la stabilité de l'état intermédiaire et du dimère en général nous a conduit à étudier le cœur hydrophobe de ces protéines. Les structures de différentes HU ont permis de mettre en évidence une cavité hydrophobe située entre le sous-domaine α du HTH et le sous domaine β . Pour étudier les tailles de ces cavités et les corrélérer à la stabilité connue de chaque dimère, nous avons utilisé l'outil CASTp (<http://sts-fw.bioengr.uic.edu/castp/index.php>)¹⁸⁰⁻¹⁸². Le volume et la surface des cavités sont répertoriés dans le **Tableau 9** et représentés graphiquement dans la **Figure 57**. Deux cavités peuvent être mises en évidence pour certaines structures et sont dénommées cavité 1 (**en vert figure 57**) et cavité 2 (**en magenta figure 57**).

Tableau 9: Surface et volume des cavités hydrophobes des dimères de HU, IHF et TF1

Protéine	Organisme	PDB	Cavité 1		Cavité 2	
			Surface (\AA^2)	Volume (\AA^3)	Surface (\AA^2)	Volume (\AA^3)
HU α_2	<i>E.coli</i>	1MUL	157,4	138,3	91,1	54,7
HU $\alpha\beta$	<i>E.coli</i>	2O97	162,6	168,2	-	-
HU β_2	<i>E.coli</i>	4P3V	272,7	256,6	-	-
HU	<i>L.lactis</i>	-	-	-	98,9	64,2
	<i>G.stearothermophilus</i>	1HUU	62,8	39,1	84,1	48,0
	<i>M.tuberculosis</i>	4PT4	165,9	229,5	-	-
	<i>B.anthraxis</i>	3RHI	254,0	230,3	-	-
	<i>Anabaena sp</i>	1P71	170,3	152,4	-	129,5
	<i>T.maritima</i>	1RIY	189,3	185,5	-	-
IHF	<i>E.coli</i>	1IHF	72,0	52,2	-	-
TF1	SPO1	1WTU	219,4	163,7	-	-

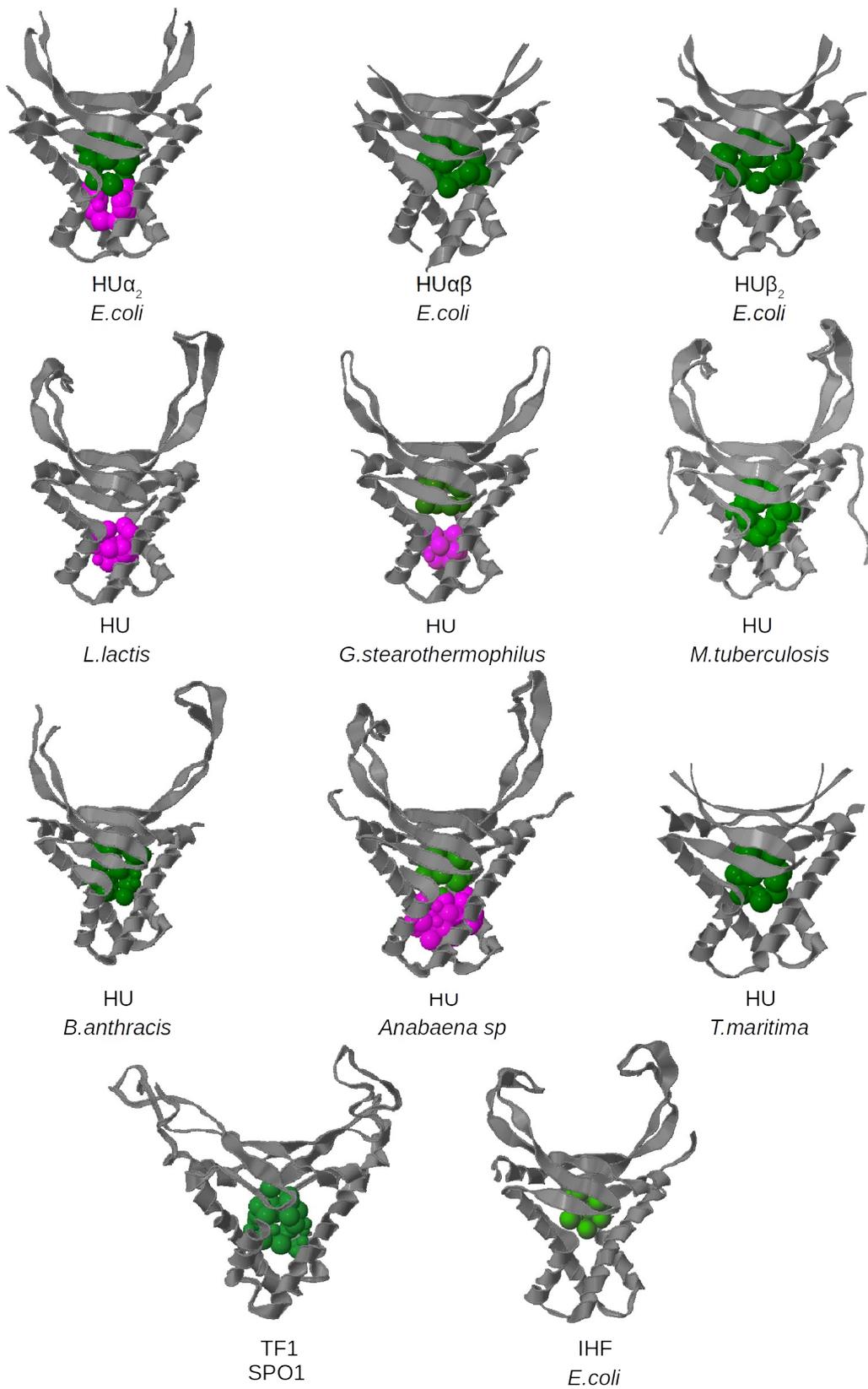


Figure 57: Cavités hydrophobes des dimères de HU, IHF et TF1 (cavité 1 en vert et cavité 2 en magenta)

On distingue différents groupes de structures en fonction des types et des tailles de cavités hydrophobes observées. Si l'on prend en compte les structures des dimères qui échangent de façon active ou qui possèdent un état intermédiaire stable (les trois HU de *E.coli*, HU de *T.maritima* et TF1) on a toujours une cavité 1 (en vert) de grande taille ($>100 \text{ \AA}^3$). Dans le cas des HU de *E.coli*, on observe également une corrélation inverse entre la taille de la cavité (ainsi que le Δ^iG) et le Tm_1 (**Tableau 10**). A l'inverse, pour les dimères qui n'échangent pas activement (HU de *G.stearothermophilus*, IHF) la cavité 1 est très petite ($\sim 50 \text{ \AA}^3$) et les dimères sont très stables. La cavité 2 n'est pas détectée dans chaque structure et est souvent relativement petite (sauf pour *Anabaena*). A ce stade, nous ignorons si la présence de cette cavité 2 a un impact sur la stabilité ou l'échange des dimères.

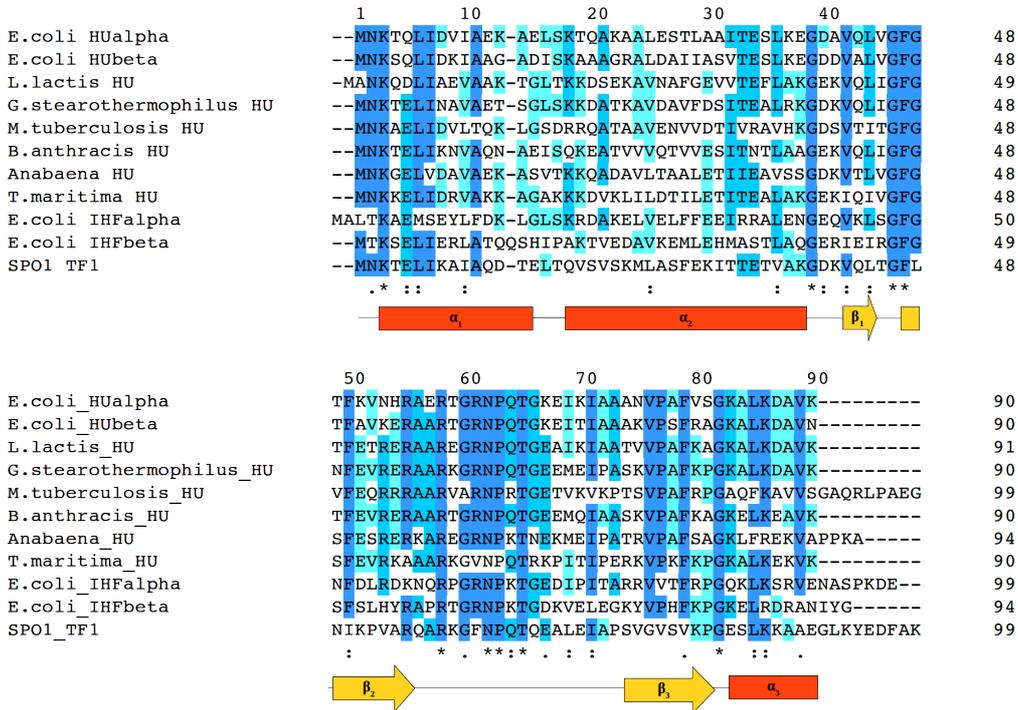
Tableau 10: Corrélation entre le Tm_1 ($N_2 \rightarrow I_2$) et le volume de la cavité 1 et l'énergie de solvation Δ^iG représente la différence d'énergie libre de solvation entre les monomères et le dimère. Cette énergie ne tient pas compte des liaisons hydrogène et ponts salins à l'interface.

Protéine	HU β_2 <i>E. coli</i>	HU $\alpha\beta$ <i>E. coli</i>	HU α_2 <i>E. coli</i>	HU <i>G. stearothermophilus</i>
Tm_1 (°C)	27	38	41	aucun
V(cavité1) (\AA^3)	256,6	168,3	138,3	39,1
Δ^iG (kcal/mol)	-31,6	-36,1	-36,3	-40,7

La cavité 1 se trouvant à l'interface entre les sous-domaines hélices (HTH) et feuillets, on peut faire l'hypothèse qu'une cavité de petite taille stabilise le dimère et contribue à la coopérativité de la dénaturation ($N_2 \rightarrow 2D$, sans intermédiaire). Dans le cas d'une cavité 1 de grande taille, on diminue la stabilité globale de la protéine et on sépare physiquement les cœurs hydrophobes des deux sous-domaines. On peut ainsi dénaturer l'un des sous-domaines puis l'autre de manière découplée et indépendante ($N_2 \rightarrow I_2 \rightarrow 2D$). En suivant cette hypothèse, on doit pouvoir éliminer l'existence de l'état intermédiaire en comblant cette cavité. Pour cela, on peut réaliser de la mutagenèse dirigée en ciblant les résidus hydrophobes bordant cette cavité et en augmentant leur taille. La **Figure 58** présente l'alignement des séquences primaires des structures étudiées. La région la moins conservée est localisée dans l'hélice α_2 qui doit être à l'origine des différences observées entre les dimères (**Figure 58 A**). Si l'on se concentre sur les résidus qui bordent les cavités (**Figure 58 B**), les

mutations semblent de faible impact (I → V, L → V, I → F ...). Mais même si on conserve souvent le caractère hydrophobe, la longueur et l'encombrement des chaînes latérales bordant la cavité varient. Ces changements d'acides aminés se répercutent sur la taille des cavités observées et probablement sur la stabilité des dimères.²²² Pour les dimères qui possèdent les cavités 1 les plus petites (ou non détectables), HU de *L.lactis* et *G.stearothermophilus*, on remarque la présence d'une phénylalanine supplémentaire dans l'hélice α_2 . L'encombrement généré par l'ajout des 2 noyaux aromatiques au sein des dimères de HU pourrait être suffisant pour combler les poches des HU d'*E.coli* et supprimer leur état intermédiaire. Pour vérifier cette hypothèse, il serait intéressant de produire le mutant L29F/I29F de chaque dimère HU d'*E.coli* et de déterminer l'effet de cette mutation sur leur T_{m1} . Malheureusement, si la mutagenèse et la production de ces mutants ont été réalisées, nous n'avons pas pu aller jusqu'à l'analyse de ces mutants au moment de la rédaction de cette thèse et ces hypothèses ne seront explorées que dans des études ultérieures.

A- Conservation



B- Résidus bordant les poches

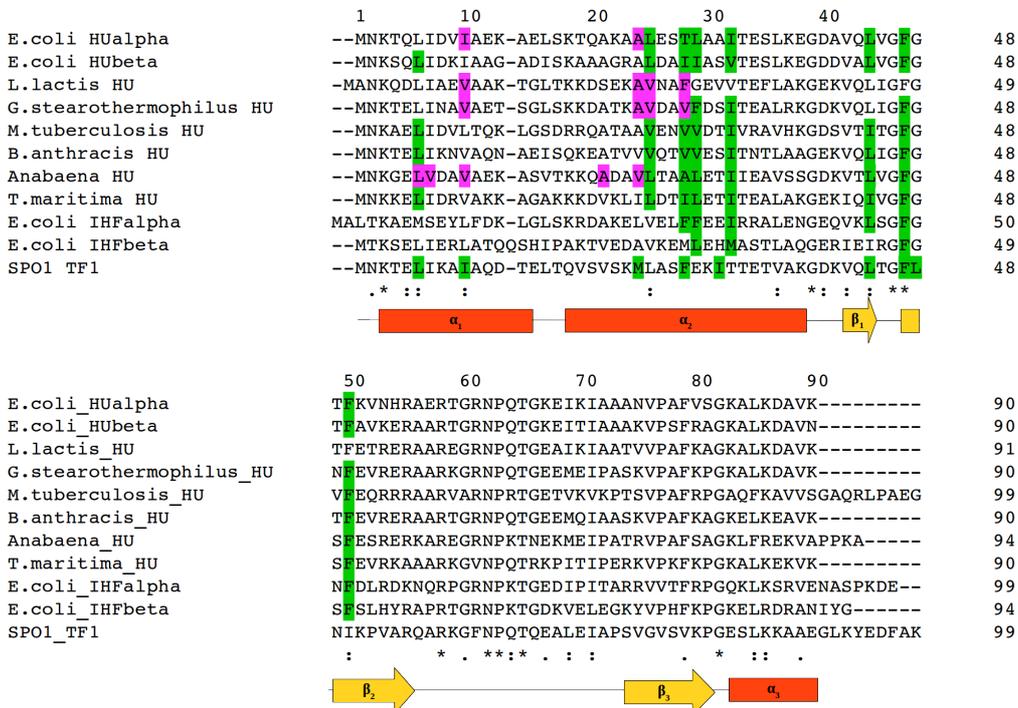


Figure 58: Alignement des séquences primaires des structures de HU, IHF et TF1 A- Pourcentage d'identité de séquence. B- Résidus bordant la cavité 1 (en vert) et la cavité 2 (en magenta).

4.4. RMN : Attribution des dimères de HU de *E. coli* et caractérisation de l'homodimère HU α_2 natif (N $_2$)

Même si les structures cristallographiques de N $_2$ des trois dimères de HU sont maintenant connues, il est nécessaire de caractériser ces mêmes structures en solution (RMN). C'est en effet en solution que nous pourrions étudier par la suite la transition N $_2 \rightarrow I_2$, la structure dynamique de I $_2$ étant inaccessible par cristallographie. Par ailleurs, en plus des informations apportées par le **calcul de la structure en solution**, la RMN permet d'obtenir des informations sur la dynamique interne (**R $_1$, R $_2$, NOE**) ou encore d'évaluer la présence et de la stabilité des liaisons hydrogène du squelette peptidique (**calcul des coefficients de température des H $_N$**). Pour notre problématique, on cherche à évaluer la stabilité et la dynamique des homodimères à l'état natif (N $_2$), afin de mettre en évidence les acides aminés impliqués dans la transition vers I $_2$ et l'échange de chaînes nécessaires à la formation de l'hétérodimère.

Dans le cas de transitions conformationnelles, les changements structuraux modifient l'environnement chimique et électronique des atomes et entraînent une perturbation de leur **déplacement chimique**. Cette perturbation peut se traduire de différentes façons sur les spectres en fonction de la vitesse d'échange (k) et de la différence de fréquence de résonance entre les états ($\Delta\omega$). Si $k \ll \Delta\omega$, la vitesse d'échange est lente, le nombre de pics pour chaque atome est multiplié par le nombre d'état, chacun ayant un déplacement chimique propre à son environnement chimique local dans une conformation donnée. L'intensité relative de chaque pic est pondérée par la population de chaque état. Dans le cas où $k \gg \Delta\omega$, la vitesse d'échange est rapide, on observe alors un seul pic pour chaque atome ayant un déplacement chimique moyenné entre ceux des conformations explorées. Le déplacement chimique moyen est pondéré par la population de chaque état.

Afin de pouvoir réaliser toutes ces analyses, la première étape a consisté à attribuer les trois dimères de HU dans leur conformation native (N $_2$).

4.4.1. Attribution des dimères de HU dans leur conformation native

Un à deux échantillons doublement marqués $^{15}\text{N}/^{13}\text{C}$ et un échantillon non marqué (300 μL avec : 1 mM HU ; 200 mM NaCl ; 20 mM NaP pH 7,6 ; 0,05 mM DTT) ont été constitués pour chaque dimère afin de collecter les spectres nécessaires à leur attribution et/ou à la résolution de

structure et l'analyse de la dynamique interne. L'avancement de l'enregistrement des expériences RMN est décrit dans le **tableau 11**.

Tableau 11: Enregistrement des expériences RMN sur les dimères de HU de *E.coli* « √ »:
 expérience réalisée ; « - » expérience non réalisée ; « NA » expérience non applicable

	Expériences \ Protéine	HU α_2 N $_2$	HU β_2 N $_2$	HU $\alpha^*\beta$ N $_2$	HU $\alpha\beta^*$ N $_2$
	^{15}N -HSQC	√	√	√	√
	^{13}C -HSQC	√	√	√	√
Squelette peptidique	HNCO	√	√	√	√
	HNCA	√	√	√	√
	HN(CO)CA	√	√	√	√
	HN(CA)CO	√	√	√	√
	HNCACB	√	√	√	√
	CBCA(CO)NH	√	√	√	√
	Cc(CO)NH	√	√	√	√
Chaînes Latérales	HBHA(CO)NH	√	√	√	√
	HBHANH	√	√	√	√
	H(CcCO)NH	√	√	√	√
	^{15}N -HSQC-TOCSY	√	√	√	√
	HCCH-TOCSY	√	√	√	√
NOESY (structure)	^{13}C -HSQC-NOESY	√	√	-	-
	^{15}N -HSQC-NOESY	√	√	-	-
	^1H -NOESY	√	√	-	-
	NOESY Filtrés	NA	NA	√	√
Relaxation	R1, R2, NOE	√	√	-	-
Effet Température (15°C à 51°C) pas = 3°C	^{15}N -HSQC	√	√	-	-
	^1H -NOESY	√	√	-	-

Ces expériences nous ont permis d'attribuer les déplacements chimiques des atomes du squelette peptidique de chaque espèce dans leur conformation native. Le taux d'attribution de chaque dimère est présenté dans le **tableau 12**. Ce travail a fait l'objet d'un article soumis au journal « Biomolecular NMR Assignment ».

Tableau 12: Attribution des dimères de HU dans leur conformation native

Attribution		HU α_2 N ₂	HU β_2 N ₂	HU $\alpha^*\beta$ N ₂	HU $\alpha\beta^*$ N ₂
Code BRMB*		25328	25329	25330	25330
Squelette peptidique	Totaux	96 %	97 %	95 %	98 %
Chaînes latérales	H	91 %	93 %	-	-
	Non-H	69 %	69 %	-	-
NOEs	¹ H-NOESY	1610	1861	-	-
	¹⁵ N-HSQC-NOESY	776	0	-	-
	¹³ C-HSQC-NOESY	1143	1204	-	-

*Biological Magnetic Resonance Bank

4.4.2. Article I : Attribution du squelette peptidique des trois dimères de HU à 293K :
ECHU α_2 , ECHU β_2 et ECHU $\alpha\beta$

Backbone Assignment of the three dimers of HU from *Escherichia coli* at 293 K: EcHU α_2 , EcHU β_2 and EcHU $\alpha\beta$

Rémy Le Meur^{1#}, Karine Loth^{1#*}, Françoise Culard¹, Bertrand Castaing¹ and Céline Landon¹

Soumis à Biomolecular NMR Assignments

1 : Centre de Biophysique Moléculaire, Centre National de la Recherche Scientifique UPR 4301, Université d'Orléans, rue Charles Sadron, F-45071 Orléans Cedex 2, France.

: Equivalent contribution to the work

* Corresponding Author : Karine Loth

Address : Centre de Biophysique Moléculaire, CNRS UPR 4301, rue Charles Sadron, F45071 Orléans Cedex 2, France.

Mail : karine.loth@cnrs-orleans.fr

Tel : +33 2 38 25 55 71

Fax : +33 2 38 69 70 31

Abstract

HU is one of the major nucleoid-associated proteins involved in bacterial chromosome structure and in all DNA-dependent cellular activities. Similarly to eukaryotic histones, this small dimeric basic protein wraps DNA in a non-sequence specific manner, promoting DNA super-structures. In most bacteria, HU is a homodimeric protein encoded by a single gene. However, in enterobacteria such as *Escherichia coli*, the presence of two genes coding for two peptidic chains, HU α and HU β , lead to the coexistence of three forms : two homodimers EcHU α_2 and EcHU β_2 , as well as a heterodimer EcHU $\alpha\beta$. Genetic and biochemical investigation suggest that each EcHU dimer plays a specific physiological role in bacteria.. Their relative abundance depends on the environmental conditions and is driven by an essential, yet unknown, fast outstanding chain-exchange mechanism at physiological temperature. Our goal is to understand this fundamental mechanism from a structural and kinetics standpoint using NMR. For this purpose, the first steps are the assignment of each dimer in their native and intermediate states. Here, we report the backbone assignment of each HU dimers from *E. coli* at 293 K in their native state.

Keywords : HU, Histone-like, dimers, *E.coli*

Biological Context

The histone-like HU protein is the major nucleoid-associated protein involved in the dynamics and structure of the bacterial "chromosome". It is a small dimeric basic protein (19 kDa, pI = 10.5) and, similarly to eukaryotic histones, it wraps DNA in a non-sequence specific manner, promoting DNA super-structures (Castaing et al 1995; Pinson et al 1999; Vitoc and Mukerji 2011). Because of its peculiar DNA binding properties, HU is involved in various physiological and cellular processes, such as replication, initiation, DNA protection against radiation, site-specific recombination, as well as DNA transcription modulation in response to heat and cold shocks, anaerobiosis, acid stress, high osmolarity, and DNA damage signaling and repair. HU protein is essential for the survival and the growth of bacteria.

HU peptidic chains are highly conserved in bacteria. Whereas HU is present in most bacteria as a single homodimer, in enterobacteria such as *E. coli*, HU is present as two homodimers EchU α_2 , EchU β_2 , and a heterodimer EchU $\alpha\beta$. In *E. coli*, the *hupA* and *hupB* genes code for the EchU α and EchU β polypeptide chains (70% of identity, 90% of homology – Top of Figure 1), respectively. Genetic studies and DNA binding properties of HU species suggest that each EchU dimer has a different biological purpose (Pinson et al 1999; Oberto et al 2009; Prieto et al 2012). For example, EchU β_2 is required for *E. coli* cold shock adaptation (Giangrossi et al 2002). Indeed, the relative abundance of the different dimers depends on the environmental conditions and is driven by an essential, yet unknown, fast chain-exchange mechanism at physiological temperature (Claret and Rouviere-Yaniv 1997).

We have previously shown (Ramstein et al 2003; Garnier et al 2011) that, unlike the other HU proteins studied so far, which melt through a single step ($N_2 \leftrightarrow 2D$), the *E. coli* dimers follow a two-step thermal unfolding ($N_2 \leftrightarrow I_2 \rightarrow 2D$) in which the native dimer, N_2 , partially melts into a dimeric intermediate, I_2 , which in turn yields the unfolded monomers, D . The partially unfolded dimer I_2 is unable to complex with high-affinity, single-stranded break-containing DNA. These results suggest that the $N_2 \leftrightarrow I_2$ equilibrium plays a central role in the physiology of the *E. coli* HU proteins. Moreover, the intermediate state I_2 might be required in for heterodimer formation from homodimers.

Our objective is to understand this fundamental chain-exchange mechanism from a structural and

kinetics standpoint using NMR. For this purpose, assignment of each dimer in their native and intermediate state is necessary. Here, we report the backbone assignment of each HU dimers from *E. coli* at 293 K in their native state.

Methods and experiments

NMR sample preparation

EcHU α_2 and EcHU β_2 homodimers production and purification (in natural abundance and ^{15}N - ^{13}C labelled) has been performed as previously reported (Coste et al 1999; Ramstein et al 2003; Garnier et al 2011).

EcHU $\alpha^*\beta$ heterodimer (*: ^{15}N - ^{13}C labelled) was prepared by mixing the ^{15}N - ^{13}C labelled EcHU α_2 homodimer with an excess of the unlabelled EcHU β_2 homodimer. Symmetrically, EcHU $\alpha\beta^*$ heterodimer was prepared by mixing EcHU β_2^* and an excess of EcHU α_2 .

NMR spectroscopy

NMR samples contained 0.5 mM of dimers in 95 % H $_2$ O/5 % D $_2$ O at pH 7.5 in 20 mM phosphate buffer with 200 mM NaCl and 0,05 mM DTT. NMR spectra were acquired at 293K using a 600 MHz Varian ^{UNITY}INOVA spectrometer equipped with a non-cryogenic pentaprobe or a Bruker 700 MHz spectrometer equipped with a TCI CryoProbe. Sequential assignments were carried out using ^{15}N -HSQC, HNCOC, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, following classical procedures. NMR data were processed using NMRPipe software (Delaglio et al 1995) or Bruker's Topspin 3.2TM and analyzed with CcpNmr software (Vranken et al 2005).

Assignment and data deposition

Figure 1 shows the ^1H - ^{15}N HSQC spectra of EchU α_2 homodimer, EchU β_2 homodimer, EchU $\alpha^*\beta$ and EchU $\alpha\beta^*$ heterodimers.

EchU α_2 homodimer

Chemical shift assignments were made for protein backbone resonances: H^N (98.9 %), N (98.9 %, proline residues excluded), CO (93.3 %), C α (100 %) and C β (96.4 %).

EchU β_2 homodimer

Chemical shift assignments were made for protein backbone resonances: H^N (97.7 %), N (97.7 %, proline residues excluded), CO (99 %), C α (100 %) and C β (97.6 %).

EchU $\alpha\beta$ heterodimer

Chemical shift assignments were made for protein backbone resonances: H^N (95.4 % and 97.7 % for the α and β chain respectively), N (95.4 % and 97.7 % for the α and β chain respectively, proline residues excluded), CO (94.5 % and 94.5 % for the α and β chain respectively), C α (99.8 % and 100 % for the α and β chain respectively) and C β (81.9 % and 95.1 % for the α and β chain respectively).

The secondary structure elements of each dimer have been predicted using the ^{13}C chemical shift index (Wishart and Sykes 1994) and the DANGLE (Cheung et al 2010) programs. Results are schematically summarized in **Figure 2**. As expected, the fold of each monomer consists of a helix-turn-helix motif (α_1 : 3-13, α_2 : 18-37), an antiparallel β -sheet (β_1 : 42-44, β_2 : 48-55, β_3 : 74-81) and a C-terminal α -helix (α_3 : 83-88). The β -sheet includes a very flexible part connecting β_2 to β_3 and a β -hairpin structure β_2 :(58-61) - β_3 :(68-71).

For each EchU α or EchU β chain, the H^N-N chemical shift perturbations between their homodimeric and heterodimeric states have been determined and are showed at the bottom of **Figure 2**. In this case, CSPs are only due to the differences between EchU α and EchU β sequences, at the proximity to contact points between the chains. The chemical shifts assignments for the EchU α_2 homodimer, the EchU β_2 homodimer and the EchU $\alpha\beta$ heterodimer have been deposited in the BioMagResBank

(<http://www.brmb.wisc.edu>) under the accession number 25328, 25329 and 25330 respectively.

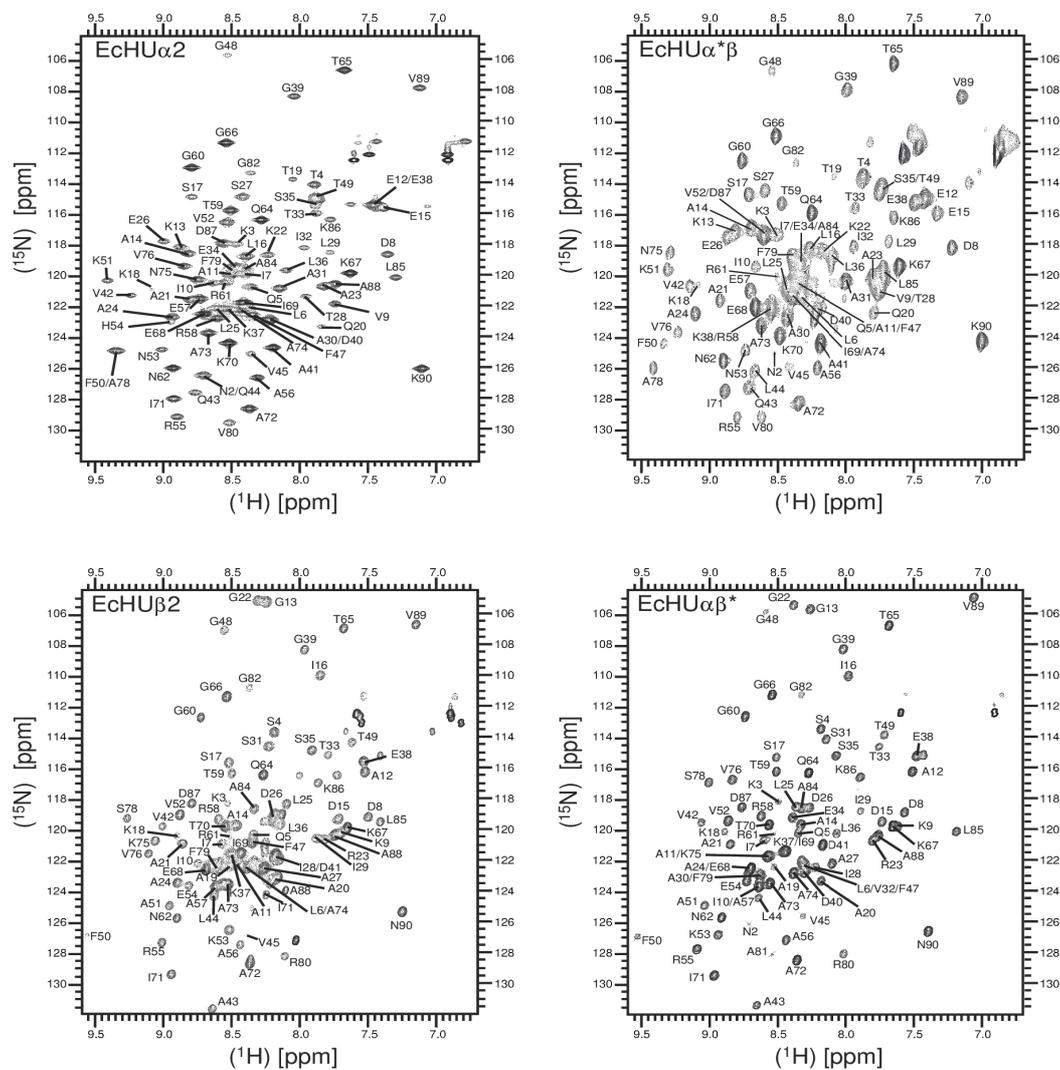


Figure 1: ¹H-¹⁵N-HSQC spectra of the three EchU dimers. Left, homodimers (Top : EchU α_2 – Bottom : EchU β_2). Right, heterodimer (Top : EchU $\alpha\beta$ – Bottom : EchU $\alpha\beta^*$).

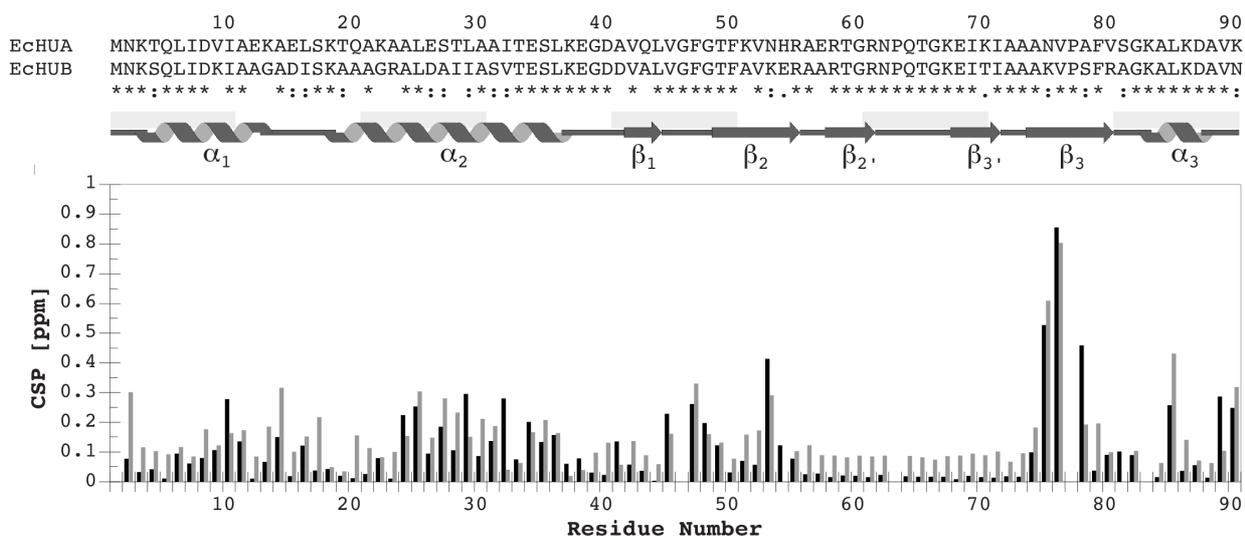


Figure 2: Chemical Shift Perturbation of HU depending on its quaternary structure

Top – Sequence alignment of EchU α and EchU β chains (Identity and homology are indicated by * and : respectively) and schematic representation of secondary structure elements from CSI and DANGLE. Bottom – H^N -N Chemical Shifts perturbations between the EchU α and EchU β chains in both their homodimeric and heterodimeric states (EchU α in grey bars and EchU β in black bars).

References

Castaing B, Zelwer C, Laval J, Boiteux S (1995) HU protein of Escherichia coli binds specifically to DNA that contains single-strand breaks or gaps. *J Biol Chem* 270:10291–6.

Cheung M-S, Maguire ML, Stevens TJ, Broadhurst RW (2010) DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson* 202:223–33. doi: 10.1016/j.jmr.2009.11.008

Claret L, Rouviere-Yaniv J (1997) Variation in HU composition during growth of Escherichia coli: the heterodimer is required for long term survival. *J Mol Biol* 273:93–104. doi: 10.1006/jmbi.1997.1310

Coste F, Hervouet N, Oberto J, et al (1999) Crystallization and preliminary X-ray diffraction analysis of the homodimeric form α_2 of the HU protein from Escherichia coli. *Acta Crystallogr Sect D Biol Crystallogr* 55:1952–1954. doi: 10.1107/S0907444999011361

Delaglio F, Grzesiek S, Vuister GW, et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–93.

Garnier N, Loth K, Coste F, et al (2011) An alternative flexible conformation of the *E. coli* HU β 2 protein: structural, dynamics, and functional aspects. *Eur Biophys J* 40:117–29. doi: 10.1007/s00249-010-0630-y

Giangrossi M, Giuliadori AM, Gualerzi CO, Pon CL (2002) Selective expression of the beta-subunit of nucleoid-associated protein HU during cold shock in *Escherichia coli*. *Mol Microbiol* 44:205–16.

Oberto J, Nabti S, Jooste V, et al (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One* 4:e4367. doi: 10.1371/journal.pone.0004367

Pinson V, Takahashi M, Rouviere-Yaniv J (1999) Differential binding of the *Escherichia coli* HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *J Mol Biol* 287:485–97. doi: 10.1006/jmbi.1999.2631

Prieto AI, Kahramanoglou C, Ali RM, et al (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res* 40:3524–37. doi: 10.1093/nar/gkr1236

Ramstein J, Hervouet N, Coste F, et al (2003) Evidence of a Thermal Unfolding Dimeric Intermediate for the *Escherichia coli* Histone-like HU Proteins: Thermodynamics and Structure. *J Mol Biol* 331:101–121. doi: 10.1016/S0022-2836(03)00725-3

Vitoc CI, Mukerji I (2011) HU binding to a DNA four-way junction probed by Förster resonance energy transfer. *Biochemistry* 50:1432–41. doi: 10.1021/bi1007589

Vranken WF, Boucher W, Stevens TJ, et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687–96. doi: 10.1002/prot.20449

Wishart DS, Sykes BD (1994) The ¹³C chemical-shift index: a simple method for the identification of protein secondary structure using ¹³C chemical-shift data. *J Biomol NMR* 4:171–80.

Les squelettes peptidiques des chaînes α et β étant attribués dans des contextes homodimériques et hétérodimériques, on peut calculer la perturbation de déplacement chimique (CSP) provoqué par le changement de chaîne dans les dimères (**Article I : Figure 2**). La structure globale étant conservée entre les dimères de HU, les changements d'environnements chimiques sont principalement dus aux mutations entre α et β au niveau des points de contacts entre les chaînes. Ces perturbations de déplacement chimique peuvent ainsi mettre en évidence les points de contacts interchaînes de manière expérimentale et aider l'attribution des NOE par la suite.

4.4.3. Structures des homodimères en solution (en cours d'affinement)

Les atomes des chaînes latérales et les spectres NOESY ont également été attribués pour les homodimères, HU α_2 et HU β_2 (**Tableau 12 p 135**). Des structures en solution des conformations natives des homodimères ont été calculées mais nécessitent encore quelques étapes d'affinement pour obtenir des statistiques satisfaisantes, en particulier au niveau du nombre de violations des NOEs (**Tableau 13**), et déposer les coordonnées à la PDB. Cependant, elles témoignent de notre capacité à calculer un modèle de dimère symétrique où les chaînes entrelacées compliquent considérablement la différenciation des NOE intra- et inter-chaînes. Les structures actuelles sont déjà assez proches des structures cristallographiques obtenues précédemment (**Figure 59**). En effet, le repliement global est conservé et le RMSD entre les structures cristallographiques et les structures préliminaires en solution est de 1 Å pour HU α_2 et 2 Å pour HU β_2 au niveau du squelette peptidique du corps (résidus 1-55 et 75-90). Les bras sont dynamiques et explorent un espace conformationnel important. Leur apparente coordination est due au fait que le calcul de structure est réalisé, à ce stade, en imposant la symétrie d'ordre C2. La dynamique importante des bras explique qu'ils sont invisibles dans la plupart des structures cristallographiques (comme celles d'*E. coli*). Pour celles où ils sont visibles (*L. lactis*, *M. tuberculosis*, *B. anthracis* etc), les bras sont impliqués soit dans l'interaction avec l'ADN ou soit dans des contacts inter-dimères au sein de la maille cristalline.

L'obtention des modèles N₂ actuels est très encourageant dans l'optique du calcul des structures des conformations intermédiaires I₂ en solution. En effet, dans le cas où une partie de la structure serait conservée entre les conformations N₂ et I₂ (notre hypothèse), l'attribution correspondante pourra être directement transposée d'une forme à l'autre et facilitera le calcul des structures des formes intermédiaires.

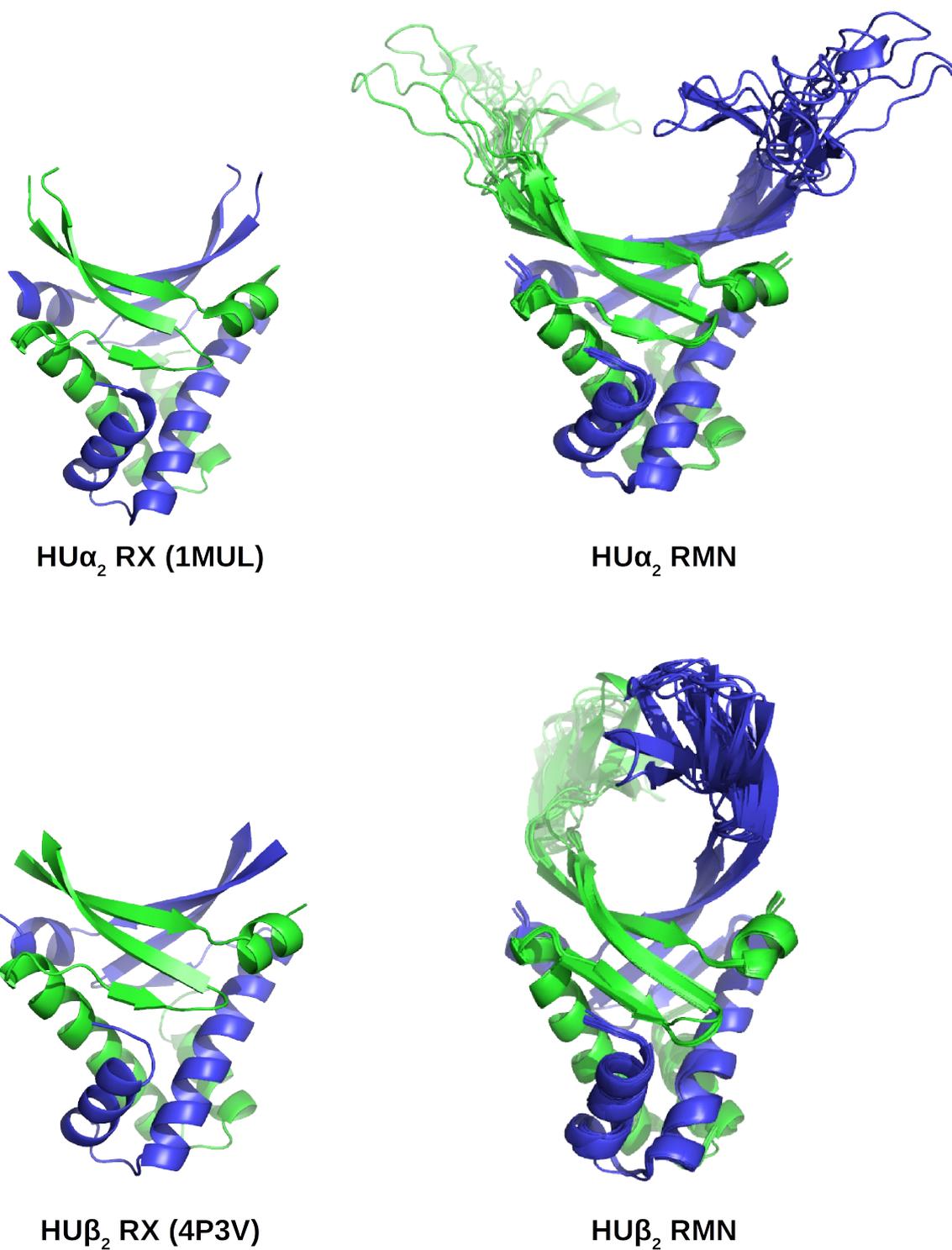


Figure 59: Comparaison des structures RMN préliminaires avec les structures cristallographiques (RX) des homodimères de HU

Tableau 13: Contraintes RMN et statistiques structurales

	HU α_2	HU β_2
Contraintes RMN		
Contraintes de distance par chaîne		
NOE totaux	1638	1846
Liaisons hydrogène	0	101
Contraintes d'angle dièdres	168	163
Statistiques structurales		
	Modèles 1-10	Modèles 1-10
Violations moyennes par structure		
NOEs ³ 0.3 Å	1,2%	10%
Liaisons hydrogène ³ 0.5 Å	0	0
Angles dièdres ³ 10°	3,6%	9,2%
RMSD moyen (Å) (1-55 75-90)	0,3	0,24
Diagramme de Ramachandran		
Régions les plus favorables	87,6%	77,9%
Régions favorables	11,9%	21,9%
Régions autorisées	0,5%	0,2%
Régions interdites	0%	0%

4.4.4. Relaxation des homodimères

Les expériences de relaxation (R_1 , R_2 , NOE) permettent de suivre la dynamique interne des molécules à l'échelle ps-ns. Ce type d'analyse a déjà été réalisé sur l'homodimère $HU\beta_2^{167}$ et peut être comparée aux résultats obtenus avec l'homodimère $HU\alpha_2$ (**Figure 53**).

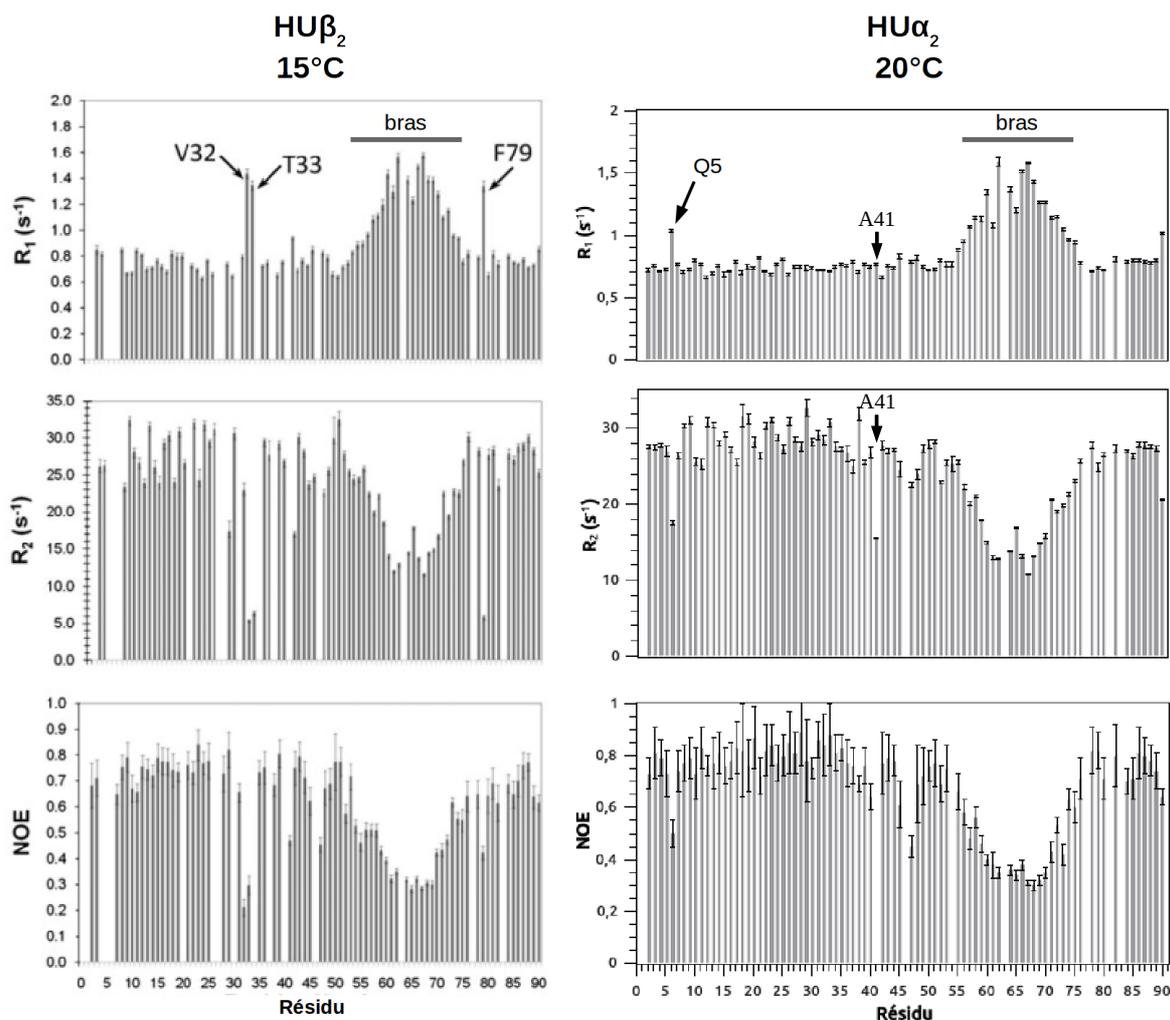


Figure 60: Relaxation ^{15}N R_1 , R_2 et NOE du squelette peptidique des homodimères de HU de *E. coli*. Les valeurs inférieures de NOE et R_2 indiquent des régions de flexibilité accrue à l'échelle ps-ns et inversement pour R_1 .

Pour $HU\beta_2$, les expériences ont été enregistrées à 15°C et présentent des profils relativement plats sauf pour les résidus du bras et les résidus V32, T33 et F79. La dynamique importante des bras est compatible avec l'exploration conformationnelle importante mise en évidence sur les structures préliminaires en solution (**Figure 59**). La dynamique du résidu 79 a été expliquée par des

expériences de simulation de dynamique moléculaire.¹⁶⁷ Dans ces expériences la phénylalanine 79 alterne entre une orientation enfouie dans le cœur hydrophobe et une orientation exposée au solvant ce qui est cohérent avec la dynamique importante observée en RMN. De même la dynamique des résidus V32 et T33 est également observée en dynamique moléculaire où un kink de l'hélice α_2 est observable au cours des simulations.

Dans le cas de $HU\alpha_2$, il faut travailler à 20°C pour obtenir des valeurs moyennes de R_1 , R_2 et NOE du même ordre que pour $HU\beta_2$. $HU\alpha_2$ est globalement plus rigide que $HU\beta_2$. Au niveau des bras, on a une situation similaire entre les deux homodimères avec une augmentation de la dynamique pour les résidus correspondants. Cependant, on ne retrouve pas, à cette échelle de temps, de dynamique plus importante pour les résidus I32, T33 et F79. La valine 32 de $HU\beta_2$ est remplacée par une isoleucine dans $HU\alpha_2$. Or une isoleucine a une bien meilleure propension à stabiliser une hélice α qu'une valine.²²³ Le remplacement V32I est donc stabilisant pour l'hélice α_2 et permet d'expliquer le niveau normal de dynamique interne des résidus 32 et 33. Par ailleurs, les valeurs mesurées pour la phénylalanine 79 sont dans la moyenne et nous portent à croire que le changement d'orientation de la chaîne latérale de ce résidus n'a pas lieu pour $HU\alpha_2$ ou alors à une échelle de temps plus grande. On observe cependant que la glutamine 5 semble avoir une dynamique interne plus importante (**Figure 60 panneau de droite**).

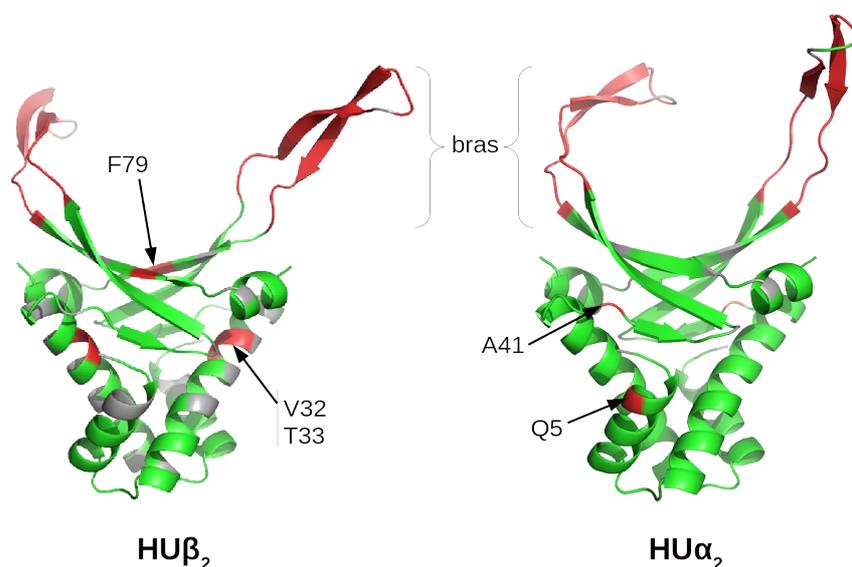


Figure 61: Résidus des homodimères présentant une dynamique interne importante d'après les expériences de relaxation

On peut suspecter que la glutamine 5 constitue un point de fragilité du HTH de HU α_2 tout comme le sont la valine 32 et la thréonine 33 chez HU β_2 . L'alanine 41 a également un comportement inattendu avec R_2 inférieur à la moyenne. Ces expériences doivent être reproduites afin de vérifier que cette différence est significative et reproductible.

De manière générale, à l'exception des bras, HU α_2 est relativement rigide et on ne peut mettre en évidence des résidus présentant une dynamique suffisamment importante pour expliquer la capacité de HU α_2 à transiter vers sa conformation intermédiaire. L'échelle de temps ps-ns ne semble pas adaptée pour ce système.

4.4.5. Effet de la température sur les spectres ^{15}N -HSQC de HU α_2

L'expérience ^{15}N -HSQC enregistre les corrélations proton-azote des fonctions amides de la protéine. On obtient donc au minimum un signal par acide aminé de la protéine à l'exception du résidu N-terminal et des prolines. Nous avons enregistré des spectres ^{15}N -HSQC en faisant varier la température afin d'analyser son effet sur le déplacement chimique et l'intensité des pics des résidus de la protéine. Ce type d'analyse a déjà été réalisé en partie sur l'homodimère HU β_2 ¹⁶⁷ et nous pouvons donc comparer les analyses des deux homodimères.

Suivi des déplacements chimiques

Les liaisons hydrogène sont les liaisons faibles qui sont notamment responsables de la formation des structurations secondaires dans les protéines. Les structurations en hélices α , 3_{10} , π , feuillets β et tours sont toutes stabilisées par des liaisons hydrogène au sein du squelette peptidique. La variation de déplacement chimique du proton H_N en fonction de la température est un bon indicateur de la présence et de la stabilité de ce type de liaison faible.^{224,225} Les protons des fonctions amides impliquées dans des liaisons hydrogène sont influencés par la présence de la fonction carbonyle qui les déblindent. Si l'on augmente la température et que la liaison hydrogène est déstabilisée, la distance entre le H_N et le carbonyle augmente et l'effet de déblindage des fonctions carbonyles diminue. On observe alors une variation de déplacement chimique de ces protons vers les hauts champs.

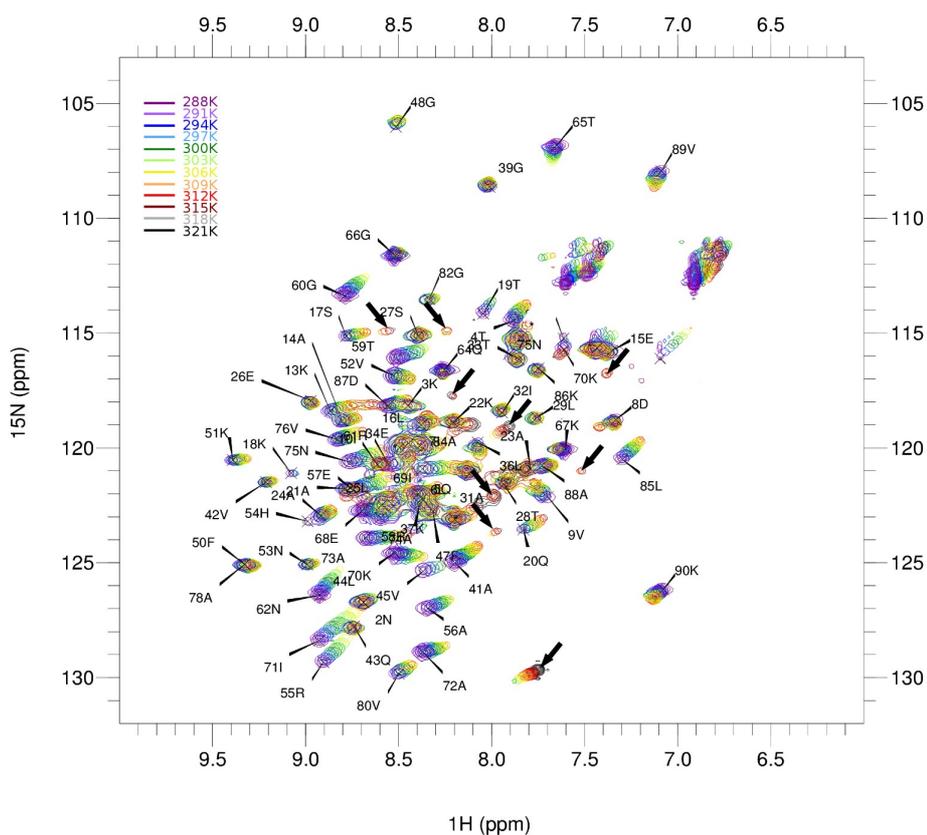


Figure 62: Expériences ^{15}N -HSQC de $\text{HU}\alpha_2$ enregistrées à différentes températures De 288 à 321 K par pas de 3 K. Des pics apparaissant au cours de l'augmentation de la température sont indiqués par des flèches.

Le coefficient de température des protons amides est ainsi dépendant de son contexte structural. D'après le travail de Cierpicki et collaborateurs,²²⁴ en absence de liaison hydrogène, la valeur de ce coefficient est comprise entre -11 et -4,6 ppb/K (**exemple : A41 Figure 63**). Au dessus de -4,6 ppb/K il existe plus de 85 % de chance d'avoir une liaison hydrogène et ce chiffre monte à plus de 93 % au dessus de -4 ppb/K (**exemple : E38 Figure 63**). Dans notre cas, on considère qu'au dessus de -4,6 ppb/K une liaison hydrogène est formée. De plus le coefficient de température est inversement proportionnelle à la longueur de la liaison hydrogène. Plus le coefficient est élevé et plus la liaison hydrogène est stable (et la distance entre l'amide et le carbonyle est courte). Le déplacement chimique des H_N a été suivi au cours de la montée en température (pas de 3 K de 291 K à 321 K) sur des spectres ^{15}N -HSQC.

Le coefficient de température a été calculé pour chaque résidu par ajustement linéaire du

déplacement chimique H (ppm) en fonction de la température (K). Les résultats des ajustements sont résumés dans le tableau 26 en annexe p312. Les graphiques correspondant sont présentés dans les Figures 114 à 116 également en annexe p313 à p315. Les coefficients de température obtenus sont représentés sur la **Figure 64**.

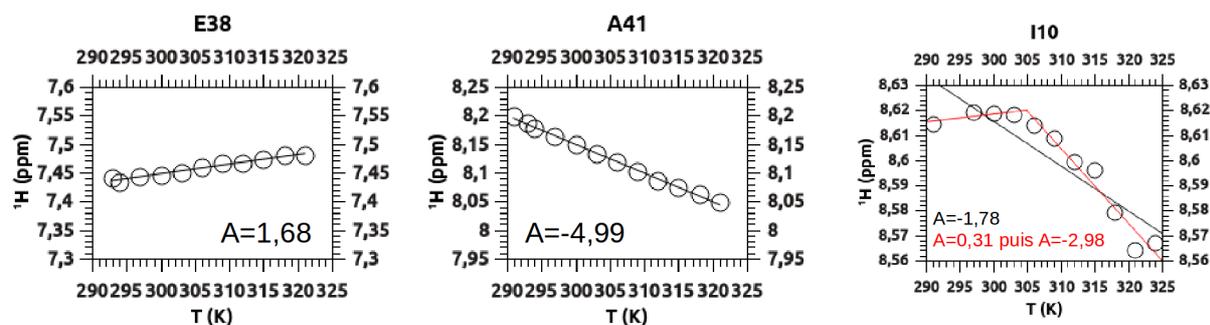


Figure 63: Exemples d'ajustement des courbes de déplacement chimique en fonction de la température pour l'extraction du coefficient de température E38 appartient à l'hélice α_2 , A41 à une portion en pelotte statistique et I10 est positionnée en bas de l'hélice α_1 à l'interface entre les hélices α_2 du HTH.

Comme attendu, une grande partie des résidus dont le coefficient de température est situé sous le seuil sont présents dans des régions exemptes de structure secondaire (**Figure 64**). C'est le cas de la région des bras à l'exception de l'extrémité où des liaisons hydrogène stabilisent le tour α (62-66). Les résidus du petit feuillet à deux brins bordant ce tour α sur la structure ont des coefficients de température inférieurs à $-4,6$ ppb/K. Il est probable que ce petit feuillet soit assez instable et qu'on ne l'observe pas dans tous l'ensemble de structures RMN. Le reste du corps de $HU\alpha_2$ est assez stable et comporte de nombreuses liaisons hydrogène. On remarque que la phénylalanine 79 est juste sous le seuil de liaison hydrogène. Or, un mouvement de type flip-flop de cette phénylalanine, dont la chaîne latérale oscille entre une orientation enfouie dans le cœur hydrophobe et une orientation exposée au solvant, a été mis en évidence dans $HU\beta_2$ par dynamique moléculaire.¹⁶⁷ Si ce mouvement existe également dans $HU\alpha_2$, l'orientation moyenne du squelette peptidique peut apparaître défavorable pour le maintien de la liaison hydrogène du brin β_3 au niveau de ce résidu et expliquer cette valeur.

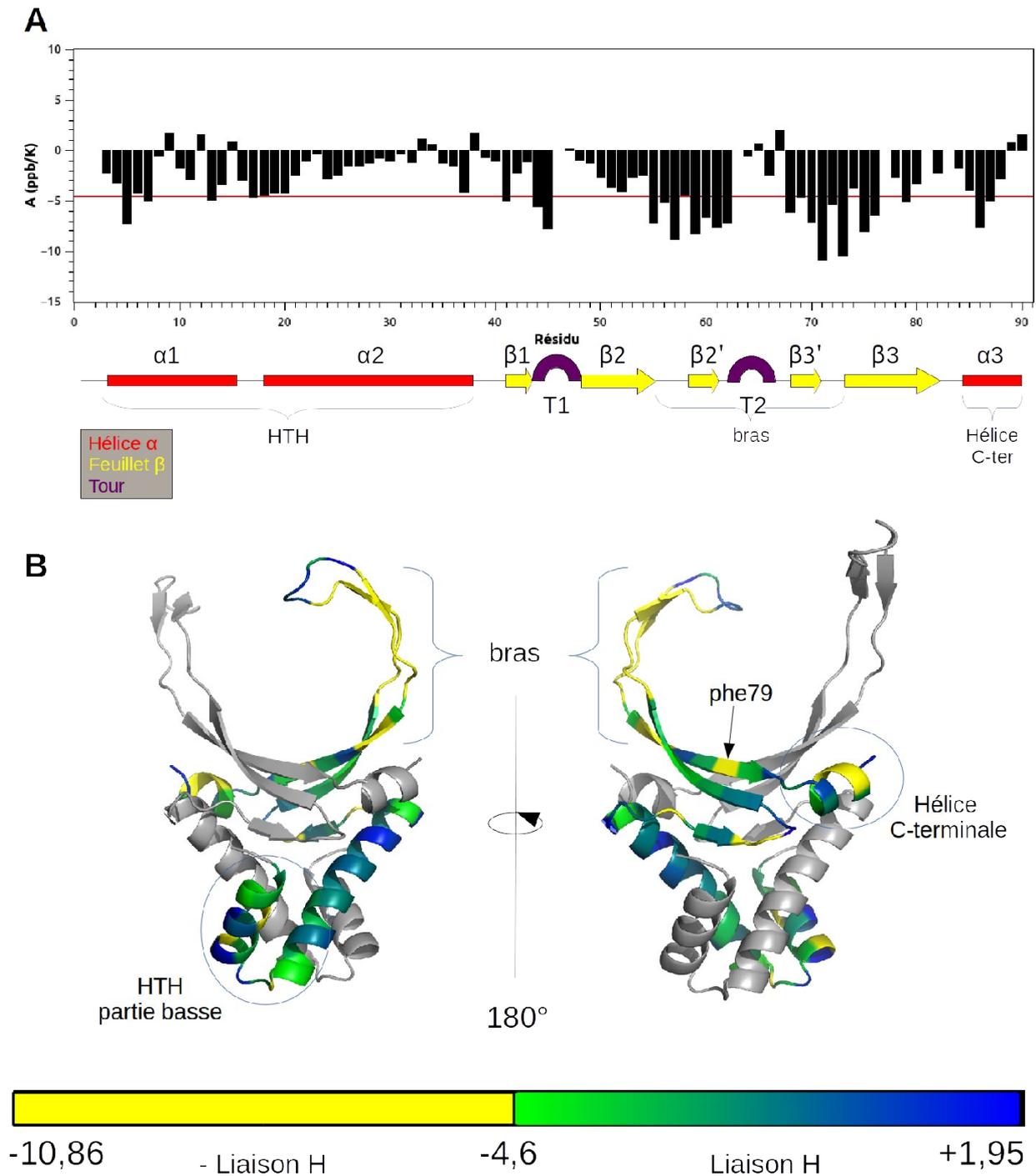


Figure 64: Coefficients de température de HU α_2 en fonction de la structure primaire, secondaire (A) et tertiaire (B)

On remarque également que la petite hélice C-terminale α_3 et la partie basse du HTH comportent des résidus avec des liaisons hydrogène faibles, voire prédites comme absentes (**Figure 64 B**). Il est possible que ces régions constituent des zones de fragilité qui pourraient permettre la transition vers

la conformation intermédiaire I₂.

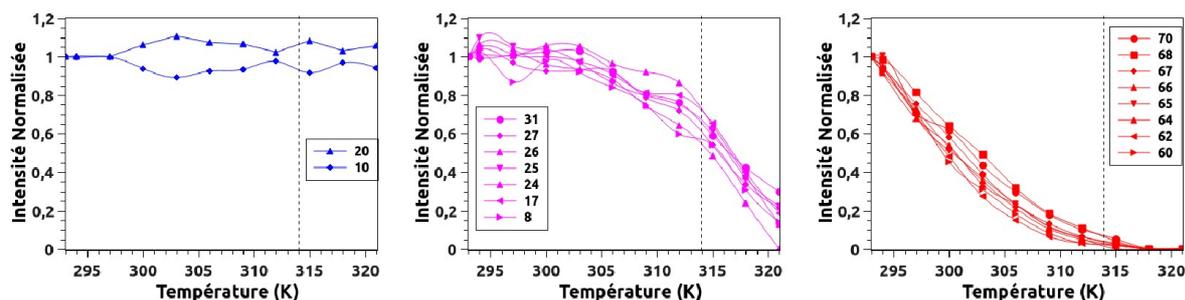
Si l'on regarde plus en détail les courbes d'ajustement, on remarque qu'ils ne sont pas tous linéaires (par exemple I10 **Figure 63** ; ou en annexe : tableau 26 p 312, et figure 114 p 313). Le coefficient peut varier en fonction de la température et conduire à des courbes biphasiques. Ceci a été observé pour les résidus 10, 23, 28 et 29 qui appartiennent au sous-domaine HTH. On remarque que les résidus 10, 28 et 29 font également partie des résidus qui bordent la poche (**Figure 58 p 132**). La non linéarité des coefficients de température a déjà été observée pour des protéines subissant des transitions de conformations à faible énergie libre.²²⁶ Dans notre cas, on peut faire l'hypothèse que l'on observe les premiers changements structuraux de la forme N₂ conduisant à une transition de grande ampleur vers la conformation I₂. Dans ce cas, les premiers résidus impliqués sont ceux de la partie basse du HTH et ceux de l'hélice α_3 C-terminale.

Suivi des intensités

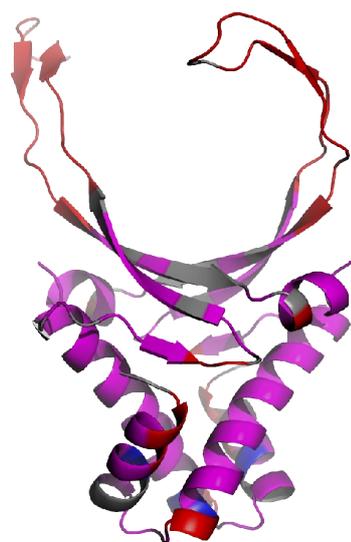
Le second paramètre que l'on peut analyser dans ces expériences ¹⁵N-HSQC à différentes températures est l'évolution de l'intensité. Pour corriger l'effet de la température sur les intensités, on réalise une double normalisation : par l'intensité de chaque résidu à 293K, puis par l'intensité moyenne des résidus qui restent stables jusqu'aux plus hautes températures. Ce type d'analyse a déjà été réalisé pour HU β_2 pour laquelle trois profils d'intensités ont pu être mis en évidence (**Figure 65 A**). Les résidus du premier profil ont une intensité qui reste stable avec l'augmentation de la température. Ces résidus sont situés majoritairement au début de l'hélice α_2 . Le second profil correspond aux résidus dont l'intensité reste stable avec l'augmentation de la température jusqu'à atteindre un seuil correspondant au T_{m1} de l'homodimère HU β_2 (300 K, en pointillé sur la Figure 65 A). L'intensité diminue rapidement lorsque ce seuil est atteint. Les résidus appartenant à ce second profil sont presque tous localisés au niveau du corps de la protéine. Le troisième profil correspond aux résidus déjà dynamiques à basse température (les bras et les extrémités des structures secondaires). L'intensité diminue continuellement avec l'augmentation de température et correspond à une augmentation de la dynamique interne de ces résidus.

En réalisant le même type d'analyse sur HU α_2 , on obtient des profils très similaires (**Figure 65 A**). Les profils d'intensité constante, constante avec seuil puis diminution ou décroissance continue sont également observés (en bleu, magenta et rouge respectivement, **Figure 65 B**). Peu de différences sont observables si ce n'est que le seuil a augmenté au niveau du T m_1 de HU α_2 (314 K, en pointillé sur la **Figure 65 B**).

A

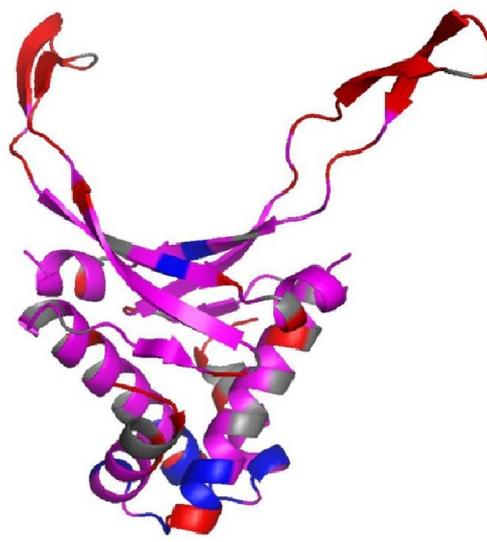


B



HU α_2

C



HU β_2

Figure 65: Evolution de l'intensité des pics ^{15}N -HSQC des résidus des homodimères de HU en fonction de la température. En bleu les résidus présentant des intensités constantes. En magenta les résidus présentant des profils d'intensité constants puis décroissants passé le seuil du T m_1 . En rouge les résidus ayant des intensités diminuant constamment avec l'augmentation de la température. A- Courbes représentatives des profils de température pour HU α_2 . B- Position des profils de température sur le modèle de HU α_2 . C- Position des profils de température sur le modèle de HU β_2 (adapté de Garnier et collaborateurs)¹⁶⁷

L'interprétation de ces résultats conduit à penser que la totalité du corps de HU subit un changement de structure important en atteignant le T_{m1} qui s'accompagne d'une augmentation de la dynamique interne et donc d'une diminution de l'intensité observée.

Toutefois cette interprétation devra être modulée à la lumière des résultats présentés ci-après.

Analyse des pics apparaissant avec l'augmentation de la température

En observant les spectres ^{15}N -HSQC présentés **Figure 62 p 151**, on remarque l'apparition de nouveaux pics au cours de l'augmentation de température. Nous avons d'abord fait l'hypothèse que ces pics étaient dus à une dégradation de la protéine induite par la température mais deux paramètres ont invalidé cette hypothèse.

Tout d'abord, le phénomène est réversible. Lorsque l'on redescend la température, la majorité des pics apparus ne sont plus visibles. On retrouve un spectre HSQC standard complètement superposable avec l'HSQC enregistrée avant la montée en température. Les intensités des pics de HU reviennent à des niveaux d'intensité normaux (**Figure 66 A**). On n'obtiendrait pas ce type de réversibilité dans le cas d'une dégradation de la protéine (en cas de protéolyse par exemple).

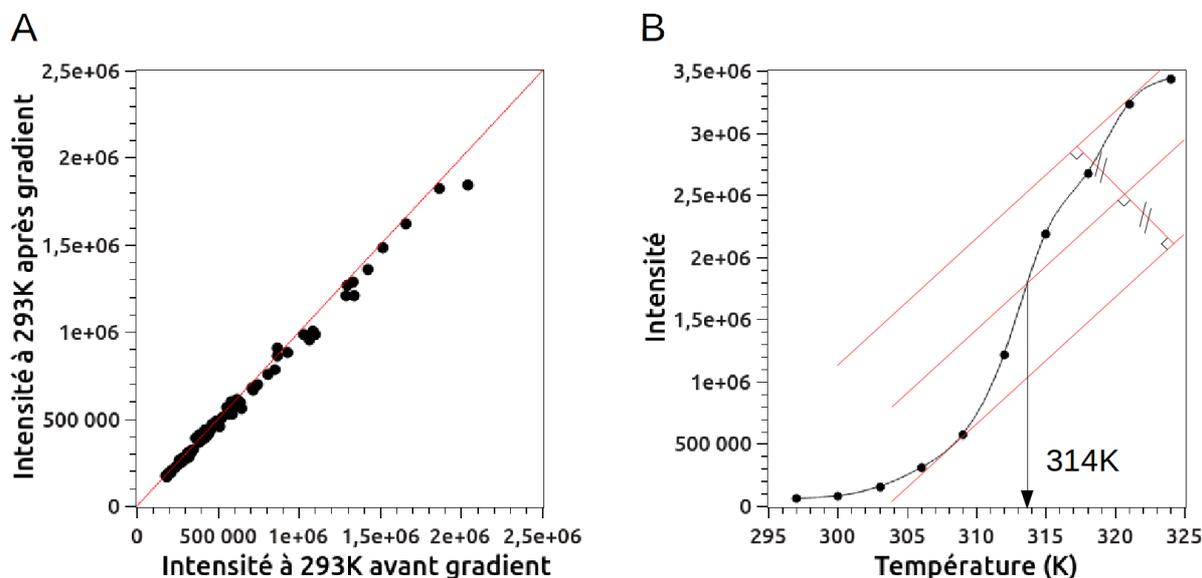


Figure 66: Reversibilité du gradient en température (A) et profil d'intensité d'un pic apparu (B)

De plus, si on analyse l'évolution de l'intensité des pics qui apparaissent au cours de la montée en température on obtient des profils tels que celui présenté **Figure 66 B**. Cette courbe décrit une sigmoïde avec un point d'inflexion situé à 314K, ce qui correspond au T_{m1} de l'homodimère $HU\alpha_2$.

Ainsi, on peut conclure que les pics qui apparaissent au cours de la montée en température ne sont pas dûs à une dégradation irréversible de l'échantillon mais représentent des résidus appartenant à la conformation intermédiaire I_2 . Les deux conformations N_2 et I_2 sont donc en échange lent à l'échelle de la RMN, au moins dans le cas de $HU\alpha_2$ et lorsque la température avoisine le T_{m1} , les deux conformations sont visibles. Ce résultat altère l'interprétation du suivi de l'intensité des pics en fonction de la température. L'évolution de cette intensité pouvant à la fois être la conséquence de la dynamique mais également l'effet de titrage de la conformation N_2 vers I_2 . Dans ce dernier cas, l'évolution de l'intensité en fonction de la température doit suivre une sigmoïde inversée par rapport à la courbes présentée **Figure 66 B**, avec un point d'intersection au niveau du T_{m1} .

Dans ce chapitre, nous avons décrit les structures et les stabilités des dimères de HU dans leur conformation native. Les résidus du HTH bordant la cavité hydrophobe et les résidus de l'hélices C-terminales semblent des points de fragilité potentielle dans ces structures qui pourraient permettre la transition vers l'état partiellement déstructuré I_2 . L'échange lent entre les conformations N_2 et I_2 nous permet de mettre en évidence des signaux spécifiques de cette dernière conformation en RMN. L'objectif principal du prochain chapitre sera la caractérisation structurale et cinétique de la transition vers la conformation intermédiaire I_2 .

4.5. À retenir

<p>Objectif du chapitre : Identifier les paramètres structuraux et/ou les résidus essentiels pour la transition vers l'état intermédiaire I₂ et l'échange de chaînes.</p>
<p>Dans la 1^{ère} partie du chapitre, j'ai :</p>
<ul style="list-style-type: none"> - Déterminé la structure RX du dimère HUα_2 d'<i>E. coli</i>, sous sa forme native N₂
<ul style="list-style-type: none"> - Réalisé le clonage/production/purification de la protéine HU de <i>L. lactis</i>, selon le protocole optimisé dans le chapitre 3.
<ul style="list-style-type: none"> - Déterminé la structure RX de HU de <i>L. lactis</i> (en cours d'affinement)
<ul style="list-style-type: none"> - Comparé les structures des protéines HU des différents organismes (disponibles dans la PDB et déterminés ici), en terme de : <ul style="list-style-type: none"> - Surface d'interaction entre les monomères - Nbre de liaisons hydrogène et ponts salins - Cavités hydrophobes → Corrélation de ces paramètres structuraux avec les données de stabilité des dimères, et leur capacité à échanger leurs chaînes.
<p>Dans la 2^{ème} partie du chapitre, j'ai :</p>
<ul style="list-style-type: none"> - Attribué les δ du squelette des 3 dimères HUα_2, HUβ_2 et HU$\alpha\beta$
<ul style="list-style-type: none"> - Attribué les δ des chaînes latérales des homodimères HUα_2 et HUβ_2, - Calculé leurs structures 3D (en cours d'affinement) <ul style="list-style-type: none"> → Structures comparables aux structures RX, avec bras visible
<ul style="list-style-type: none"> - Analysé la dynamique interne des homodimères HUα_2 et HUβ_2, <ul style="list-style-type: none"> → Identification des 1^{ers} résidus qui conduiraient à la forme intermédiaire I₂
<ul style="list-style-type: none"> - Analysé l'effet de la température sur les homodimères HUα_2 et HUβ_2, <ul style="list-style-type: none"> → Identification des 1^{ers} changements structuraux qui conduiraient à la forme intermédiaire I₂ → Mise en évidence d'un échange lent N₂ ↔ I₂ à l'échelle de la RMN → Identification les 1^{ers} signaux appartenant la forme intermédiaire I₂

Chapitre 5 : La conformation intermédiaire

La première étape du mécanisme d'échange envisagé est la transition des homodimères de HU de leur conformation native vers leur conformation intermédiaire. Dans le chapitre précédent, nous avons analysé les protéines HU dans leur conformation native et avons mis en évidence une région de fragilité potentielle de la structure au niveau du HTH autour d'une cavité hydrophobe, ainsi que dans la région de l'hélice C-terminale. L'hypothèse que la conformation intermédiaire implique une déstructuration dans ces régions est cohérente avec le travail de Ramstein et collaborateurs.¹⁶⁶ Les auteurs ont montré que cette transition, médiée par la température, s'accompagne d'une déstructuration partielle dans les structures secondaires en hélice α , grâce à des expériences de dichroïsme circulaire (**Tableau 14**).

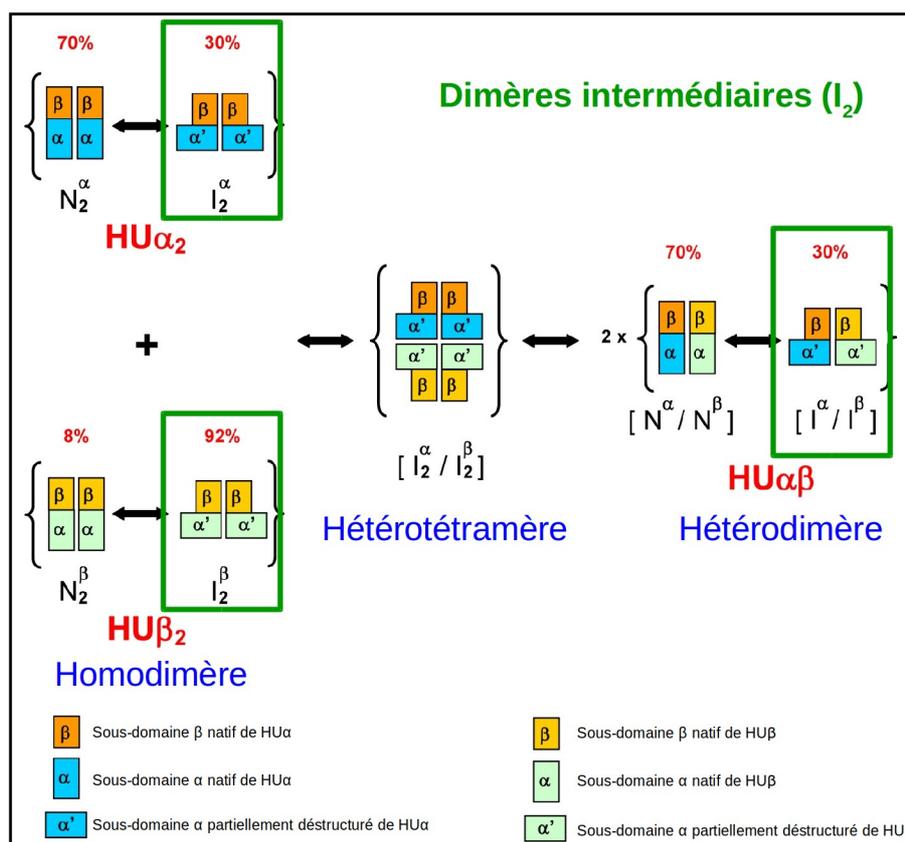


Figure 67: Dimères intermédiaires (I_2) dans le mécanisme putatif d'échange de chaînes des dimères de HU

Tableau 14: Structuration Secondaire des dimères de HU de *E. coli* (Adapté de Ramstein *et al*)

Protéine		Structuration Secondaire (%)			
		Hélice- α	Feuillet- β	Tour	Autre
<i>Dichroïsme Circulaire</i>					
EcHU α_2	Native	43,2	16,7	17,1	22,6
	50°C	17,0	20,7	17,6	44,2
EcHU β_2	Native	54,8	20,5	4,7	20,1
	43°C	25,1	22,6	20,3	31,6
EcHU $\alpha\beta$	Native	42,7	15,8	18,2	22,9
	50°C	21,0	21,5	18,8	38,2
<i>Modèles structuraux</i>					
EcHU α_2 (pdb : 1MUL)	Native	42,2	21,1	7,8	30,0
	Native	42,2	18,9	7,8	31,1
EcHU β_2 (pdb: 4P3V)	Native	42,2	18,9	7,8	31,1
	Native	43,3 (HU α) 40,0 (HU β)	14,4 (HU α) 16,6 (HU β)	6,6 (HU α) 7,7 (HU β)	34,4 (HU α) 35,6 (HU β)

Cette déstructuration au niveau des hélices α a également été suggérée par des expériences de simulation de dynamique moléculaire effectuées sur un modèle de HU β_2 par Garnier et collaborateurs.¹⁶⁷ Dans ce travail, les auteurs ont proposé un modèle de la conformation intermédiaire de HU β_2 sur la base de simulations à très hautes températures. Mais, si ce modèle permet d'expliquer certaines propriétés de HU comme sa susceptibilité à la protéase Lon et la déstabilisation des hélices, la structure proposée ne permet pas, à elle seule, d'envisager une association spécifique avec l'autre homodimère pour l'échange de chaîne (**Figure 68**).

Une description structurale complète des états intermédiaires (I_2) des homodimères nous permettrait de faire des hypothèses sur leur mode d'association pour l'échange de chaînes. A la fin du chapitre précédent, des expériences RMN ^{15}N -HSQC enregistrées à températures croissantes ont permis de mettre en évidence que la conformation I_2 est visible en RMN et est en échange lent avec la conformation N_2 . Dans ce chapitre, nous exploitons les données acquises par RMN et de nouvelles expériences de dynamique moléculaire pour caractériser l'état intermédiaire du point de vue cinétique et structural.

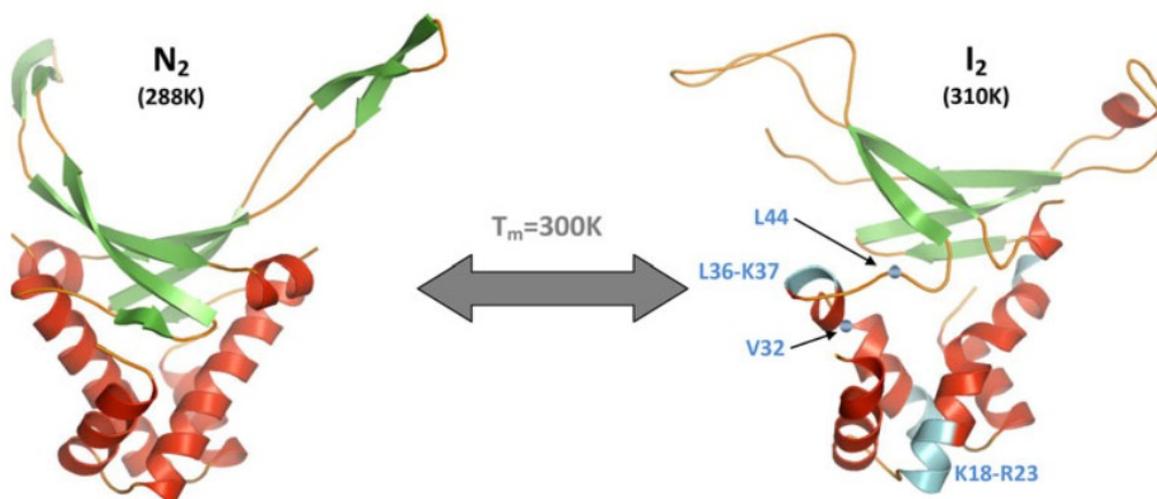


Figure 68: Modèle structural de la transition $N_2 \rightarrow I_2$ de $HU\beta_2$ proposé par garnier *et al.*¹⁶⁷ Les résidus en bleu indiquent les régions de clivages préférentiels de la protéase Lon à 37°C.^{163,265} Ces résidus sont plus exposés dans le modèle de I_2 que dans celui de N_2 .

5.1. Caractérisation des conformations intermédiaires par RMN

Les conformations N_2 et I_2 étant en échange lent, les deux états sont visibles et donnent des signaux distinguables en RMN lorsque la température est suffisamment proche du T_{m1} du dimère observé. Il faut faire un compromis entre l'augmentation de la proportion de la forme intermédiaire à haute température et la perte de signal associée à l'élévation de température. Nous nous sommes placés à 37°C (310 K) pour $HU\alpha_2$ et 27°C (300 K) pour $HU\beta_2$.

5.1.1. Attribution de la conformation intermédiaire des homodimères

Une série d'expériences standards nécessaires pour l'attribution du squelette peptidique, des chaînes latérales et des NOEs a été enregistrées pour les deux homodimères (**Tableau 15**).

En plus des expériences standards, nous avons enregistré une série d'expériences ZZ-¹⁵N-HSQC. Ce type d'expérience a été utilisé pour étudier des cinétiques de repliement de protéines telles que celles du domaine SH3 ou du domaine FF.^{186,227} La ZZ-¹⁵N-HSQC est similaire à une ¹⁵N-HSQC standard mais un délai (T) est inséré entre les deux pulses INEPT responsables du transfert

d'aimantation $H \rightarrow N$ et $N \rightarrow H$ (**Figure 69 A**).¹⁸⁵ Durant ce délai, si l'échange entre les deux conformations a lieu dans la bonne gamme de temps (de l'ordre de la milliseconde à la seconde) la fréquence de résonance de l'azote passe de sa fréquence de résonance N_2 à sa fréquence de résonance I_2 et inversement. Cela se traduit par l'apparition de pics de corrélation entre les deux pics d'un résidu en échange entre deux conformations. C'est ce que l'on peut voir pour le pic H_N-N de la glycine 46 du domaine SH3 (**Figure 69 B**) en équilibre entre une conformation native (N) et une conformation dénaturée (D).

Tableau 15: Enregistrement des expériences RMN sur les homodimères de HU de *E.coli*

	Expériences \ Protéine	HU α_2	HU β_2
		I_2 37°C	I_2 27°C
	^{15}N -HSQC	√	√
	^{13}C -HSQC	√	√
Squelette peptidique	HNCO	√	√
	HNCA	√	√
	HN(CO)CA	√	√
	HN(CA)CO	√	√
	HNCACB	√	√
	CBCA(CO)NH	√	√
	Cc(CO)NH	√	√
Chaîne Latérales	HBHA(CO)NH	√	√
	HBHANH	√	√
	H(CcCO)NH	√	√
	^{15}N -HSQC-TOCSY	√	√
	HCCH-TOCSY	√	√
NOESY (structure)	^{13}C -HSQC-NOESY	√	√
	^{15}N -HSQC-NOESY	√	√
	1H -NOESY	√	√
ZZ	zz- ^{15}N -HSQC	√	√

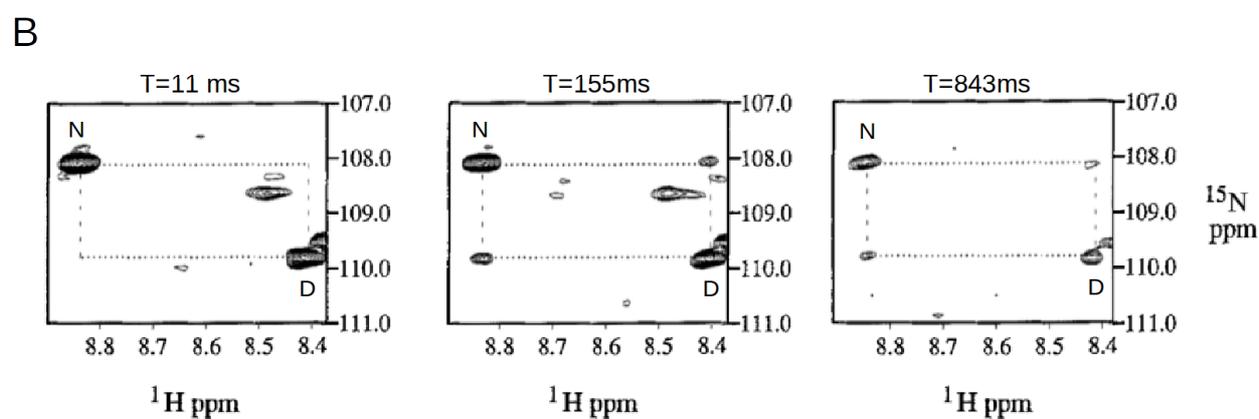
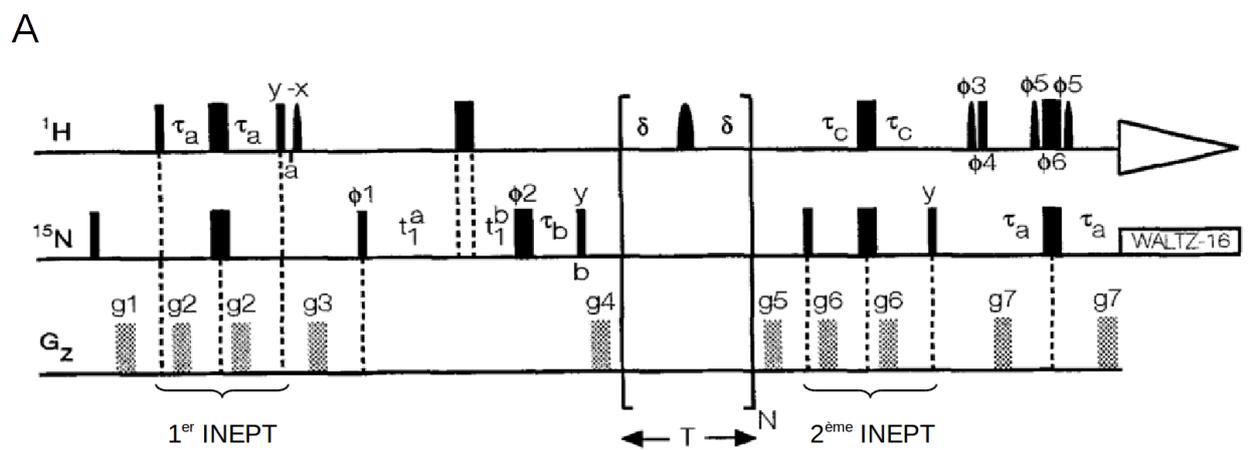


Figure 69: ZZ- ^{15}N -HSQC Adapté de Farrow N. et collaborateurs.¹⁸⁵ A- Séquence de pulses B- Spectre dans la région de la Glycine 46 de SH3 en équilibre entre une forme native (N) et une forme dénaturée (D).

Un exemple de spectre ZZ-¹⁵N-HSQC est présenté **Figure 70** pour chaque homodimère de HU de *E. coli*. Ces spectres ont été enregistrés pour des valeurs croissantes de T (de 50 à 1000 ms) à 310 K et 300 K pour HU α_2 et HU β_2 , respectivement.

Dans le cas de l'homodimère HU α_2 , de nombreuses corrélations ont été observées sur l'ensemble de la gamme de T et ont facilité l'attribution de la conformation intermédiaire, en particulier lorsque les pics sont bien isolés (par exemple K90, **Figure 70**). Une confirmation par une attribution séquentielle standard a été nécessaire dans les cas limites où une des corrélations est superposée avec le pic d'un autre résidu (A14). Par ailleurs, certains résidus présentant le même déplacement chimique azote entre les deux conformations ou présentant des superpositions n'ont pu être attribués que séquentiellement.

Dans le cas de l'homodimère HU β_2 , très peu de corrélations ont pu être observées. Seuls les résidus 89 et 90 ayant des intensités et des écarts de déplacement chimique importants ont permis de mettre en évidence des pics de corrélation, et seulement pour les valeurs de T les plus élevées (**Figure 70**). La différence entre les expériences réalisées sur les deux homodimères est probablement liée à l'effet de la température sur la cinétique d'échange. Si l'on travaille au T_{m1} de HU β_2 , les deux conformations sont suffisamment peuplées mais la cinétique d'échange est trop lente pour observer des pics de corrélation intenses. Nous avons essayé d'enregistrer les mêmes expériences pour des températures supérieures (306 et 310 K) mais la perte du signal de la conformation native est trop importante. De ce fait, la majorité des résonances de la conformation intermédiaire de HU β_2 a été attribuée par la méthode d'attribution séquentielle classique.

Le taux d'attribution actuel des homodimères HU α_2 et HU β_2 , à 310K et 300K respectivement, est présenté dans le **Tableau 18**. Au moins 37 % des résidus présentent un déplacement chimique différent de N₂, au niveau du squelette peptidique, dans la conformation I₂. Or, comme la grande majorité des pics de I₂ apparus sur les spectres ont pu être attribués, on fait l'hypothèse que les résidus restants n'ont pas de variation de déplacement chimique (même environnement chimique et électronique dans les conformations natives et intermédiaires).

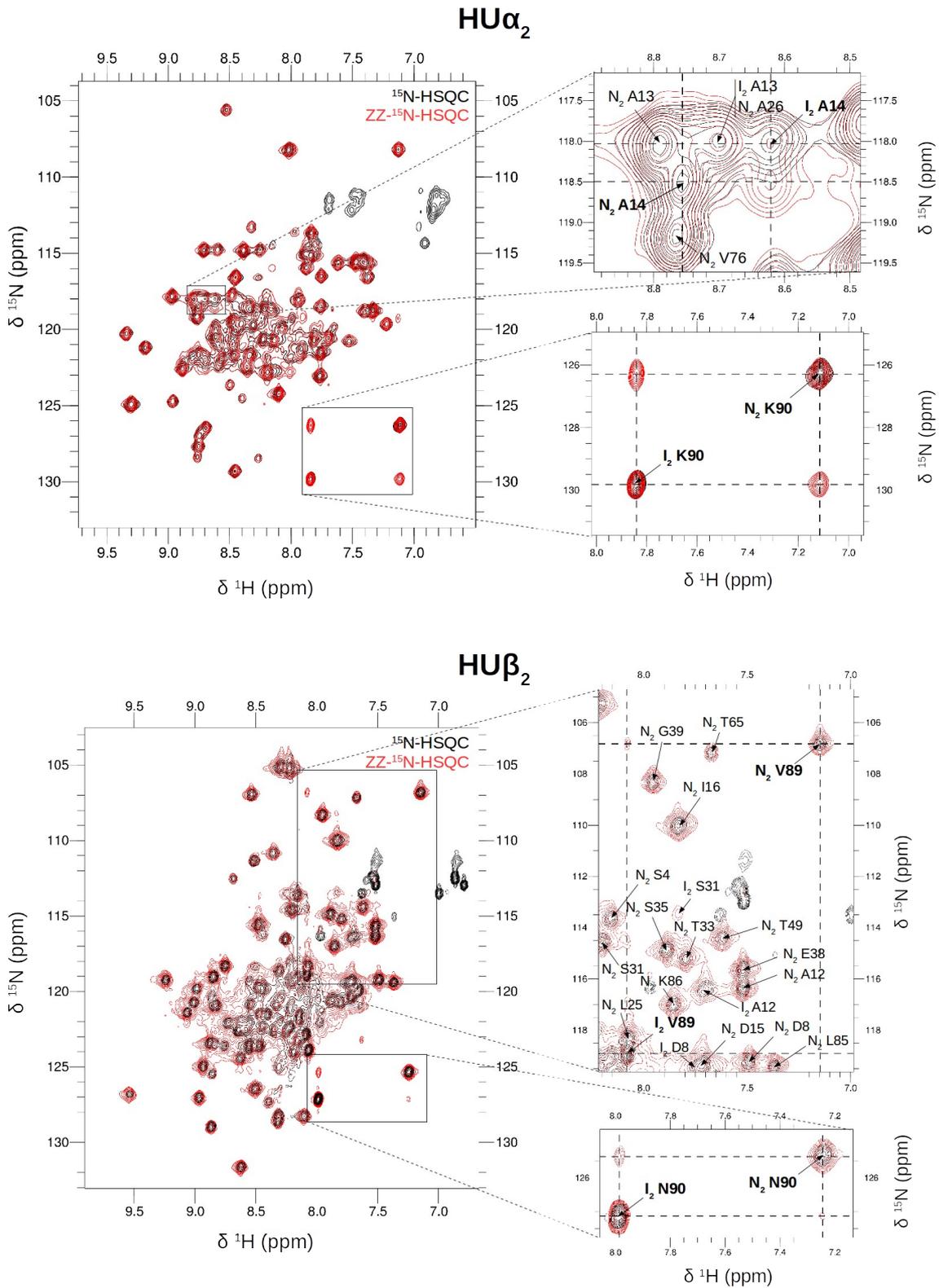


Figure 70: Superposition des spectres ^{15}N -HSQC et ZZ- ^{15}N -HSQC des homodimères de HU Spectres enregistrés à 310K à T= 500ms pour HU α_2 et à 300K et T=1000ms pour HU β_2 .

Tableau 16: Attribution des dimères de HU dans leur conformation native et intermédiaire

Attribution \ Protéine		HU α_2 37°C / 310 K		HU β_2 27°C / 300 K	
		N ₂	I ₂	N ₂	I ₂
Squelette peptidique	Totaux	89 %	37 %* / 81 %**	98 %	37 %* / 84%**
Chaînes latérales	H	57 %	21 %* / 51 %**	87 %	29 %* / 65 %**
	Non-H	51 %	19 %* / 44 %**	70 %	25 %* / 52 %**

* Taux d'attribution si l'ensemble des résidus avait un déplacement chimique différent de N₂ dans la conformation I₂.

** Taux d'attribution si l'ensemble des résidus manquant pour I₂ ont en réalité les mêmes déplacements chimiques que pour la conformation N₂.

5.1.2. Calcul des constantes cinétiques d'échange

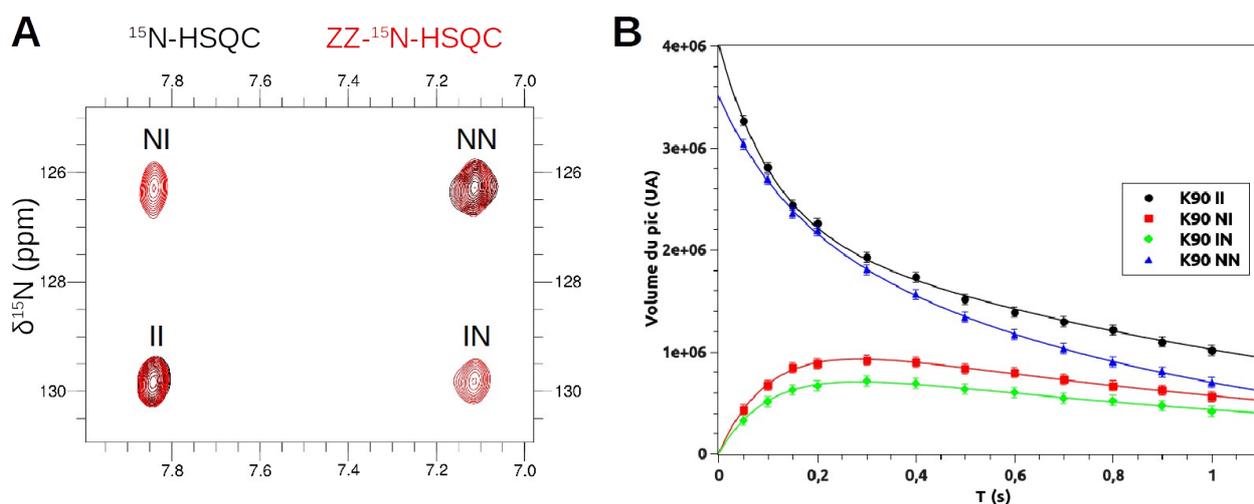


Figure 71: Déduction des paramètres cinétiques d'échange N₂/I₂ par ZZ-¹⁵N-HSQC de HU α_2

Les constantes cinétiques d'échange entre les conformations N₂ et I₂ de HU α_2 peuvent être extraites à partir des expériences ZZ-¹⁵N-HSQC enregistrées à différentes valeurs de T (50ms à 1000ms). Le résidu K90 a été utilisé car les pics des formes N₂ et I₂ ainsi que ceux des corrélations sont intenses et bien résolus sans superposition avec ceux d'autres résidus (**Figure 71 A**). Les volumes des pics auto (NN et II) et croisés (NI et IN) sont tracés en fonction de T puis ajustés en utilisant les équations 1.A à 1.D comme indiqué dans le chapitre 2 (matériels et méthodes p 72). Ceci nous permet d'extraire les valeurs des constantes de vitesse k₁ et k₋₁ représentant la réaction N₂ → I₂ et le retour I₂ → N₂, respectivement. Ainsi, on obtient k₁=2,4 s⁻¹ et k₋₁=7,7 s⁻¹. On peut en déduire la

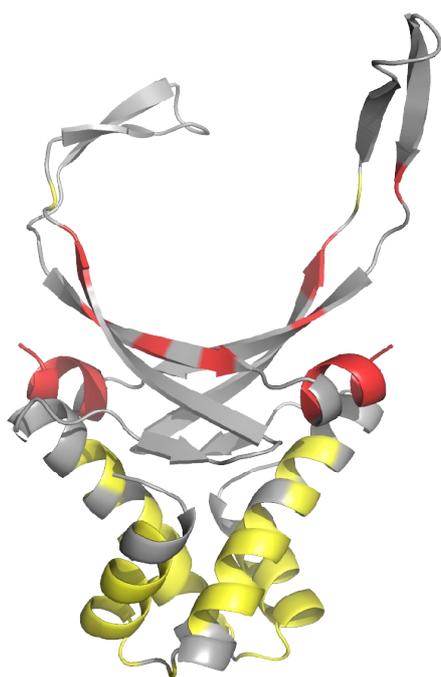
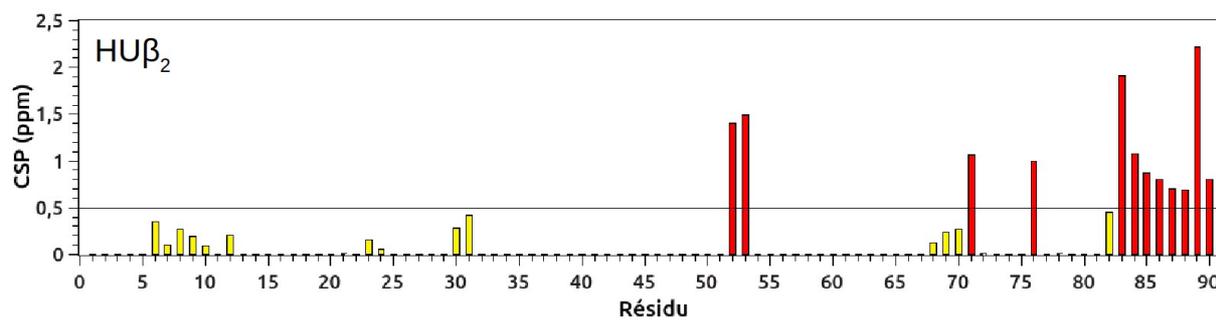
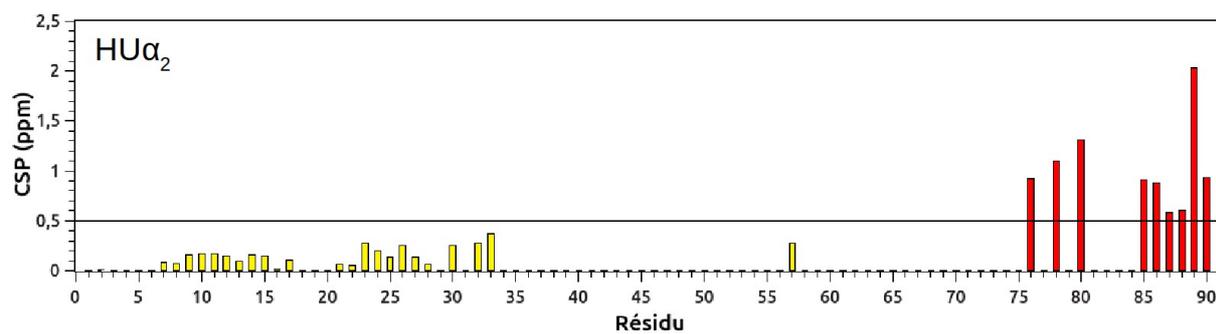
constante d'équilibre $K = k_1/k_{-1} = 0,31$ ce qui est cohérent avec la valeur de K obtenue par des expériences de DSC réalisées précédemment ($K=0,29$ à 310K pour $HU\alpha_2$).¹⁶⁷

Malheureusement, ces constantes de vitesse d'échange n'ont pas pu être évaluées pour l'homodimère $HU\beta_2$. A basse température (300K) on obtient seulement le début du profil obtenu pour $HU\alpha_2$ à 310K (augmentation des volumes des pics de corrélation). La réaction est probablement trop lente dans cette gamme de température. Si on augmente la température jusqu'à 310K ($T_{m1} + 10K$) pour accélérer l'échange, on perd le signal de la conformation N_2 . Nous n'avons pas pu obtenir de compromis satisfaisant entre la vitesse d'échange et l'abondance relative des deux conformations pour extraire des constantes cinétiques, mais la seule présence des pics de corrélation observés nous permet d'affirmer que la vitesse d'échange a lieu également dans la gamme de la milliseconde à la seconde pour $HU\beta_2$.

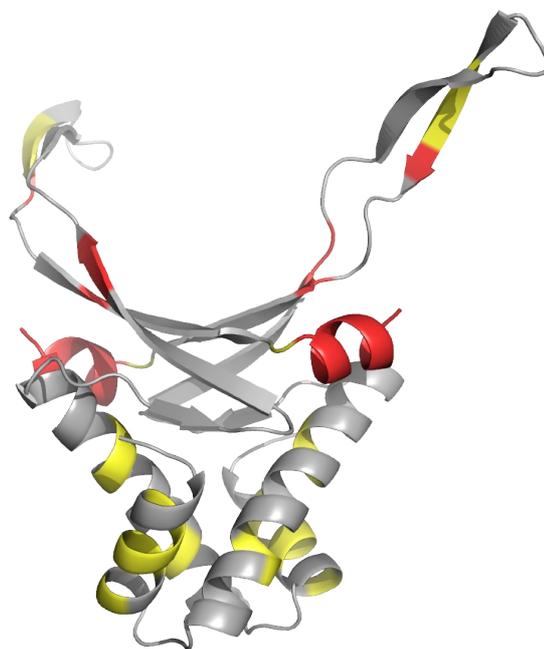
5.1.3. Perturbations de déplacement chimique entre les conformations N_2 et I_2

Les attributions des conformations natives et intermédiaires des homodimères nous permettent d'évaluer les perturbations de déplacements chimiques (CSP). Les CSP (H_N-N) ont été calculés comme indiqué dans le chapitre 2 (matériels et méthodes p 73) et sont représentés sur la **Figure 72**. Les CSP traduisent des changements d'environnement chimique qui peuvent être dus à des variations de structures secondaires et/ou tertiaires. Les valeurs des CSP mesurées permet de rassembler les résidus des deux homodimères dans trois classes différentes. La première classe correspond à des résidus ayant des CSP nuls (ou non mesurables). Les résidus de cette classe présente un signal unique pour les conformations N_2 et I_2 (ou pour lesquels une attribution est manquante) (**en gris Figure 72**). La seconde classe correspond à des résidus ayant des CSP relativement faibles ($<0,5$ ppm, **en jaune Figure 72**) et appartenant principalement au domaine HTH. Les résidus ayant des CSP élevés ($>,5$ ppm, **en rouge Figure 72**) constituent la troisième classe et sont situés majoritairement dans l'hélice α_3 C-terminale ainsi que dans certaines parties des feuillets β environnants.

La contribution dans les CSP des variations de structures secondaires ou tertiaires ne peut pas être déduite directement. Pour pouvoir discerner les deux effets, on peut exploiter l'attribution des squelettes peptidiques pour prédire la structuration secondaire des deux conformations des homodimères.



$HU\alpha_2$



$HU\beta_2$

Figure 72: CSP H_N -N entre les conformations natives et intermédiaires des homodimères de HU

5.1.4. Prédictions des structures secondaires à partir des déplacements chimiques

Deux méthodes ont été utilisées pour prédire les structures secondaires des homodimères dans leurs différentes conformations : le CSI²²⁸ (Chemical Shift Index), et DANGLE¹⁸⁴.

Le déplacement chimique des atomes du squelette peptidique est fortement influencé par la structuration secondaire. Le déplacement chimique des carbonyles et C α augmente, par rapport à leur valeur en absence de structure secondaire et dans les structures en hélice alors qu'il diminue dans des structures étendues en feuillet β .^{229,230} On observe l'effet inverse dans le cas des déplacements chimiques des H α et des C β .^{228,231}

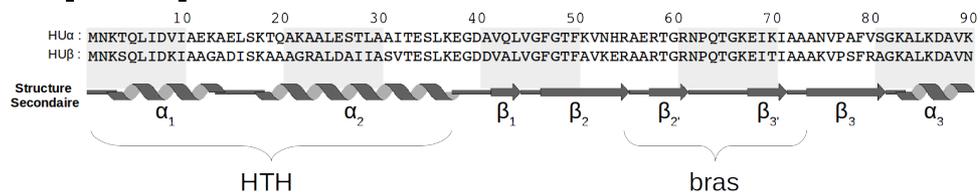
Ces effets sont exploités pour déterminer le CSI, les différences entre les déplacements chimiques attribués et leur valeur de référence en pelote statistique sont annotées $\Delta\delta$ (**Figure 73**). Dans le cas d'une hélice α , on aura $\Delta\delta^{13}\text{C}\alpha > 0$, $\Delta\delta^{13}\text{C}\beta < 0$, $\Delta\delta^{13}\text{CO} > 0$ et $\Delta\delta^1\text{H}\alpha > 0$, et le CSI vaudra -1. Dans le cas d'un feuillet β , on aura $\Delta\delta^{13}\text{C}\alpha < 0$, $\Delta\delta^{13}\text{C}\beta > 0$, $\Delta\delta^{13}\text{CO} < 0$ et $\Delta\delta^1\text{H}\alpha < 0$, et le CSI vaudra +1. En l'absence de structure secondaire, on aura $\Delta\delta^{13}\text{C}\alpha \approx 0$, $\Delta\delta^{13}\text{C}\beta \approx 0$, $\Delta\delta^{13}\text{CO} \approx 0$ et $\Delta\delta^1\text{H}\alpha \approx 0$, et le CSI vaudra 0.

La deuxième approche de prédiction des structures secondaires exploite le logiciel DANGLE. Ce logiciel compare les déplacements chimiques des atomes du squelette peptidique avec ceux d'une banque de données de déplacements chimiques de protéines dont les structures sont connues. Des correspondances sont réalisées sur des fragments de séquences identiques ou proches entre la protéine d'intérêt et les protéines présentes dans la banque de données. Des valeurs des angles dièdres φ et ψ sont ainsi prédites par inférence bayésienne et permettent de déduire la structuration secondaire correspondante. Les structurations secondaires prédites par DANGLE pour les conformations natives et intermédiaires des homodimères sont représentées schématiquement sur la **Figure 73**.

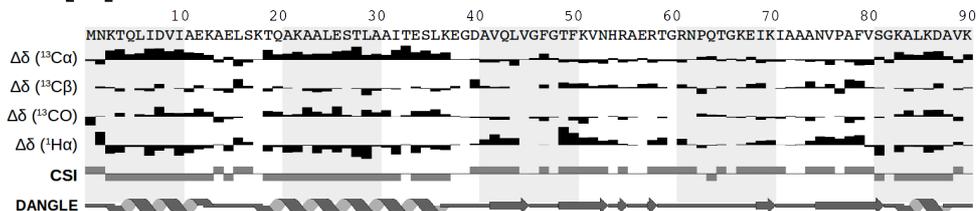
On observe un bon accord entre les prédictions de structures secondaires des conformations natives des homodimères et les structures secondaires observées dans les modèles structuraux à l'exception du feuillet du bras ($\beta 2' - \beta 3'$) qui est bien détecté en CSI mais assez mal par DANGLE. Cependant, ces brins β ne sont pas systématiquement visibles dans toutes les structures préliminaires en solution (**Figure 59 p 146**) et il est possible qu'ils ne soient que transitoires. cela expliquerait qu'ils soient difficiles à détecter. Dans le cas des conformations intermédiaires, on observe une perte de

structuration secondaire dans la partie C-terminale des homodimères (région marquée d'un crochet rouge sur la **Figure 73**) où les valeurs de $\Delta\delta$ tendent vers zéro. Contrairement à la partie C-terminale, le HTH semble conserver sa structuration secondaire entre N_2 et I_2 .

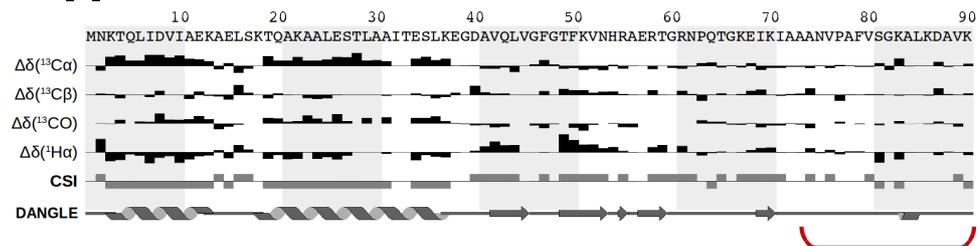
HU α_2 et HU β_2 (structures secondaires dérivée des structures natives) :



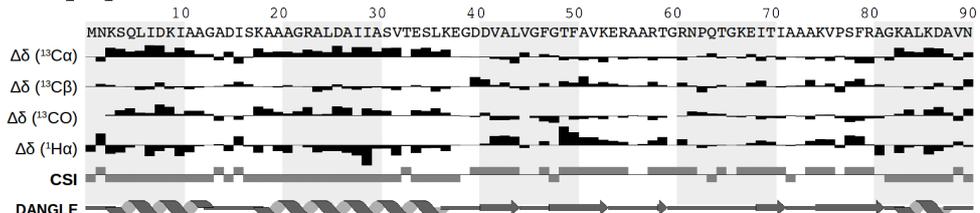
HU α_2 N $_2$ (37°C) :



HU α_2 I $_2$ (37°C) :



HU β_2 N $_2$ (27°C) :



HU β_2 I $_2$ (27°C) :

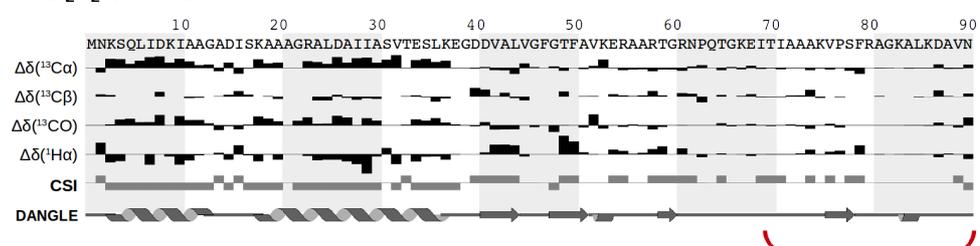


Figure 73: Prédictions des structures secondaires des homodimères dans leurs conformations natives et intermédiaires par la méthode du CSI et le logiciel DANGLE

D'après ces prédictions de structuration secondaire, on peut proposer une interprétation des profils de CSP observés précédemment. Dans le cas de la partie C-terminale et de certains résidus des feuillets β environnants, on a une forte variation de déplacement chimique (**Figure 72**) que l'on peut associer à une perte de structuration secondaire (**Figure 73**) et probablement tertiaire. Cela expliquerait les CSP importants des résidus V76, A78 et V80 de HU α_2 . Ces résidus appartenant au brin β_3 forment des liaisons hydrogène avec le brin β_2 . La déstructuration du brin β_3 entraîne ainsi la perte spécifique de ces liaisons hydrogène et entraîne des variations de déplacement chimique importantes (1 résidu sur 2). Dans le cas des résidus du HTH, on observe une perturbation de déplacement chimique plus faible alors que la structure secondaire est globalement conservée dans ce domaine structural. On peut donc en déduire que la variation de déplacement chimique du HTH est due à une variation de structure tertiaire uniquement. On peut imaginer une réorientation de ces hélices α vers une ouverture du cœur hydrophobe comme suggéré dans le modèle structural de l'échange de chaîne envisagé dans le chapitre introduction (**Figure 19 p 55**). La déstructuration de l'hélice α_3 C-terminale, à l'interface entre la fin de l'hélice α_2 du HTH et le feuillet β , serait nécessaire pour que les hélices du HTH aient un degré de liberté suffisant pour l'ouverture du cœur hydrophobe.

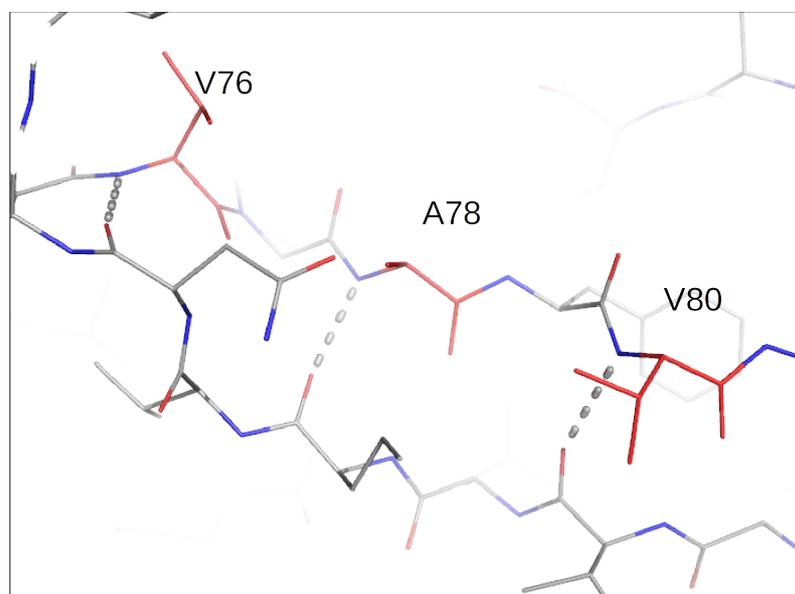


Figure 74: Liaison hydrogènes du brin β_3 de HU α_2 Les résidus en rouges présentent des CSP $N_2 \rightarrow I_2$ importants (Figure 72) et forment des liaisons hydrogène avec les brins β_2 du feuillet.

5.1.5. Analyses des spectres NOESY des conformations natives et intermédiaires des homodimères

Nous avons cherché à attribuer les spectres NOESY afin d'obtenir le maximum d'informations structurales sur les conformations I_2 et ainsi vérifier nos hypothèses. Des expériences ^{15}N -HSQC-NOESY, ^{13}C -HSQC-NOESY et ^1H -NOESY ont été enregistrées à 37°C et 27°C pour les homodimères $\text{HU}\alpha_2$ et $\text{HU}\beta_2$, respectivement. La conservation des NOEs entre les deux formes peut aisément être estimée sur les spectres ^{15}N -HSQC-NOESY. L'HSQC étant attribuée, on compare les bandes en dimensions ^1H en se plaçant aux déplacements chimiques H_N et N des résidus appartenant à N_2 ou I_2 . Les résultats sont présentés sur les **Figures 75 et 76** pour $\text{HU}\alpha_2$ et sur les **Figures 77 et 78** pour $\text{HU}\beta_2$.

Pour $\text{HU}\alpha_2$, on remarque que la majorité des NOEs est conservée entre les deux formes dans la partie HTH (résidus 7 à 29). Dans le spectre équivalent pour $\text{HU}\alpha_2$ à 293K , en moyenne un NOE sur trois est à longue distance (>5 résidus d'écart dans la séquence) et un sur dix est inter-chaînes. D'après les données de prédictions de structure secondaire, on peut raisonnablement faire l'hypothèse que les NOEs séquentiels et à courtes distances (écart inférieur ou égal à 4 résidus) sont conservés et que les NOEs longues distances disparaissent, traduisant un changement de structure tertiaire sans détruire les hélices du HTH. Dans le cas du haut de l'hélice α_2 (résidus 30 à 33) et de l'hélice α_3 C-terminale (résidus 85-90), on observe une perte quasi-totale des NOEs. L'hélice C-terminale est un point de contact entre le haut de l'hélice α_2 et le feuillet β .

La perte des NOEs pour ces résidus est probablement due à une déstructuration. La perte de structuration secondaire a été mise en évidence Figure 73 p 170, en particulier pour la plupart des résidus de l'hélice C-terminale de la conformation I_2 de $\text{HU}\alpha_2$, mais également pour les résidus 32 et 33.

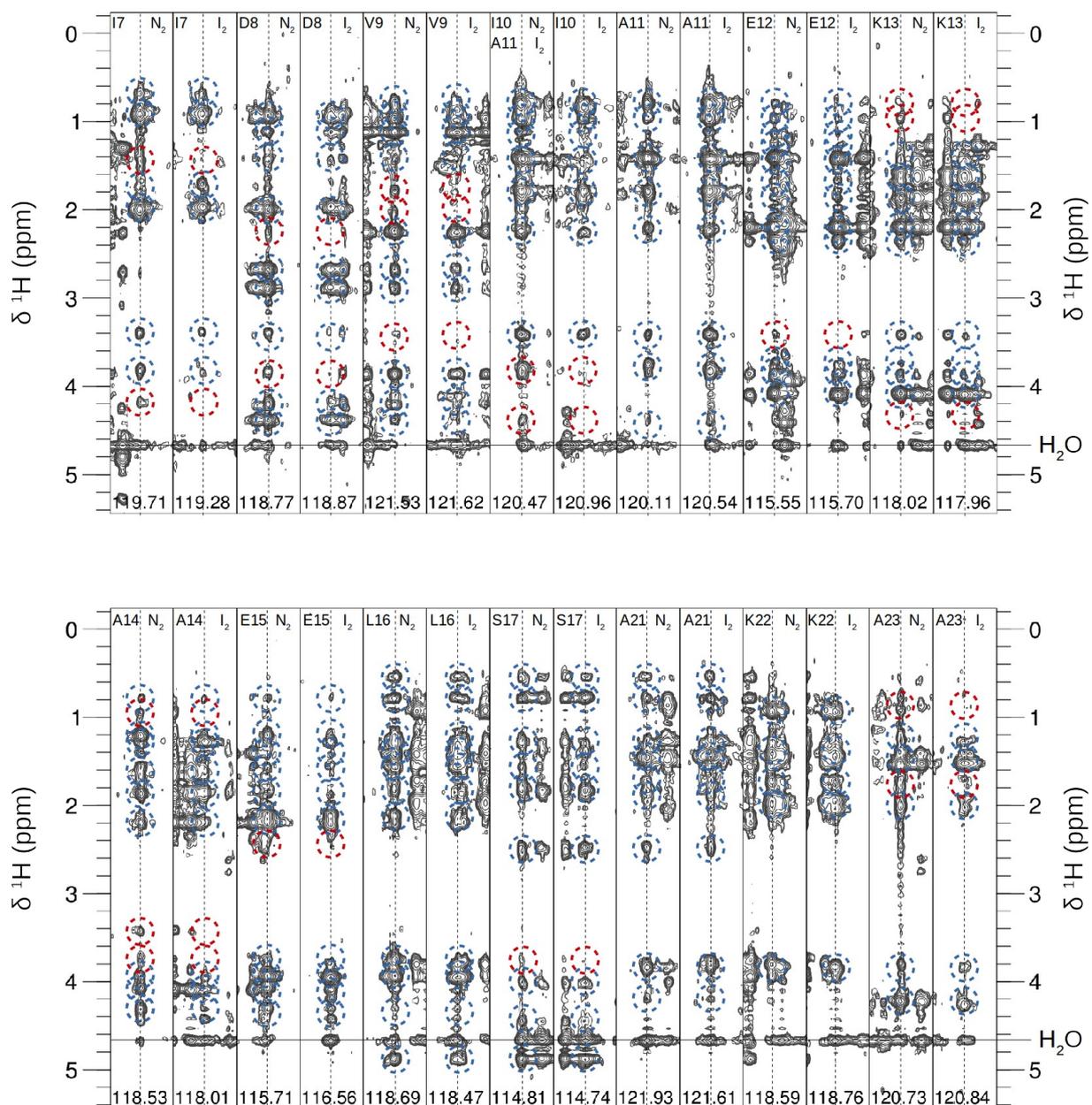


Figure 75: Comparaison des NOE $H_N-H_{\text{aliphatique}}$ des conformations N_2 et I_2 de $HU\alpha_2$ (1/2)

Des bandes extraites du spectre ^{15}N -HSQC-NOESY de $HU\alpha_2$ à 310K sont comparées entre les conformations natives et intermédiaires en limitant les régions présentant trop de superpositions. Les déplacements chimiques azote des bandes sont indiqués en bas de celles-ci. Un trait pointillé indique le déplacement chimique H_N du résidu considéré. Les cercles pointillés bleus mettent en évidence des NOEs conservés entre les deux formes. Les cercles rouges mettent en évidence des NOEs différents entre les deux formes qui correspondent le plus souvent à une disparition du signal.

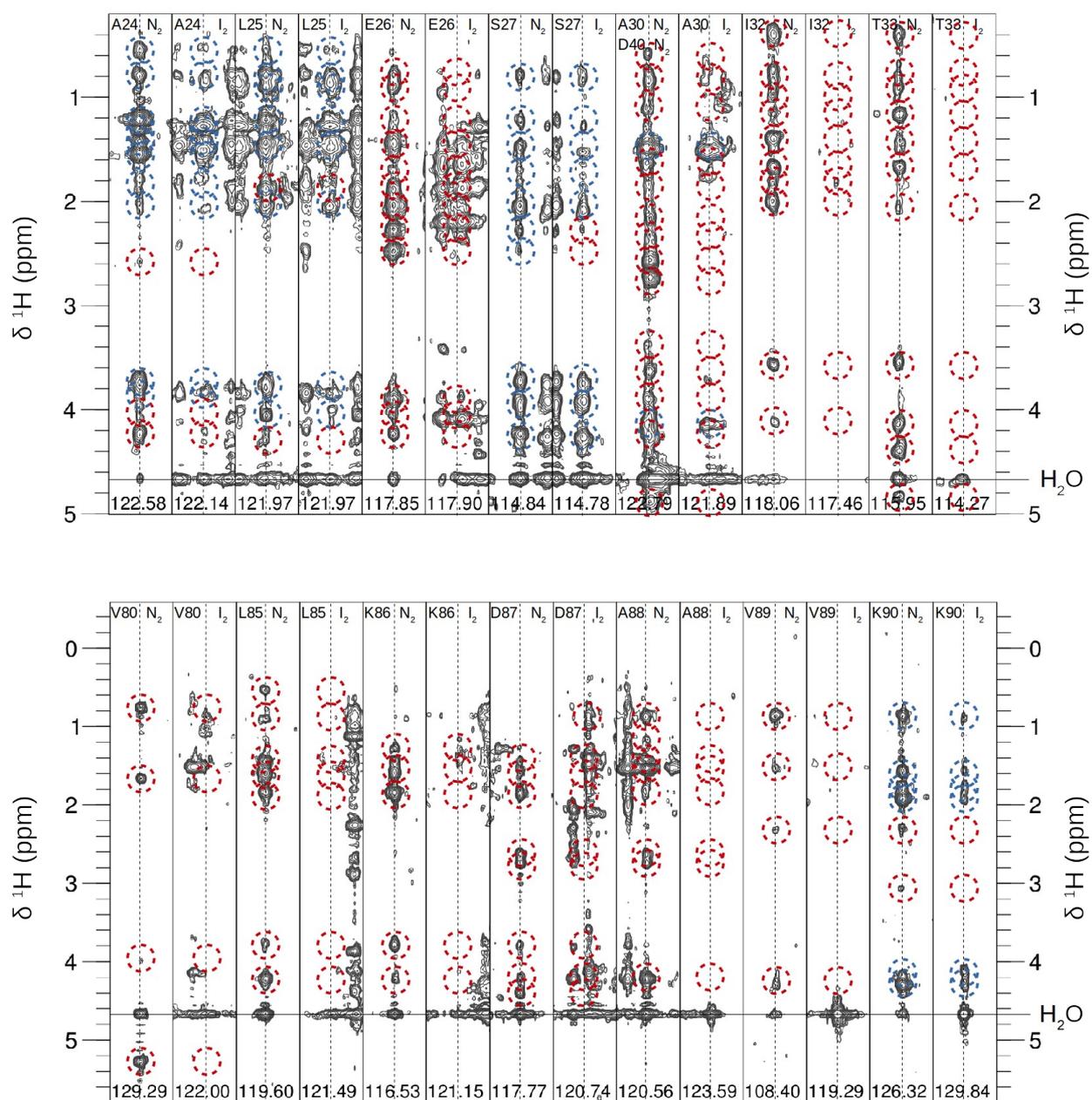


Figure 76: Comparaison des NOE H_N - $\text{H}_{\text{aliphatique}}$ des conformations N_2 et I_2 de $\text{HU}\alpha_2$ (2/2)

voir légende figure 75.

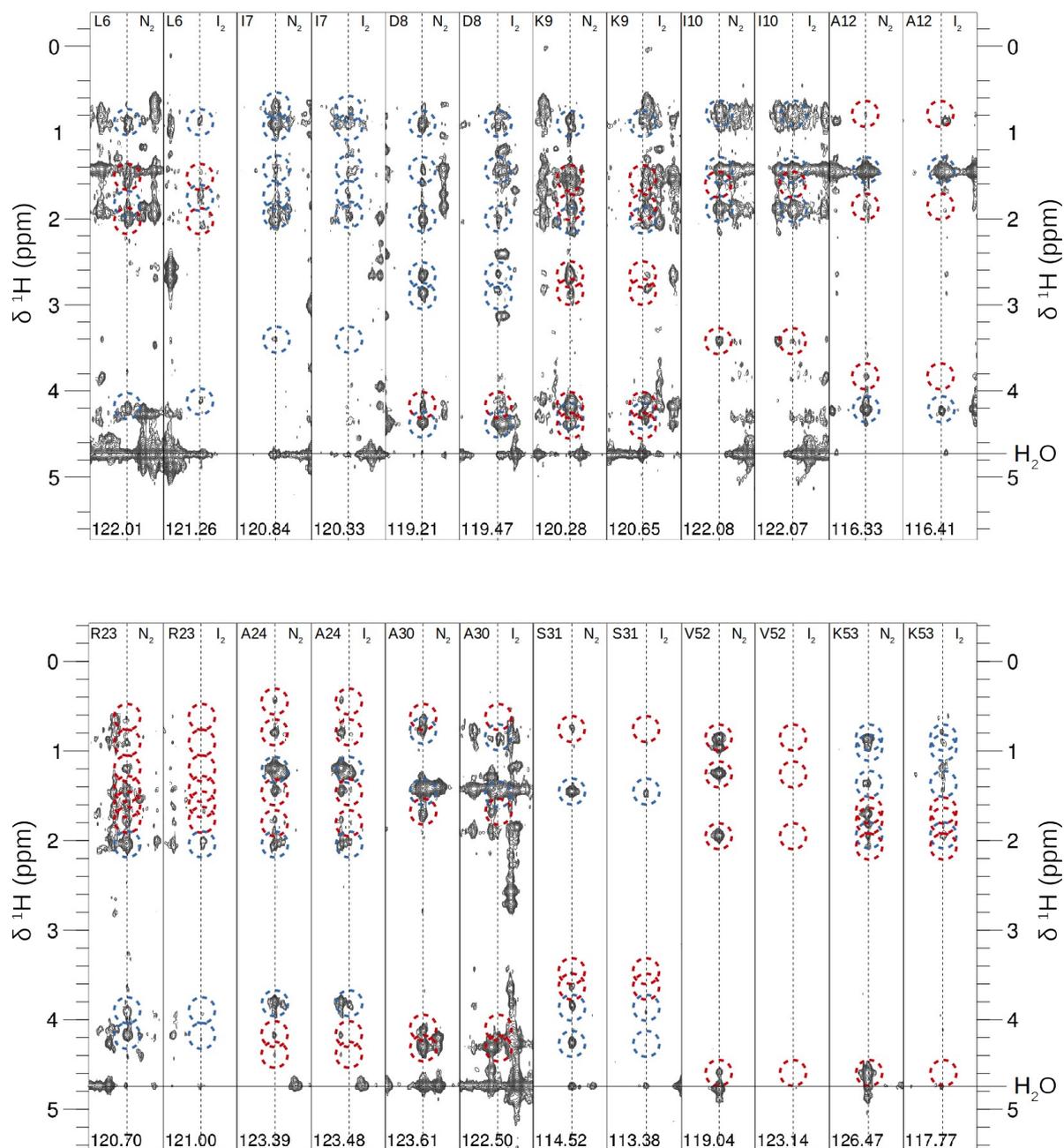


Figure 77: Comparaison des NOE H_N - $H_{\text{aliphatique}}$ des conformations N_2 et I_2 de $HU\beta_2$ (1/2)

Des bandes extraites du spectre ^{15}N -HSQC-NOESY de $HU\beta_2$ à 300K sont comparées entre les conformations natives et intermédiaires en limitant les régions présentant trop de superpositions. Les déplacements chimiques azote des bandes sont indiqués en bas de celles-ci. Un trait pointillé indique le déplacement chimique H_N du résidu considéré. Les cercles pointillés bleus mettent en évidence des NOEs conservés entre les deux formes. Les cercles rouges mettent en évidence des NOES différents entre les deux formes qui correspondent le plus souvent à une disparition du signal.

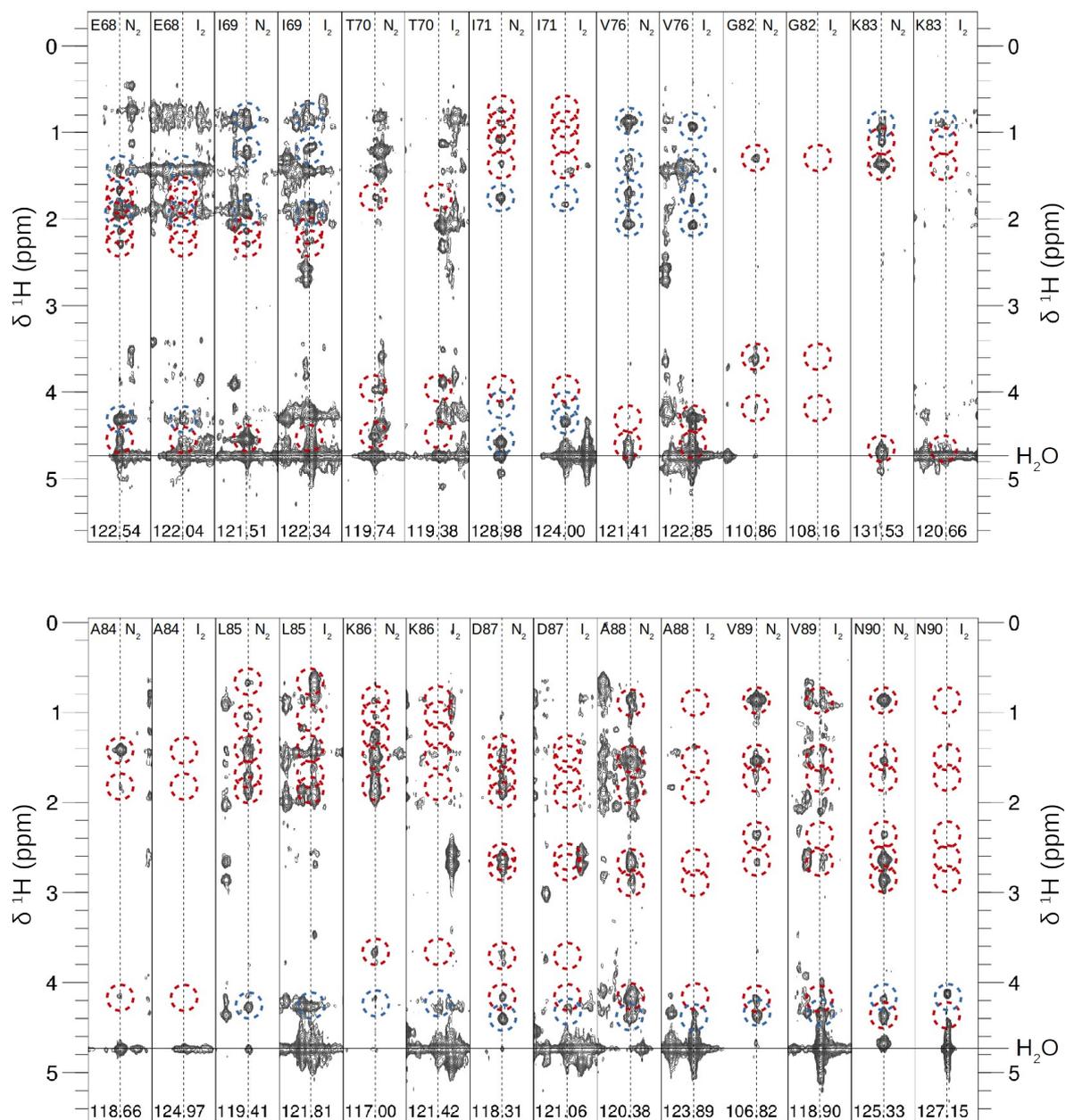


Figure 78: Comparaison des NOE $H_N-H_{\text{aliphatique}}$ des conformations N_2 et I_2 de $HU\beta_2$ (2/2)

Voir légende figure 77.

On observe un profil assez similaire pour HU β_2 mais avec des disparitions de NOEs plus nombreuses que pour HU α_2 . La région où la perte de NOE est la plus importante est la région C-terminale (résidus 84 à 90 **Figure 78**) traduisant une déstructuration importante. Au niveau du HTH l'hélice α_1 présente des profils NOEs assez conservés entre la forme N $_2$ et I $_2$ de HU β_2 (Figure 77), cependant l'hélice α_2 présente des disparitions de NOEs plus importantes que celles observées pour HU α_2 .

Les spectres disponibles présentent un nombre de NOEs insuffisant pour obtenir une structure expérimentale résolue de la conformation I $_2$ en solution, en particulier pour les régions C-terminales des homodimères. Dans le cas de la partie HTH, de faibles variations de déplacement chimique entre I $_2$ et N $_2$ des H_{aliphatiques} (et des carbones aliphatiques) rendent ambiguë l'attribution des expériences ^1H -NOESY et ^{13}C -HSQC-NOESY. On ne bénéficie pas, pour ces expériences, de séparation des résonances N $_2$ et I $_2$ par le H_N, comme pour l'expérience ^{15}N -HSQC-NOESY.

L'attribution maximale des spectres NOESY nous a permis d'obtenir de précieuses informations structurales sur la conformation I $_2$. Pour obtenir une description structurale plus poussée de la conformation intermédiaire des homodimères, des expériences de simulation par dynamique moléculaire ont été réalisées.

5.2. Étude de la transition $N_2 \rightarrow I_2$ en simulation de dynamique moléculaire dite « en gros grain »

Nous avons cherché à observer la transition vers l'état intermédiaire des homodimères à partir de leur conformation native par simulation de dynamique moléculaire. Les expériences de RMN nous ont permis de mettre en évidence que les homodimères de HU de *E. coli* effectuent cette transition à une échelle de temps de l'ordre de la milliseconde à 310K. Dans ces conditions, on peut comprendre que la simulation classique (tous les atomes explicites ; simulation sur quelques nanosecondes) requiert une très haute température pour simuler cette transition. Des températures atteignant 500K ont été utilisés dans le cas de la simulation de la transition de HU β_2 réalisée par Garnier et collaborateurs pour accélérer le processus.¹⁶⁷

Pour réaliser des simulations en atteignant des échelles de temps supérieures, à puissance de calcul égale, on peut utiliser la dynamique moléculaire « en gros grains ». Le principe est de modéliser un ensemble d'atomes par une sphère, un grain, qui possède les propriétés moyennes équivalentes au groupe d'atomes complet. On perd la résolution atomique, mais la puissance de calcul nécessaire pour la simulation est fortement diminuée. Cette approche permet de simuler beaucoup plus rapidement et d'accéder à des échelles de temps de l'ordre de la microseconde voir la milliseconde. On peut ainsi simuler des changements conformationnels de plus grande ampleur qu'en dynamique moléculaire classique.

Nous avons utilisé le champs de force « Martini » développé par Marrink et collaborateur. Bien que ce champs de force a été initialement mis au point pour simuler des lipides,^{232,195} le travail de Monticelli et collaborateurs a permis d'ajouter les protéines aux biomolécules étudiables avec ce champ de force.¹⁹⁴ Chaque acide aminé est représenté par un grain représentant le squelette peptidique et zéro à quatre grains représentant la chaîne latérale (**Figure 79**). Bien que le modèle Martini soit basé sur une réduction par quatre en moyenne entre le nombre d'atomes et le nombre de grains, ce facteur est réduit jusqu'à 2 dans le cas des acides aminés aromatiques pour en reproduire la géométrie. Chaque grain présente des propriétés différentes, de polarité notamment, afin de reproduire le plus fidèlement possible les interactions des acides aminés tout atome (**Figure 79**). L'eau dispose également d'un modèle gros grains, où un grain représente 4 molécules d'eau.

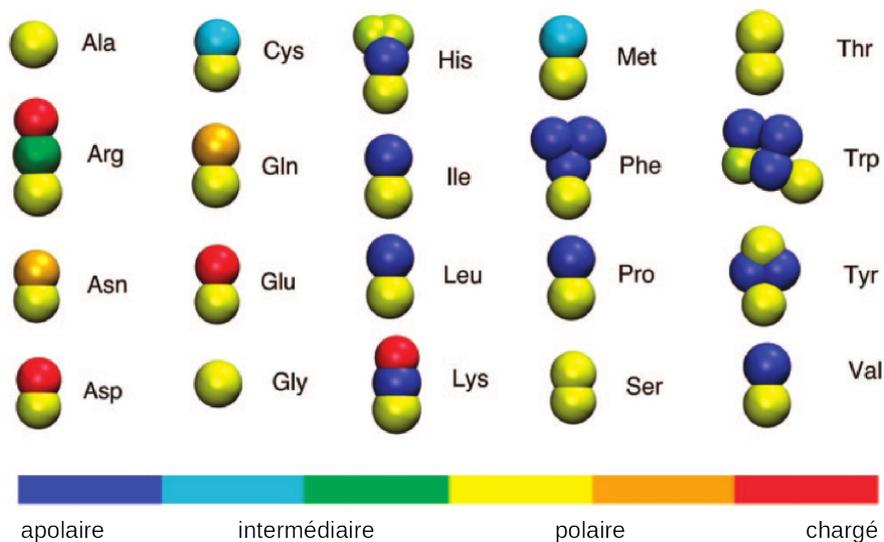


Figure 79: Représentation en « gros grains » des acides aminés dans le champ de force Martini (Adapté de Monticelli et collaborateurs)¹⁹⁴

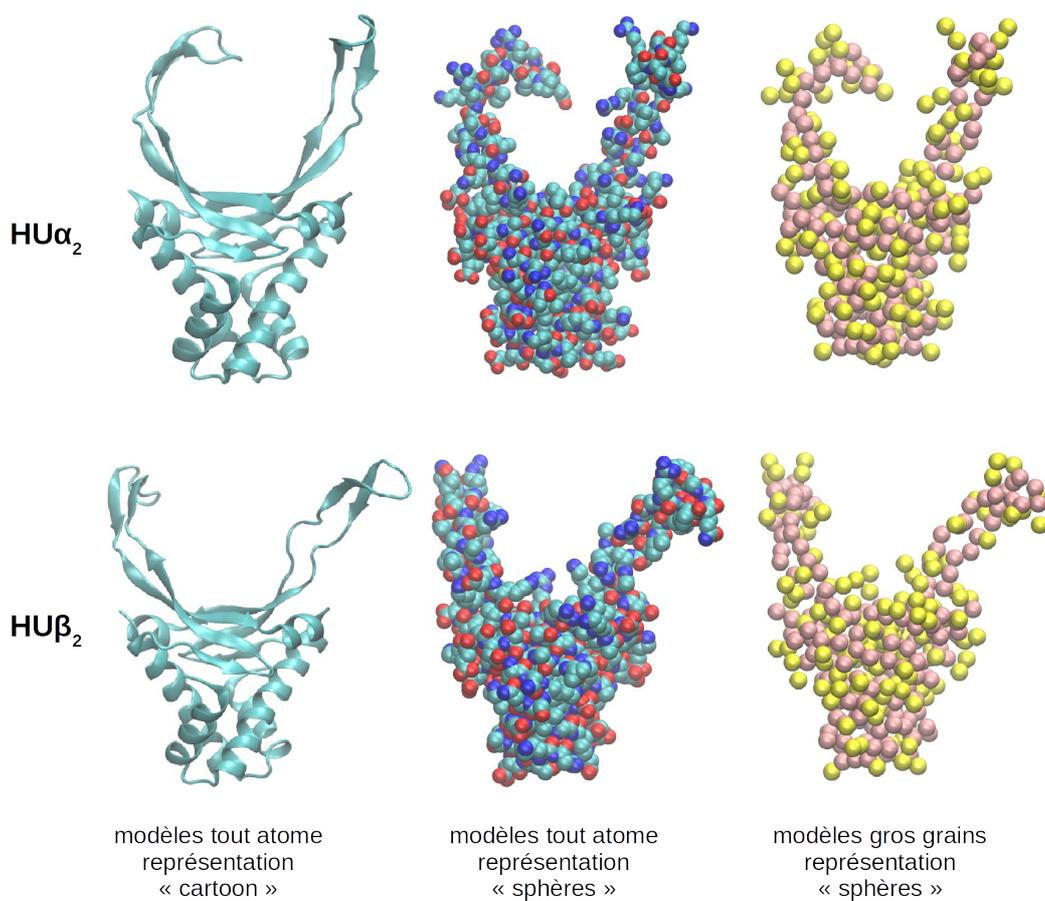


Figure 80: Modèles « tout atome » et « gros grains » des homodimères

Des modèles « gros grains » ont été générés à partir des modèles structuraux « tout atome » des homodimères. On obtient une réduction du nombre d'atomes en grains de 1340/1296 à 362/346 pour HU α_2 /HU β_2 respectivement (sans compter les hydrogènes). Avec un faible nombre de grains, nous avons pu réaliser des simulations de dynamique moléculaire des homodimères sur une durée de 10 μ s. Nous avons ainsi pu gagner trois ordres de grandeur en temps par rapport aux simulations « tout atome » classiques (de l'ordre de la dizaine de nanosecondes pour des temps de calcul équivalents). Pour s'approcher de l'échelle de temps de la milliseconde caractéristique de l'échange N $_2$ \rightarrow I $_2$, nous avons décidé de réaliser ces simulations à des températures assez élevées (350 et 370 K) qui demeurent plus raisonnables que celles utilisées pour les simulations « tout atome » sur HU β_2 (500 K).¹⁶⁷ Nous espérons que l'échelle de temps atteinte soit suffisante pour observer les premiers événements qui conduisent aux mouvements moléculaires de grande ampleur observés par RMN.

Il est important de noter que dans le champs de force Martini, les structurations secondaires sont bloquées. Même à haute température, aucune déstructuration de structure secondaire n'est possible. Nous avons fait l'hypothèse que les premiers événements à observer dans la transition étaient des mouvements de sous-domaines de la protéine ne requérant pas de déstructuration d'éléments de structure secondaire (mouvement des hélices α les unes par rapport aux autres par exemple)

5.2.1. Analyse des trajectoires

Les modèles « gros grains » ne permettent pas la visualisation aisée des structures de protéines au cours des trajectoires de simulation. Avec les outils de visualisations disponibles, seules des représentations en sphères ou en bâtons sont disponibles et permettent très difficilement d'apprécier qualitativement les changements de conformations. Pour visualiser les changements d'orientations et les mouvements des sous-domaines des homodimères, différentes distances et angles ont été mesurés sur les trajectoires des simulations.

Mouvements des hélices du HTH

Les changements de structure du HTH sont suivis par la mesure d'une série de distances résumées dans la **Figure 81**. Les variations de ces distances au cours des simulations sont présentées graphiquement sur la **Figure 82**. La mesure des variations d'angles entre les axes des hélices est également présentée **Figure 83**.

Dans les simulations réalisées sur $HU\alpha_2$, l'effet de la température est assez modéré. Les variations de distances et d'angles présentent des amplitudes comparables sur la plupart des graphes. Une variation marquante est observable en fin de simulation (9 à 10 μ s) à 350 K, au niveau des distances K3-L44', M1-D40' et K3-A25 (**Figure 82**), ainsi que sur l'angle entre les axes des hélices α_1 et α_2 (**Figure 83 en haut**). Ces variations correspondent à un changement d'orientation de l'hélice α_1 qui s'éloigne par rapport au reste du HTH puis reprend sa position initiale en fin de simulation.

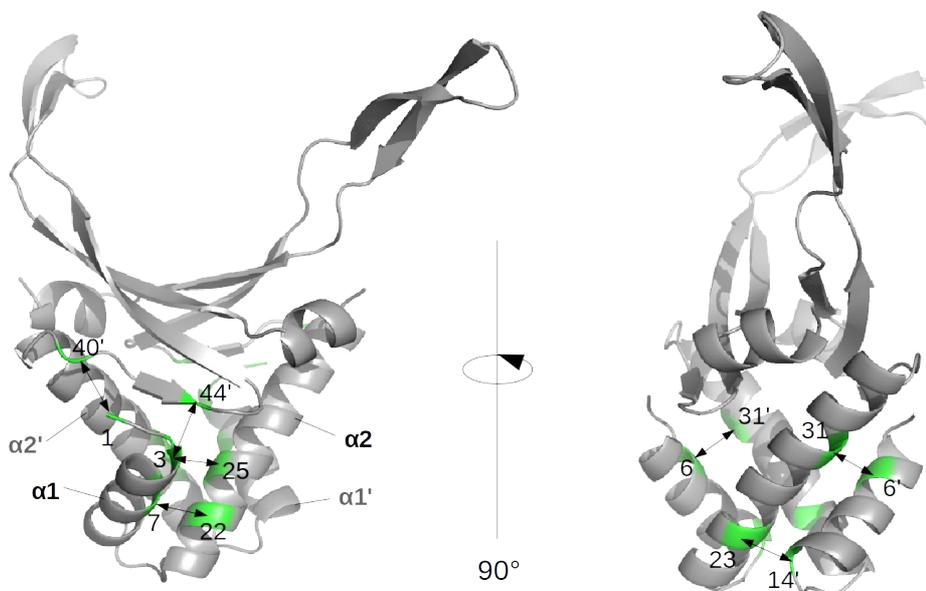


Figure 81: Distances mesurées pour le suivi des mouvements des hélices HTH Les numéros indiquent la position des résidus dans la séquence primaire. Le ' indique qu'un résidu appartient à la seconde chaîne protéique de l'homodimère.

Ce mouvement est également observé de manière plus modérée entre 3 et 4 μ s de simulation où l'augmentation de ces deux mêmes distances et la variation du même angle sont observées. A 370K, la simulation de $HU\alpha_2$ présente un changement de conformation qui a lieu vers 5 μ s. Il se traduit par une augmentation des distances A23-A14', A31-L6', D40-M1' et L44-K3' de l'ordre de 5 Å. Ces distances traduisent un éloignement global de l'hélice α_1 du reste du HTH et s'accompagne d'une variation mineure de l'angle entre les axes d'hélices α_1' et α_2' ($\sim 20^\circ$ **Figure 83**). Ce changement de conformation est assez visible sur l'une des chaînes et est mineur sur l'autre chaîne. Cependant, deux distributions d'angles sont effectivement visibles entre les hélices α_1 et α_2 des deux chaînes à 370 K (**Figure 83**). L'effet de cette transition sur l'arrangement des domaines HTH de chaque chaîne l'un par rapport à l'autre se traduit par une augmentation de la distance (~ 5 Å) visible sur les résidus équivalents des hélices α_2 (**Figure 84**) qui n'entraîne qu'une variation mineure des angles entre les axes de ces hélices α_2 (**Figure 83**).

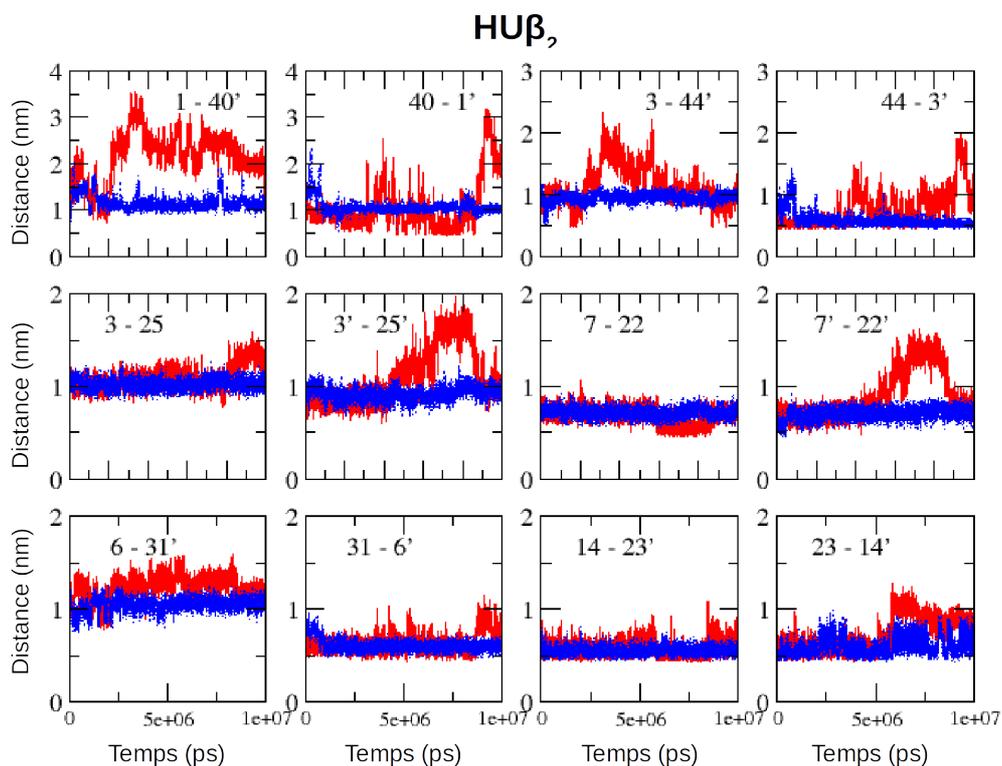
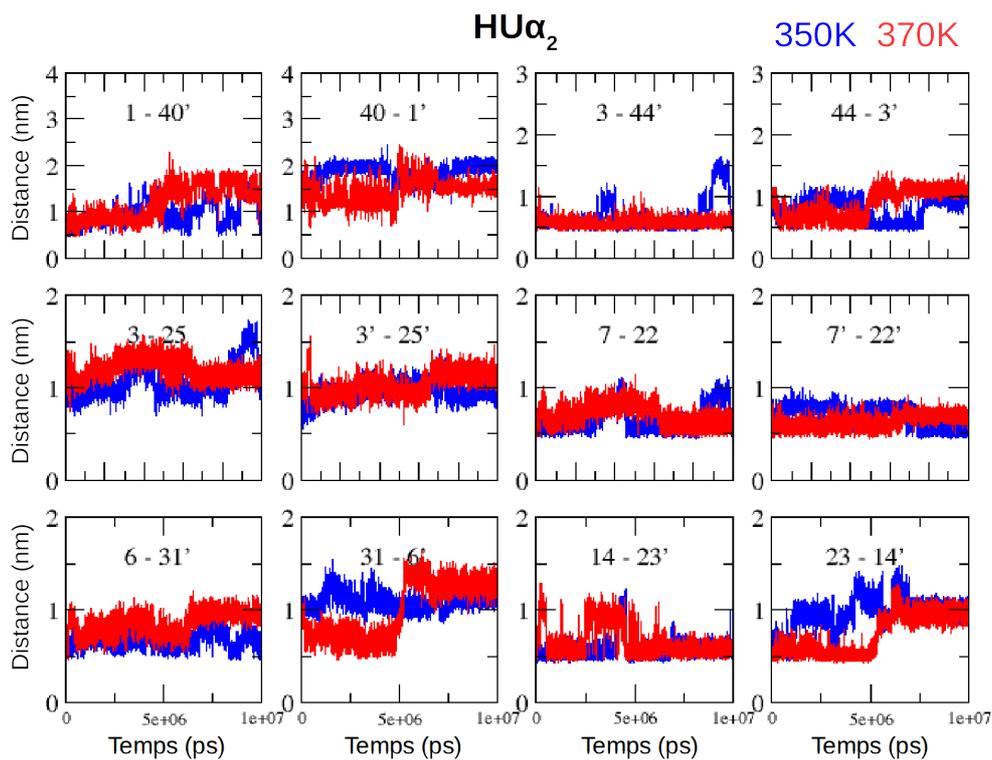
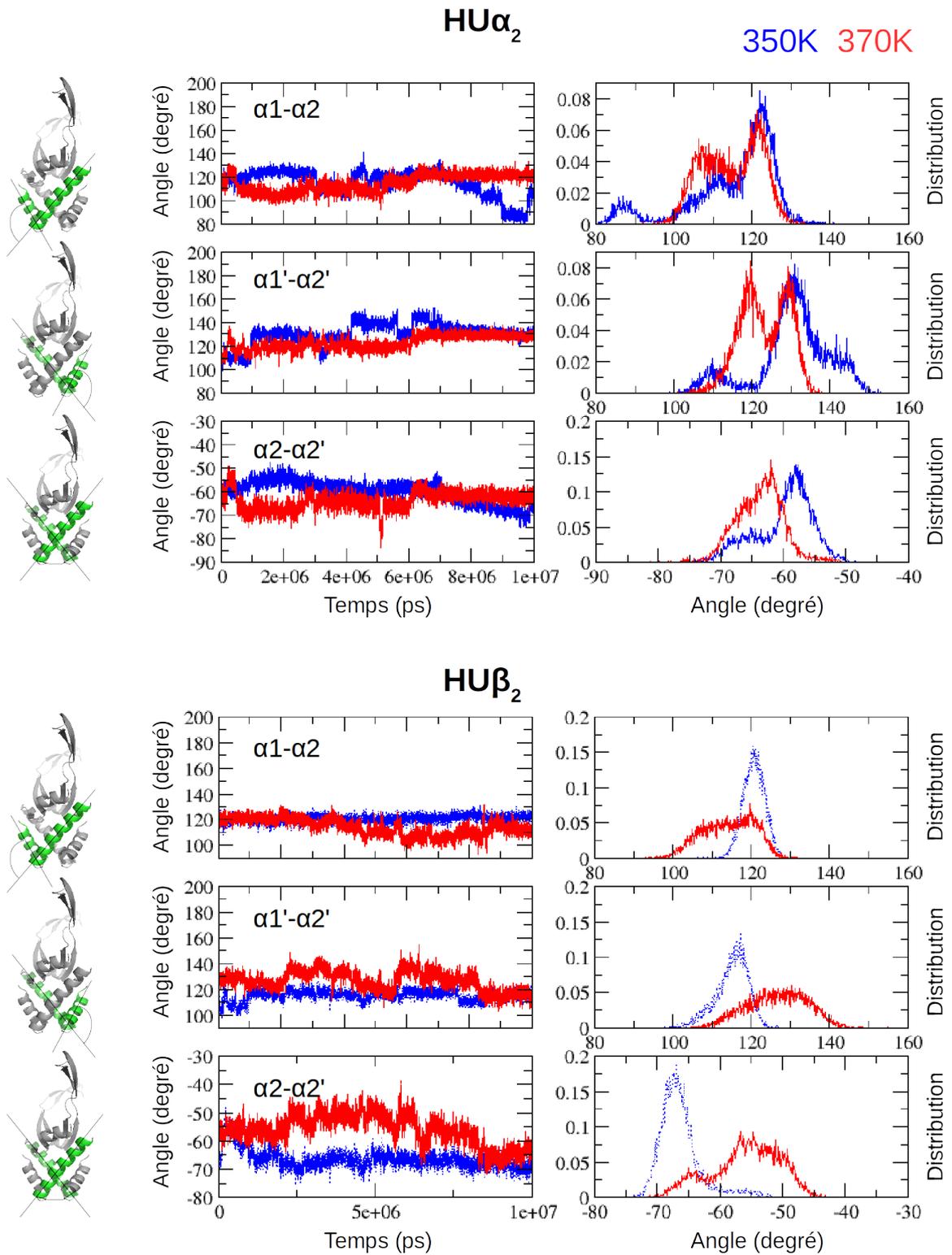


Figure 82: Variation des distances pour le suivi du mouvement des hélices au cours des simulations Les numéros indiquent la position des résidus dans la séquence primaire. Le ' indique qu'un résidu appartient à la seconde chaîne protéique de l'homodimère.



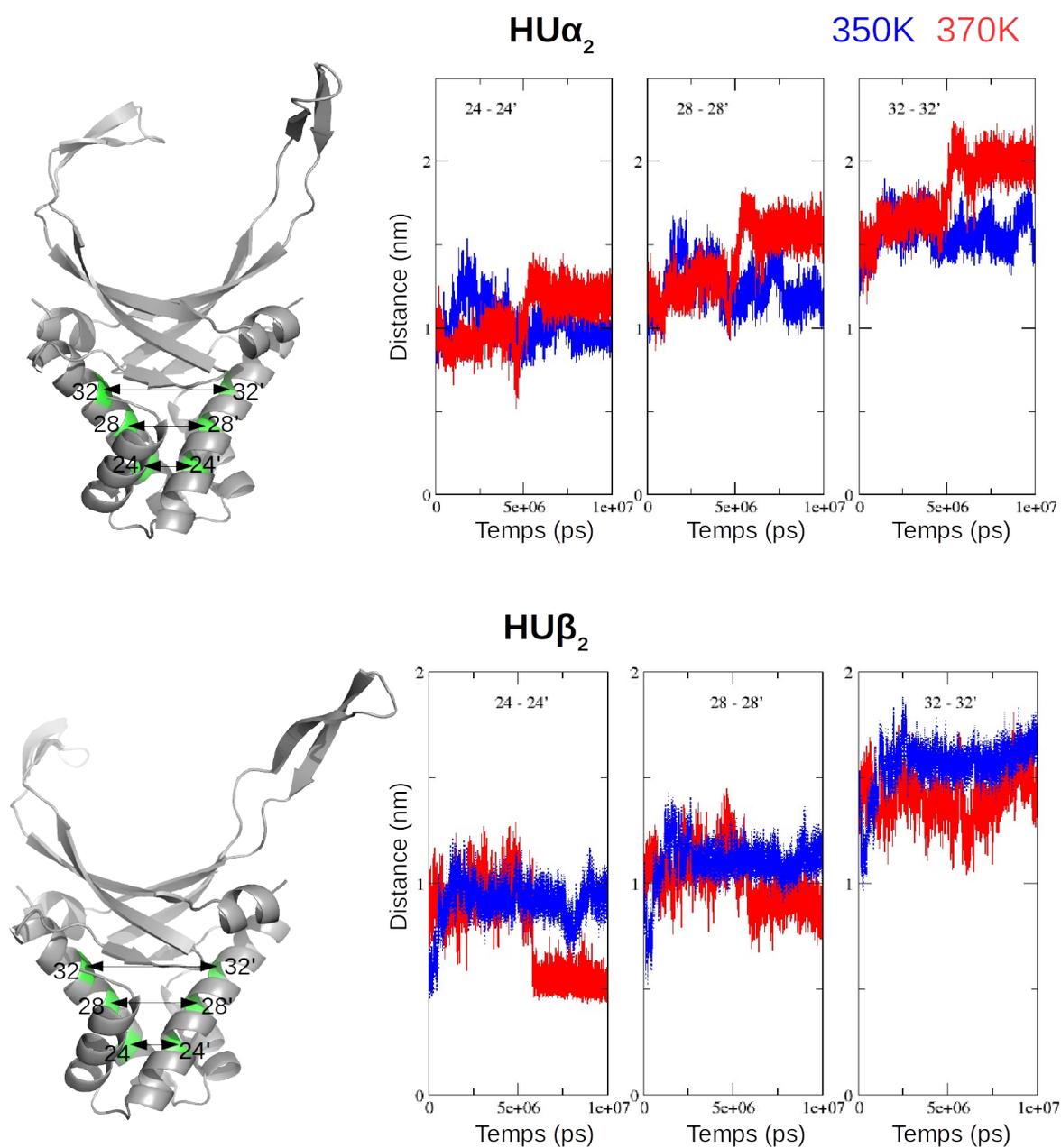


Figure 84: Variation des distances entre les hélices α_2 au cours des simulations

Pour $HU\beta_2$, l'effet de la température sur les amplitudes de variation des distances et des angles est plus important que sur $HU\alpha_2$. On observe peu de variations à 350 K avec des augmentations de distances les plus importantes de l'ordre de 5 Å (**Figures 82 et 84**) et les angles des hélices sont distribués autour d'une unique gaussienne (**Figure 83**). On observe beaucoup plus de variations à 370K. Des fluctuations des distances entre les extrémités N-terminales et la jonction entre l'hélice 2 et le brin β_1 sur l'autre chaîne (M1-D40', 40-1', K3-L44' et L44-K3' **Figures 82**) sont visibles sur la quasi-totalité de la simulation. Par ailleurs, un mouvement d'hélices est visible entre 6 et 8 μ s. On observe une augmentation des distances 3'-25' et 7'-22' de presque 10 Å qui correspond à un éloignement entre les hélices 1 et 2 sur l'une des chaînes. Cet éloignement s'accompagne d'une variation des angles entre les hélices qui fluctuent dans une gamme d'environ 20° (**Figure 83**). Ces variations d'angles se répercutent sur les distances entre les hélices 2 (**Figure 84 en bas**) où l'on observe une diminution de distance au niveau du bas de ces hélices (24-24'). L'ampleur de cette diminution s'amointrit si l'on mesure les distances entre des résidus plus haut sur ces hélices (I28-I28' et I32-I32').

Pour résumer, on peut observer des mouvements d'hélices dans les simulations de $HU\alpha_2$ et $HU\beta_2$. Ces mouvements présentent des amplitudes relativement restreintes avec des variations d'angles entre les axes des hélices de l'ordre de 20° (40° au maximum) et des variations de distances de l'ordre de 5 Å et 10 Å pour la plupart. Il est possible que l'ampleur de ces mouvements soit restreints par l'absence de dénaturation de l'hélice C-terminale à l'interface entre l'hélice α_2 et le feuillet β .

Stabilité de l'hélice C-terminale

Nous avons vu dans les expériences de RMN décrites plus tôt dans ce chapitre que l'hélice α_3 C-terminale perdait sa structuration secondaire dans la conformation I_2 . Une des limitations de la simulation en « gros grains » vient du fait que les structurations secondaires sont bloquées. Nous avons donc voulu voir si des mouvement étaient possibles entre le début du bras, l'hélice C-terminale et la fin de l'hélice α_2 . Les résultats sont présentés **Figure 85**.

Pour $HU\alpha_2$ les fluctuations de distances sont assez faibles et on observe plutôt une diminution de cette distance dans les simulations à 350 K. Cette région reste donc assez stable au cours des simulations de ce dimère. Dans le cas de $HU\beta_2$, on observe des augmentations de distances significatives pour une seule chaîne et seulement à 370 K (distances N90-K37' et A88-T33' entre 3

et 4 μ s). Cette augmentation de distance est maintenue et précède le mouvement de la partie HTH de la chaîne correspondante (distances 3'-25' 7'-22' entre 6 et 8 μ s **Figure 82**). On peut donc imaginer que la dénaturation (ou l'éloignement) de l'hélice C-terminale soit un facteur important pour augmenter le degré de liberté des mouvements du HTH.

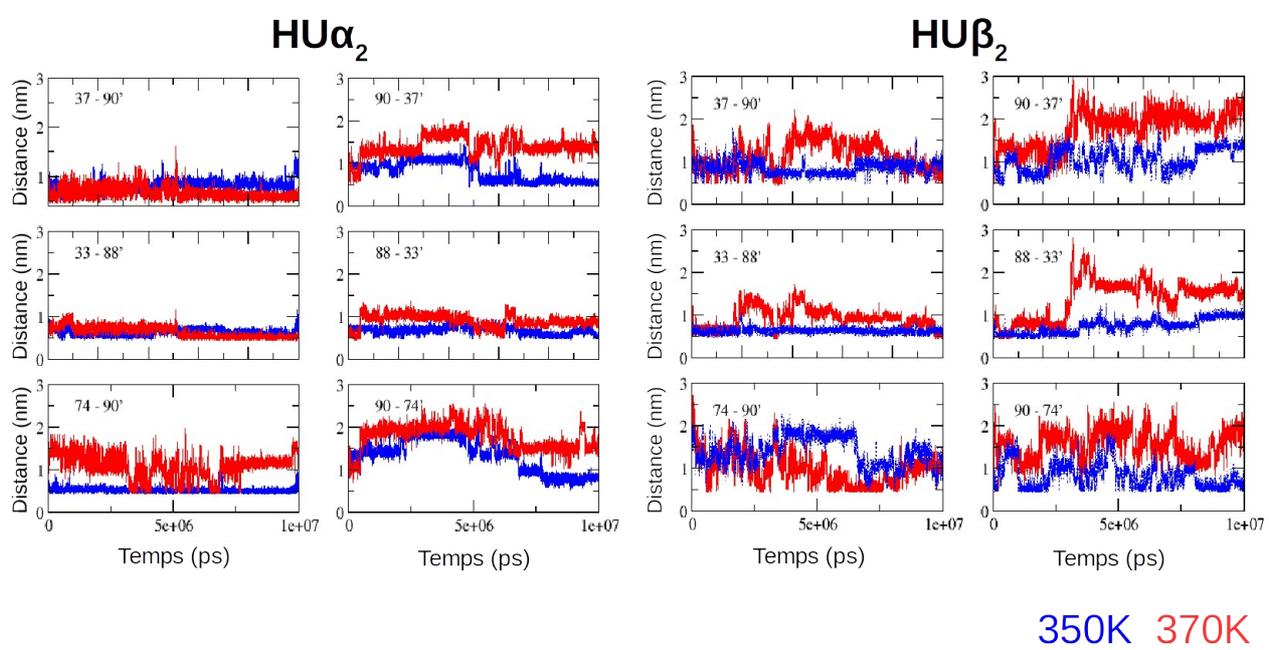
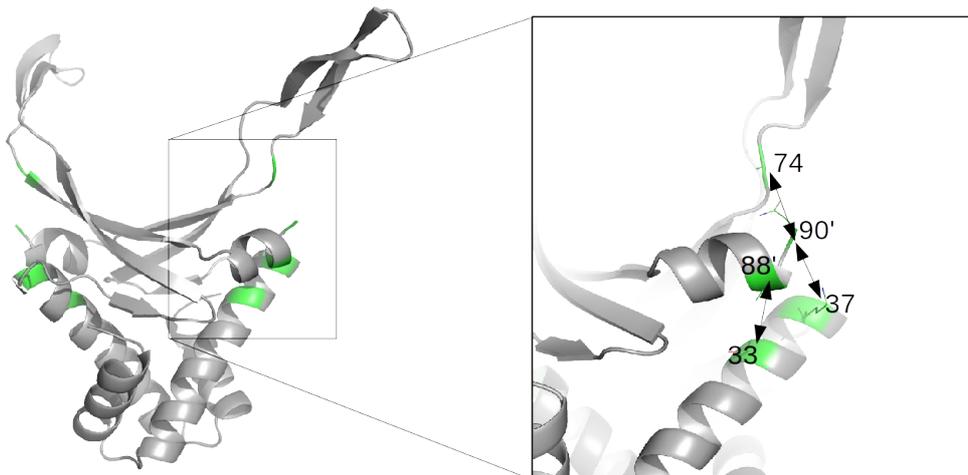


Figure 85: Stabilité de l'hélice α_3 par rapport au début du bras et à l'hélice α_2

Conclusion

Pour conclure sur ces expériences de simulation, nous nous heurtons aux limites actuelles de la dynamique en « gros grains » dans ce système. Si la rapidité de simulation en « gros grain » nous a permis de gagner trois ordres de grandeurs dans la gamme de temps simulée (on passe de la dizaine de nanosecondes en « tout atome » à la dizaine de microsecondes en « gros grains ») nous n'avons pas pu atteindre l'échelle de temps de la milliseconde où s'opère la transition $N_2 \rightarrow I_2$. Pour essayer d'augmenter la vitesse d'échange, nous avons augmenté la température des simulations (350K et 370K). Cependant, comme pour la majorité des champs de force utilisés en dynamique moléculaire, MARTINI a été paramétrisé dans une gamme de température restreinte (270 – 330 K). Les simulations en dehors de cette gamme doivent être interprétées avec précaution.¹⁹⁵ De plus, nous avons fait l'hypothèse que les premiers mouvements ne requièrent pas de déstructuration des structures secondaires (champs de force MARTINI).

Nous avons pu effectivement observer des mouvements de faible ampleur des hélices α au cours des simulations. Il est possible que l'absence de dénaturation de l'hélice α_3 C-terminale limite le degré de liberté du HTH, notamment au niveau de l'hélice α_2 . De plus, l'échelle de temps explorée dans ces simulations (μ s) est insuffisante pour la simulation de la transition $N_2 \rightarrow I_2$ qui opère à une échelle de temps de l'ordre de la milliseconde.

Si on extrapole les éloignements des hélices et les variations d'angles, on peut imaginer une ouverture du HTH rendant accessible le cœur hydrophobe de la protéine (**Figure 86 A**). D'après les données de RMN, il n'y a pas de modification des structures secondaires du HTH et les perturbations de déplacement chimique (**résidus en jaune Figure 86 B**) sont probablement dues aux modifications d'environnement chimique et électronique suite à son ouverture. Les résidus de l'hélice α_3 C-terminale et du brin β_3 subissent une perte de structure secondaire (**résidus en rouge Figure 86 B**) qui est possiblement liée à la perte des contacts avec la fin de l'hélice α_2 du HTH qui s'en éloigne dans ce modèle.

Pour aller plus loin dans la description de l'état intermédiaire, une approche intéressante serait d'utiliser la dynamique moléculaire couplée à des contraintes RMN. Ce type d'expériences a été réalisé dans une étude récente sur HU pour la description de l'espace conformationnel exploré par les bras pour la fixation sur l'ADN.²³³ Par ailleurs, cette approche a également été employée pour la résolution de la structure de l'état intermédiaire partiellement déstructuré du domaine SH3.²³⁴

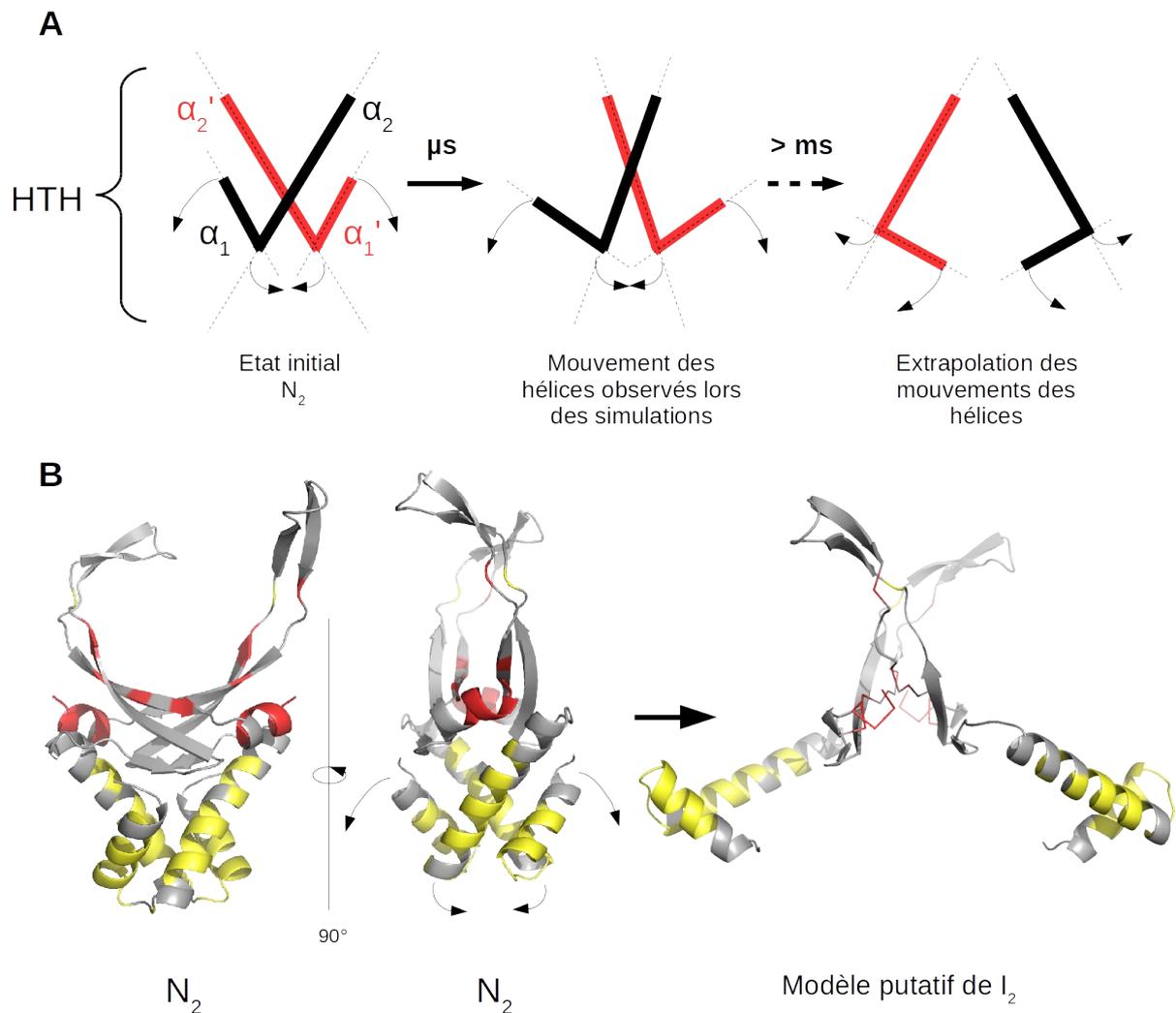


Figure 86: Modèle de la conformation I_2 déduit des expériences de RMN et en simulation de dynamique moléculaire A- Résumé des mouvements d'hélices observés dans les simulations en dynamique moléculaire et extrapolation à l'échelle de temps supérieur. B- Modèle de la transition $N_2 \rightarrow I_2$ en accords avec les résultats de RMN et de simulation. La coloration des résidus en jaune et en rouge représente les CSP $N_2 \rightarrow I_2$ de $HU\alpha_2$.

Les données RMN utilisables pour ce type d'expériences sont les déplacements chimiques, les NOEs et/ou des RDC. Au cours de la simulation ces données sont recalculées à chaque instant et une force artificielle est générée pour biaiser le champs de force (un nouveau terme énergétique est généré) afin de tendre vers un accord entre les données expérimentales et recalculées. C'est ce type d'approche qui est utilisé dans les logiciels standards de calcul de structure RMN. On peut également utiliser plusieurs modèles simulés en parallèle afin de comparer des données moyennées sur une population de structures avec les données expérimentales. On peut ainsi obtenir un meilleur

accord si un ensemble de conformations est nécessaire pour la description d'une structure. Cette approche permet d'obtenir des modèles plus réalistes en simulation, grâce à l'incorporation des données expérimentales, et permettrait de dépasser les barrières d'échelle de temps auxquelles nous nous sommes heurtés au cours de ces premières simulations de la transition $N_2 \rightarrow I_2$.

5.3. À retenir

Objectif du chapitre :
Caractériser l'état intermédiaire du point de vue cinétique et structural
- Attribution des déplacements chimiques des états intermédiaires I_2 des homodimères $HU\alpha_2$ et $HU\beta_2$ (à 310 et 300K, respectivement)
- Analyse des séries d'expériences ZZ- ^{15}N -HSQC → Évaluation des constantes cinétiques de l'échange $N_2 \leftrightarrow I_2$ (possible pour $HU\alpha_2$ seulement)
- Analyse des CSP entre conformations N_2 et I_2 , et prédiction des structures secondaires à partir des déplacements chimiques → Identification des résidus et des structures secondaires perturbés entre les conformations N_2 et I_2, pour les 2 homodimères $HU\alpha_2$, $HU\beta_2$
- Analyse des modifications des pics NOE entre les conformations N_2 et I_2 des 2 homodimères $HU\alpha_2$ et $HU\beta_2$ → Identification des zones où se produisent des pertes de structuration secondaire et/ou tertiaire.
Pour $HU\alpha_2$, la zone hélice-tour-hélice est globalement maintenue avec une conservation des structures secondaires mais une perte de la structure tertiaire. La partie C-terminale se déstructure. Le profil est le même pour $HU\beta_2$ avec une plus grande déstructuration au niveau de l'hélice 2.
- Simulation de dynamique moléculaire en gros grains (avec blocage des structures secondaires) → Visualisation des mouvements des hélices du HTH les unes par rapport aux autres. La déstructuration de l'hélice C-terminale serait un facteur important dans l'augmentation des degrés de liberté du HTH, puis l'ouverture du HTH rendrait accessible le cœur hydrophobe.
⇒ Extrapolation des données (RMN et MD) pour proposer un modèle structural de l'état intermédiaire I_2

Chapitre 6 : Le Tétramère Transitoire

Dans le mécanisme d'échange de chaînes postulé, nous avons émis l'hypothèse que l'échange pouvait s'opérer dans un hétérotétramère dans lequel les homodimères sous leurs formes I_2 interagissent de façon transitoire (**Figure 87**). Le caractère transitoire du tétramère est nécessaire pour le mécanisme d'échange de chaînes, car, dans le cas où le tétramère formerait un complexe stable en solution, l'échange de chaînes serait d'autant moins efficace.

Dans ce chapitre, nos objectifs sont de mettre en évidence la présence du tétramère de HU, et le cas échéant, d'identifier des conditions permettant de stabiliser le tétramère en vue d'une étude structurale.

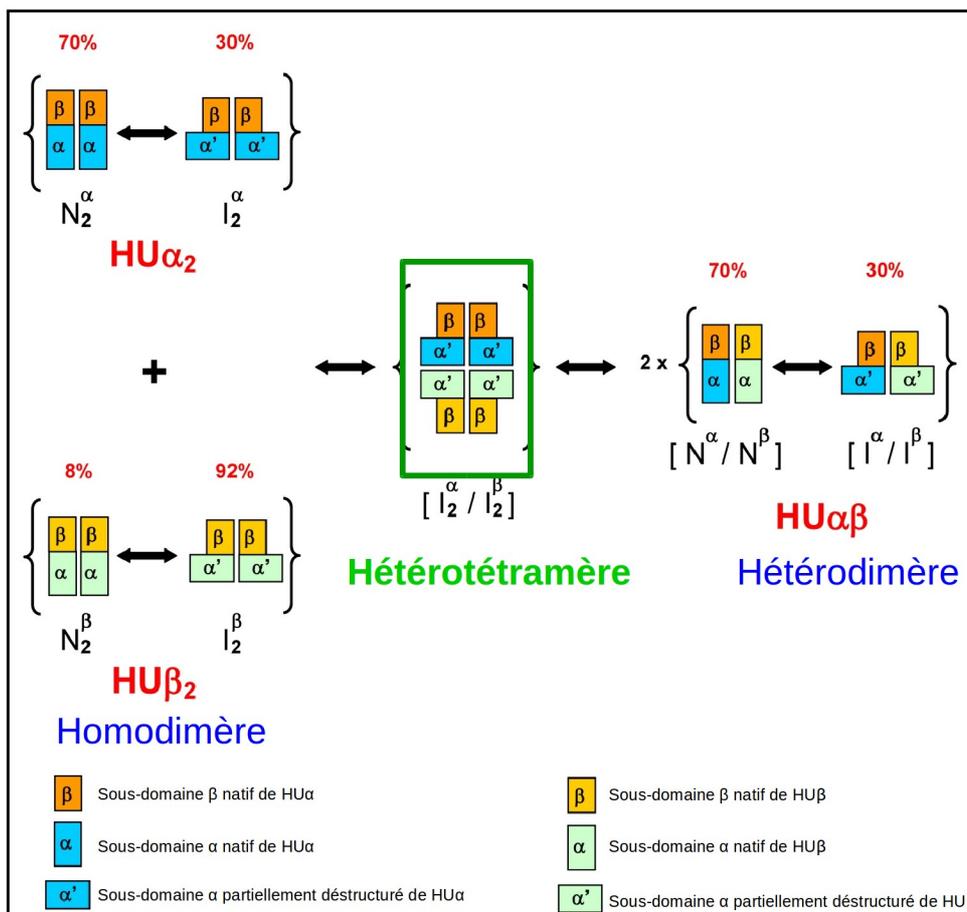


Figure 87: Hétérotétramère dans le mécanisme putatif d'échange de chaînes des dimères de HU

6.1. Mise en évidence d'un tétramère par pontage chimique

Des oligomères de HU d'ordre supérieur aux dimères ont déjà été mis en évidence dans des études antérieures par pontage chimique.¹⁵⁷ Pour déterminer les états d'oligomérisation des isoformes de HU en solution, les auteurs de ces travaux ont utilisé un réactif chimique, le diméthyl suberimidate (DMS), capable de ponter covalentement deux fonctions amines primaires relativement proches dans l'espace. Après pontage, les mélanges réactionnels ont été analysés par SDS-PAGE (**Figure 88**).

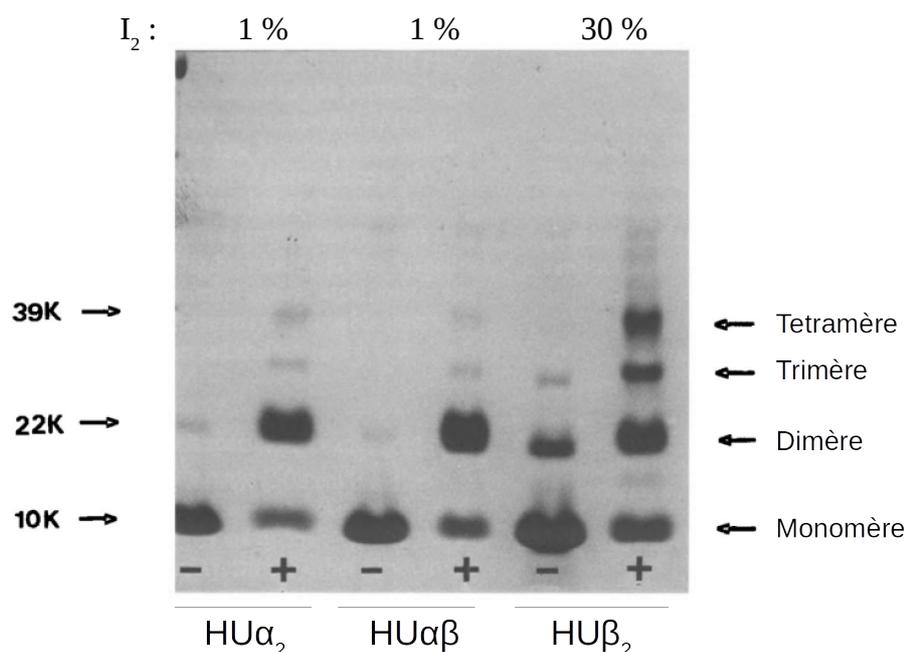


Figure 88: Pontage chimique des différentes espèces de HU de *E. coli* par le diméthyl suberimidate suivi par SDS-PAGE (Adapté de Rouvière-Yaniv J et Kjeldgaard NO)¹⁵⁶ Les poids moléculaires et les migrations des marqueurs de tailles utilisés sont repérés à gauche du gel (piste non visible). La présence (+) ou l'absence (-) d'agent pontant est indiquée en bas de chaque piste. L'abondance relative de I_2 (en %) déduite des expériences de DSC¹⁶⁶ est indiquée pour chaque dimère au dessus du gel.

Si l'on compare chaque piste de dimères pontés (+), on remarque que $HU\beta_2$ semble avoir tendance à former plus facilement des oligomères d'ordre supérieur aux dimères que $HU\alpha_2$ ou $HU\alpha\beta$. Cette observation est cohérente avec le rôle supposé de la conformation intermédiaire I_2 dans le modèle que nous avons proposé. Dans les conditions de pontage utilisées dans ces travaux, à température ambiante, $HU\beta_2$ ($T_{m1} = 27^\circ\text{C}$) présente une proportion de conformation intermédiaire bien plus grande que $HU\alpha_2$ et $HU\alpha\beta$ ($T_{m1} = 41^\circ\text{C}$ et 38°C respectivement) (**Figure 88**). On note cependant

deux problèmes sur ce gel : des formes pontées apparaissent même en absence de dimethyl suberimidate (**pistes « - » Figure 88**) et la quantité de protéine totale visible sur les pistes du gel semble supérieure pour le HU β_2 par rapport à HU α_2 ou HU $\alpha\beta$.

Pour pouvoir comparer, dans de bonnes conditions, le pontage des différents dimères de HU, nous avons réalisé des expériences comparables à celles de Rouvière-Yaniv J et Kjeldgaard NO mais en utilisant le glutaraldéhyde (disponible au laboratoire) comme agent pontant. Tout comme le dimethyl suberimidate, le glutaraldéhyde ponté les fonctions amines primaires des protéines.

6.1.1. Pontage des dimères de HU à concentration variable en glutaraldéhyde

Dans un premier temps, nous avons déterminé la concentration optimale d'agent pontant nécessaire pour l'observation des formes oligomériques d'ordre supérieur à deux. Des réactions de pontage ont été réalisées en faisant varier la concentration en glutaraldéhyde. Les mélanges réactionnels sont ensuite analysés par SDS-PAGE (**Figure 89**).

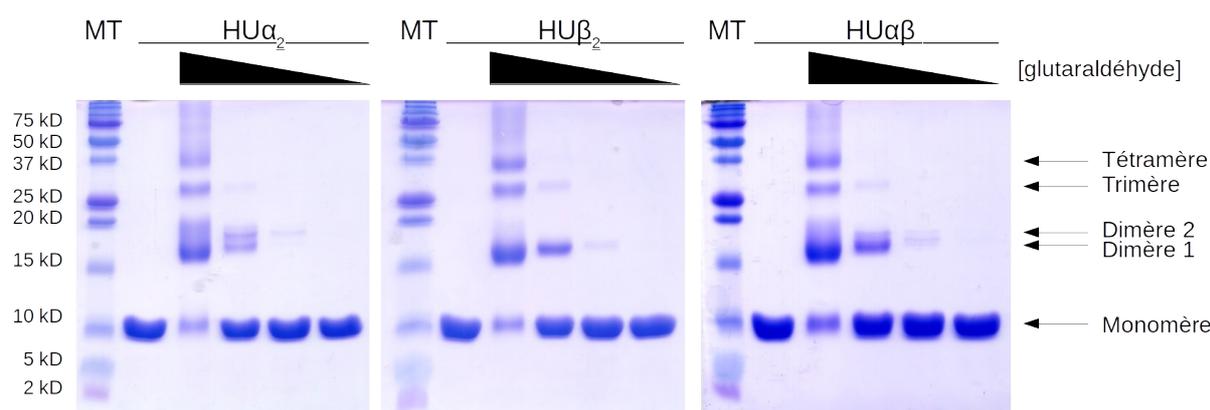


Figure 89: Analyse SDS-PAGE du pontage des dimères de HU à 37°C à concentration variable en glutaraldéhyde La concentration en protéine est fixée à 1 mg/mL et la concentration en glutaraldéhyde est de gauche vers la droite de : 0 ; 0,01 ; 0,001 ; 0,0001 et 0,00001 % final w/w. L'abréviation "MT" désigne les marqueurs de taille

Pour chaque isoforme de HU, on observe des oligomères allant jusqu'au tétramère à la concentration en agent pontant la plus élevée (0,01 % glutaraldéhyde). La traînée au dessus de la bande du tétramère suggère que des oligomères d'ordre supérieur à 4 mais moins abondants existent aussi en en solution. Cependant, à ces concentrations en glutaraldéhyde, il est difficile d'estimer si ces espèces sont pontées spécifiquement ou non spécifiquement.

Au niveau des bandes des dimères, on peut en dénombrer deux dans le cas de HU α_2 et HU $\alpha\beta$, alors qu'une seule bande est observable dans le cas de HU β_2 . Comme la réaction a été réalisée à 37°C, la température est proche du T m_1 dans le cas de HU α_2 et HU $\alpha\beta$, on a donc un mélange des deux états N $_2$ et I $_2$, alors qu'elle est très au dessus du T m_1 (+10°) de HU β_2 qui est ainsi majoritairement sous forme I $_2$ à cette température. Le glutaraldéhyde figerait les deux conformations présentes pour HU α_2 et HU $\alpha\beta$. Une différence de surface accessible au SDS entre les deux formes conduirait à la variation de migration observée et expliquerait la présence de deux bandes. Au contraire, une unique bande est observée pour HU β_2 .

6.1.2. Effets de la température sur le pontage

Des pontages ont été réalisés pour chaque dimère en utilisant la concentration en glutaraldéhyde optimale pour l'observation du tétramère définie précédemment (0,01%) et en faisant varier la température d'incubation (4, 27 ou 37°C). Les produits de pontages ont été analysés par gel SDS-PAGE et par NAT-PAGE (**Figure 90 A et B, respectivement**).

L'analyse SDS-PAGE révèle que la quantité de tétramère ponté est plus faible à basse température. Néanmoins, on observe que la quantité de dimère ponté est également plus faible lorsque la température est basse. On a donc deux effets simultanés possibles à basse température :

- (i) une diminution de la proportion de HU I $_2$ par rapport à N $_2$ qui peut entraîner une diminution de la quantité de tétramère que l'on peut ponter.
- (ii) une diminution de la réactivité du glutaraldéhyde.

Dans ces conditions, il est difficile d'évaluer l'effet de la température sur l'état d'oligomérisation en solution des différents dimères de HU de *E. coli*.

L'analyse du NAT-PAGE met en évidence l'inhomogénéité du produit de pontage des dimères de HU. Les agents pontants chimiques bifonctionnels ciblant les amines primaire des protéines, tels que le glutaraldéhyde et le DMS, présentent de très nombreuses cibles sur les protéines (amines des liaisons peptidiques et des chaînes latérales). Les combinaisons de pontages possibles, parfois d'un seul groupement fonctionnel, sont extrêmement nombreuses. L'hétérogénéité de pontage conduit à une hétérogénéité de migration dans les conditions natives qui conduit aux traînées visibles sur les pistes où le pontage est efficace (**Figure 90 B**).

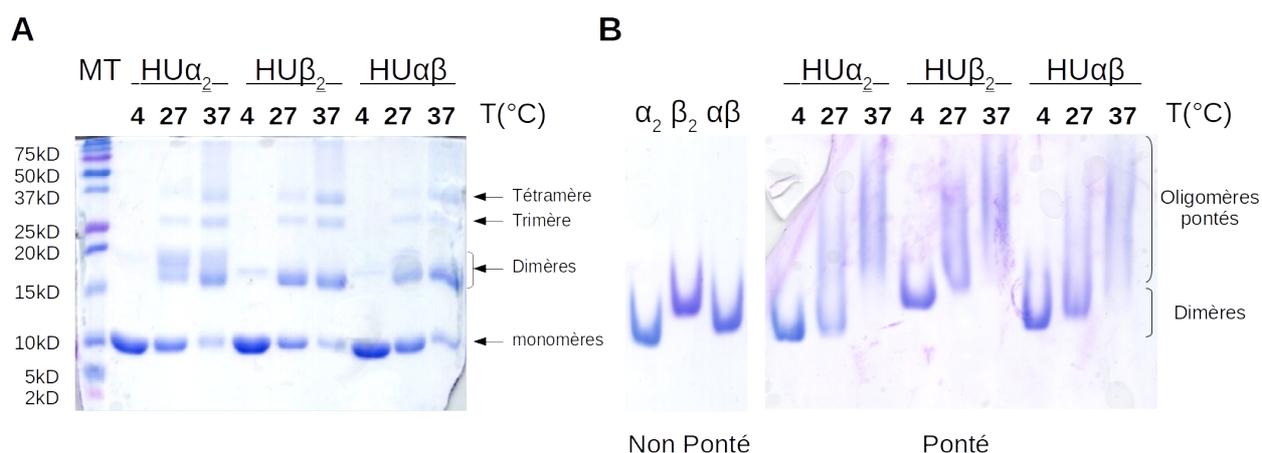


Figure 90: Effet de la température sur le pontage analysé par SDS-PAGE (A) et NAT-PAGE (B) La concentration en protéine est constante (1mg/mL) ainsi que la quantité de glutaraldéhyde (0,01 % m/v).

Dans le cadre de la mise en évidence et la caractérisation du tétramère, les pontages chimiques par le glutaraldéhyde présentent deux limitations majeures :

- (i) la haute concentration en agent pontant et les traînées visibles sur les SDS-PAGE ne permettent pas d'affirmer que le pontage du tétramère est spécifique dans ces conditions.
- (ii) l'hétérogénéité de pontage et donc des oligomères obtenus, n'est pas satisfaisante pour une analyse structurale qui requiert un échantillon pure et très homogène.

Une méthode de pontage alternative, spécifique et homogène, est nécessaire pour valider l'existence du tétramère et envisager sa caractérisation structurale.

6.2. Détection du tétramère par spectrométrie de masse en conditions natives

La spectrométrie de masse est une méthode reconnue pour la caractérisation des protéines et également de leurs complexes. Les liaisons non covalentes (hors effets de solvant) peuvent être maintenues en phase gazeuse dans des conditions douces d'ionisation, principalement en utilisant des sources de type électrospray. Ainsi, grâce aux développements méthodologiques et instrumentaux, la spectrométrie de masse peut permettre d'étudier des assemblages protéine-ligands ou protéine-protéine afin d'en extraire des données telles que l'affinité et la stœchiométrie par exemple. Le couplage de la spectrométrie de masse en conditions natives avec la séparation par mobilité ionique (sur des critères de taille et de forme) permet d'aller encore plus loin en séparant différentes conformations d'un mélange complexe (sur des critères de taille et de forme), différents oligomères et/ou de mettre en évidence des états intermédiaires de dénaturation. Des revues récentes traitant de l'utilisation de la spectrométrie de masse en conditions natives et de la mobilité ionique appliquée aux complexes protéiques ont été publiées par Hilton G. et collaborateurs et Utrecht C. et collaborateurs respectivement.^{235,236}

La spectrométrie de masse en conditions natives a été utilisée auparavant par Vis H. et collaborateurs pour l'étude de la sélectivité d'association des protéines HU issues des organismes *Bacillus subtilis* et *Geobacillus stearothermophilus*.¹⁹⁶ De manière analogue, nous avons employé cette méthode sur les dimères de *E. coli*. Des échantillons de chaque dimère de HU ont été préparés dans un tampon compatible avec la spectrométrie de masse (Acétate d'ammonium pH 5) et à faible concentration (<10 μ M) afin d'éviter une agrégation non spécifique. Le but est de mettre en évidence la possibilité de séparer les différentes espèces dimériques de HU d' *E. coli* et éventuellement de détecter des oligomères d'ordre supérieur au dimère. Les spectres de masse des protéines HU $\alpha\beta$, HU β_2 et HU α_2 sont présentés sur les **Figures 91, 92 et 93**, respectivement.

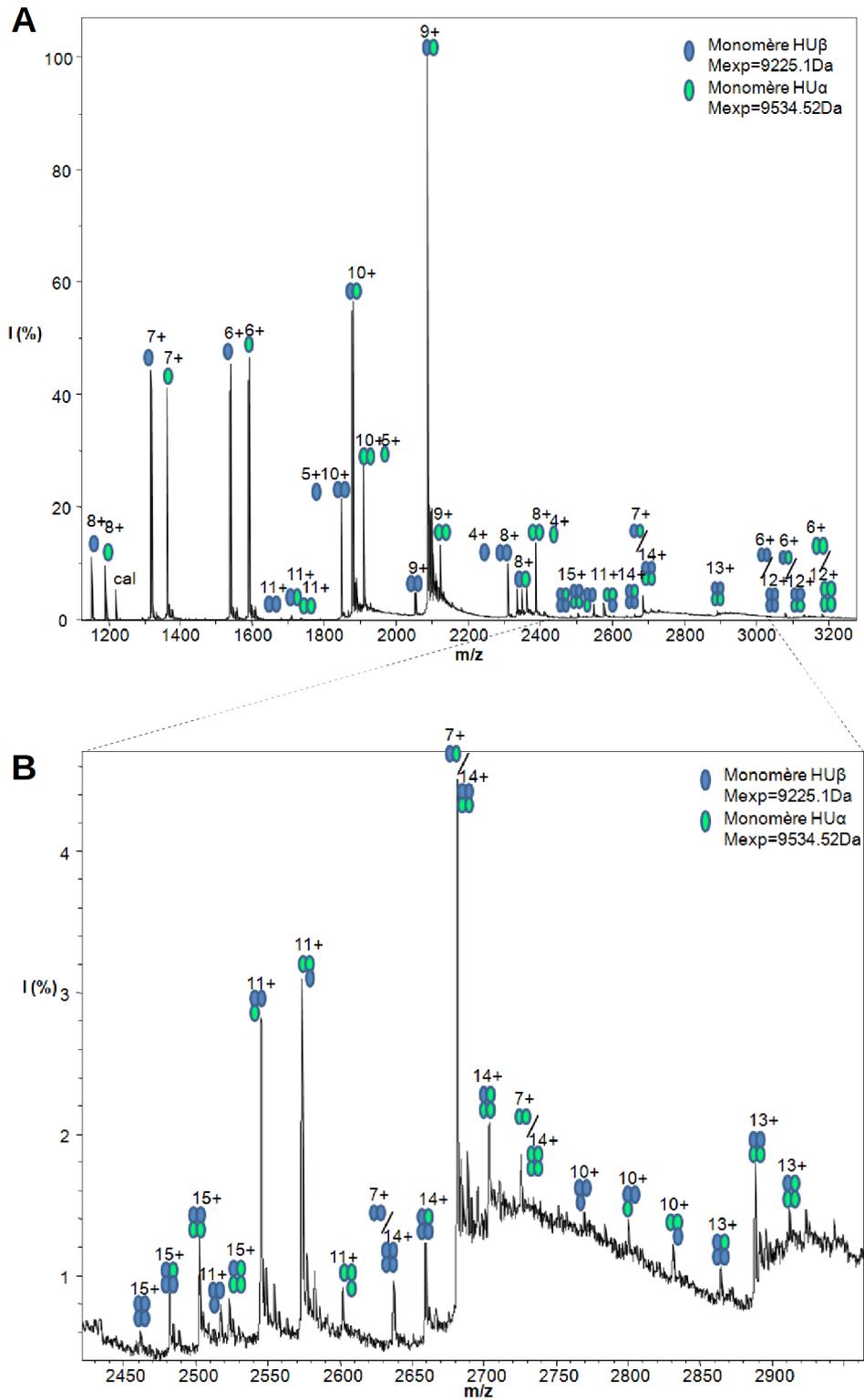


Figure 91: Spectre de masse en conditions natives de l'hétérodimère HU $\alpha\beta$ Spectre de masse dans la région de m/z des ions des dimères (A) et zoom sur la région des ions des tétramères (B).

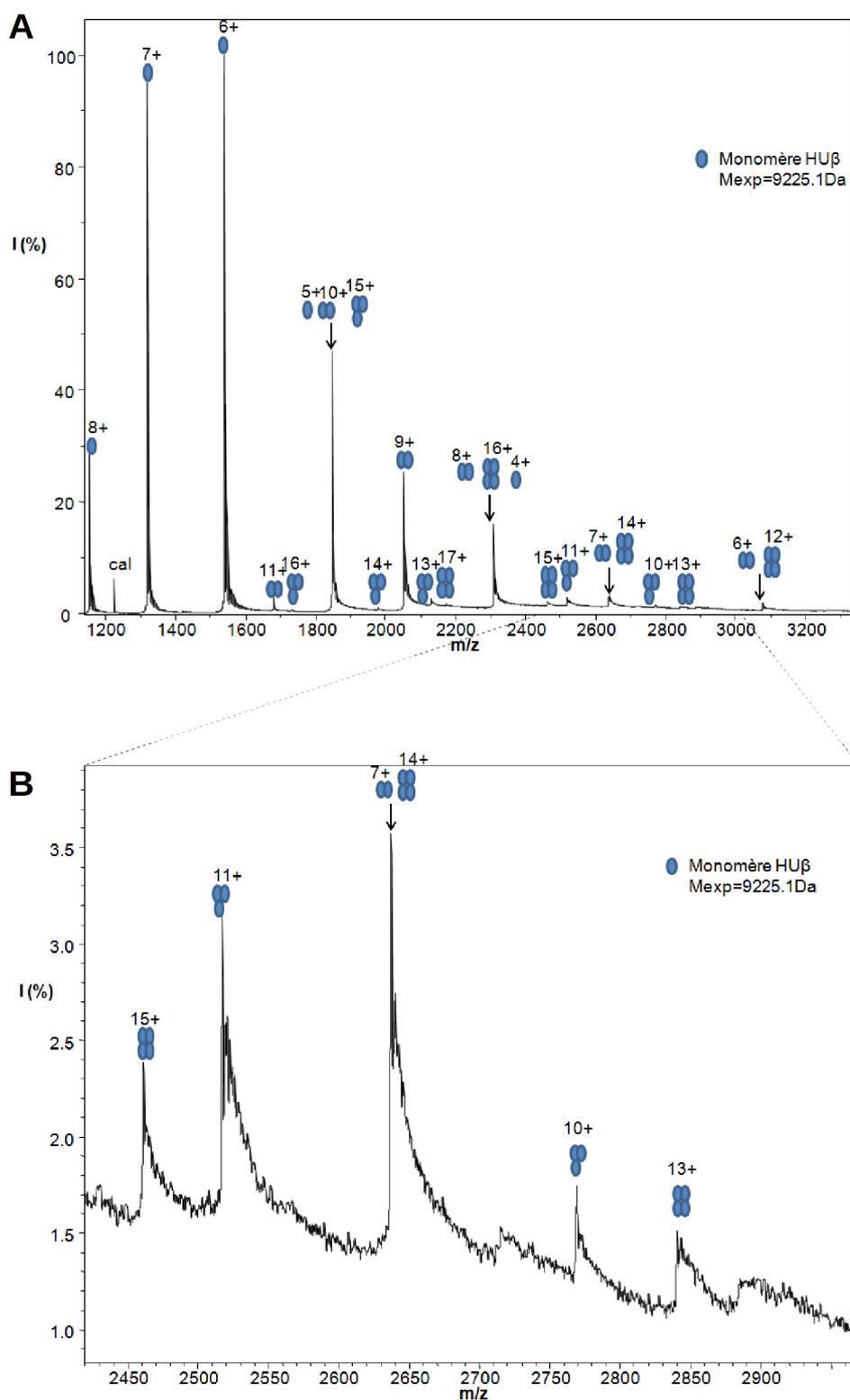


Figure 92: Spectre de masse en conditions natives de l'homodimère HUβ₂ Spectre de masse dans la région de m/z des ions du dimère (A) et zoom sur la région des ions du tétramère (B).

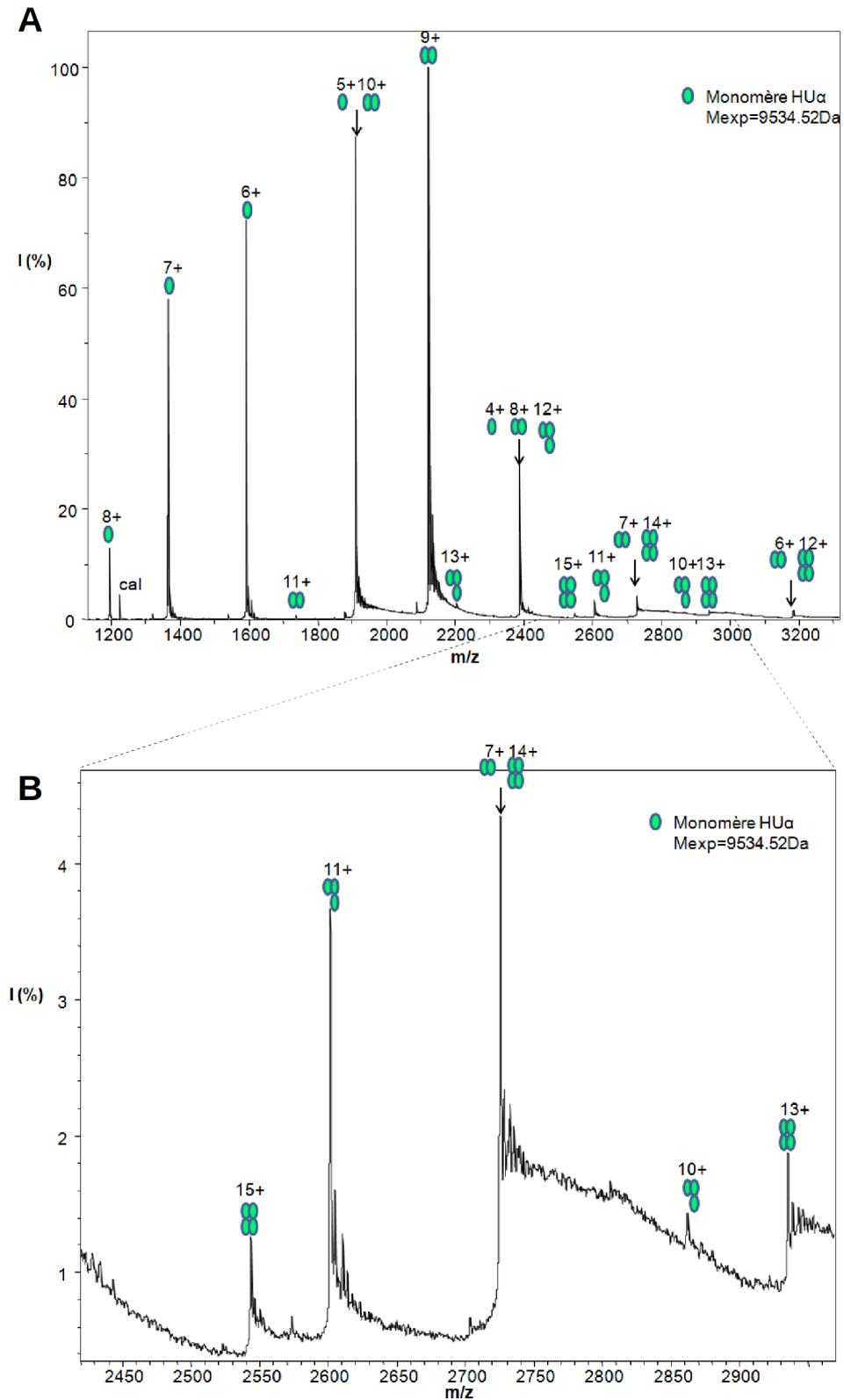


Figure 93: Spectre de masse en conditions natives de l'homodimère $HU\alpha_2$ Spectre de masse dans la région de m/z des ions du dimère (A) et zoom sur la région des ions du tétramère (B).

En spectrométrie de masse en conditions natives, en absence de séparation par mobilité ionique, les ions des dimères sont indistinguables de ceux des monomères pour les états de charge pairs. On a par exemple la même valeur de m/z du monomère chargé 4+ et du dimère chargé 8+ (**voir Tableau 17**). L'observation des dimères est faite sur les états de charges impairs pour lesquels il n'existe pas d'équivalent monomérique. Les états de charges spécifiques 7+, 9+ et 11+, sont observables pour chacun des dimères de HU et témoignent du maintien d'au moins une partie du dimère en phase gazeuse dans nos conditions expérimentales (**Figures 91 à 93**). L'état de charge 9+ est le plus intense et permet une bonne séparation des différents dimères de HU et peut donc être utilisé pour le suivi cinétique de l'échange de chaînes entre les dimères de HU (Chapitre 7).

Des oligomères d'ordre supérieur à deux sont également observables, notamment dans la gamme de m/z 2350-3000 (**Figure 91 B, 92 B et 93 B**). Des ions spécifiques des tétramères (11+ et 13+) et des trimères (10+, 11+ et 13+) sont visibles sur les spectres.

On peut conclure de ces expériences que les protéines HU de type sauvage d'*E. coli* peuvent former des tétramères en solution, et ce même en l'absence d'agent de pontage. La proportion de tétramères est très faible face aux dimères (de l'ordre de 1 % dans nos conditions) et met en évidence le caractère instable et transitoire de cette espèce par rapport aux dimères. On remarque que chaque dimère peut s'associer avec lui-même ou avec un autre type de dimère de HU donnant lieu à des homo-tétramères (**Figures 91 B, 92 B et 93 B**) ou à différents hétéro-tétramères (**Figure 91 B**). Cette observation est cohérente avec le fait que chacune des protéines HU de *E. coli* possède un état intermédiaire I_2 et est potentiellement capable de former un tétramère en suivant notre modèle d'échange. Ceci implique que chaque dimère doit pouvoir échanger ses chaînes protéiques avec les autres dimères de HU d'*E. coli* et même potentiellement avec lui-même (voir Chapitre 7). La possibilité d'un mécanisme d'échange de chaînes d'un homodimère avec lui-même était déjà suggérée par les expériences de pontages chimiques (**Figures 88 p 192 et 89 p 193**). Le pontage au DMS semblait indiquer que la propension de $HU\beta_2$ à homotétramériser corrélait avec l'abondance de la conformation I_2 (**Figure 88 p 192**). Aussi, la formation d'homotétramère de $HU\beta_2$ peut être illustrée par le mécanisme présenté dans la **figure 94** qui reprend le même formalisme que celui que nous avons proposé pour la formation de l'hétérodimère au début de ce chapitre (**Figure 87 p 191**).

La présence d'espèces trimériques et monomériques sur les spectres serait liée à un phénomène de dissociation en phase gazeuse des tétramères et des dimères, respectivement. En effet, les conditions

d'ionisation les plus douces possibles n'empêche pas qu'une partie des oligomères se dissocie en phase gazeuse.

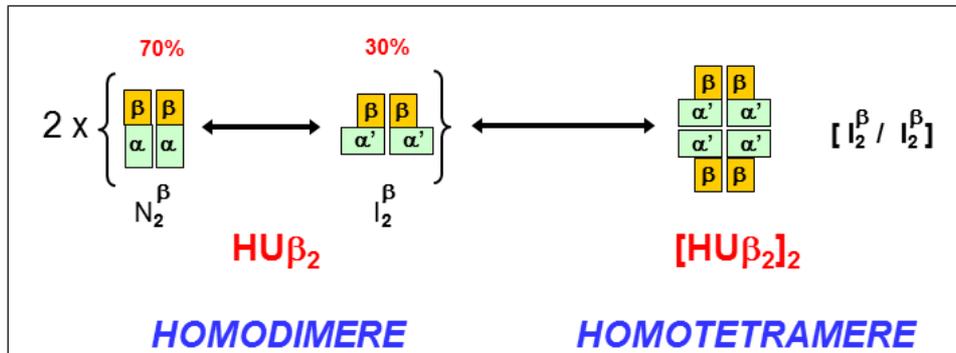


Figure 94: Mécanisme putatif de formation de l'homotétramère (illustré avec HUβ₂)

Tableau 17: Ions théoriques des oligomères de HU α_2 dans la gamme de m/z 1100-3200 Les ions annotés en gras sont effectivement visibles sur le spectre (**Figure 93**).

m/z	Monomère HU α (9535,96 Da)	Dimère HU α_2 (19070,92 Da)	Trimère HU α_3 (28605,87 Da)	Tétramère HU α_4 (38140,83 Da)
3179,3	[M+3H]³⁺	[M+6H]⁶⁺	[M+9H]⁹⁺	[M+12H]¹²⁺
2934,8				[M+13H]¹³⁺
2861,5			[M+10H]¹⁰⁺	
2725,3		[M+7H]⁷⁺		[M+14H]¹⁴⁺
2601,4			[M+11H]¹¹⁺	
2543,7				[M+15H]¹⁵⁺
2384,7	[M+4H]⁴⁺	[M+8H]⁸⁺	[M+12H]¹²⁺	[M+16H]¹⁶⁺
2244,5				[M+17H] ¹⁷⁺
2201,4			[M+13H]¹³⁺	
2119,9		[M+9H]⁹⁺		[M+18H] ¹⁸⁺
2044,2			[M+14H] ¹⁴⁺	
2008,3				[M+19H] ¹⁹⁺
1908,0	[M+5H]⁵⁺	[M+10H]¹⁰⁺	[M+15H] ¹⁵⁺	[M+20H] ²⁰⁺
1817,2				[M+21H] ²¹⁺
1788,8			[M+16H] ¹⁶⁺	
1734,6		[M+11H]¹¹⁺		[M+22H] ²²⁺
1683,6			[M+17H] ¹⁷⁺	
1244,7				[M+23H] ²³⁺
1590,2	[M+6H]⁶⁺	[M+12H] ¹²⁺	[M+18H] ¹⁸⁺	[M+24H] ²⁴⁺
1526,6				[M+25H] ²⁵⁺
1506,5			[M+19H] ¹⁹⁺	
1467,9		[M+13H] ¹³⁺		[M+26H] ²⁶⁺
1431,2			[M+20H] ²⁰⁺	
1413,6				[M+27H] ²⁷⁺
1363,1	[M+7H]⁷⁺	[M+14H] ¹⁴⁺	[M+21H] ²¹⁺	[M+28H] ²⁸⁺
1316,1				[M+29H] ²⁹⁺
1301,2			[M+22H] ²²⁺	
1272,3		[M+15H] ¹⁵⁺		[M+30H] ³⁰⁺
1244,7			[M+23H] ²³⁺	
1231,3				[M+31H] ³¹⁺
1192,8	[M+8H]⁸⁺	[M+16H] ¹⁶⁺	[M+24H] ²⁴⁺	[M+32H] ³²⁺
1156,8				[M+33H] ³³⁺
1145,2			[M+25H] ²⁵⁺	
1122,76		[M+17H] ¹⁷⁺		[M+34H] ³⁴⁺
1101,2			[M+26H] ²⁶⁺	

6.3. Mise en place d'une stratégie pour isoler le tétramère

L'étude de l'oligomérisation de HU par pontages chimiques et spectrométrie de masse en conditions natives nous a permis de confirmer que les dimères de HU d'*E. coli* pouvaient exister en solution sous la forme de tétramères. La caractérisation structurale de ces espèces n'est pas envisageable ici (abondance apparente de l'ordre de 1%). La stabilisation des tétramères, de façon homogène contrairement aux pontages chimiques, est nécessaire pour pouvoir mener l'étude structurale..

Pour réaliser un pontage homogène et spécifique, nous avons réalisé des mutations cystéines. Le but est d'oxyder ces cystéines nouvellement introduites (les chaînes d'ECHU n'en contiennent pas naturellement) en ponts disulfure afin d'obtenir des dimères covalents stables. Si les ponts disulfure sont placés de telle façon que les dimères covalents puissent toujours transiter vers leur conformation I₂, on pourrait obtenir un tétramère plus stable (un "cul-de-sac" thermodynamique), qui aurait perdu sa capacité à se dissocier en hétérodimères (**Figure 95**).

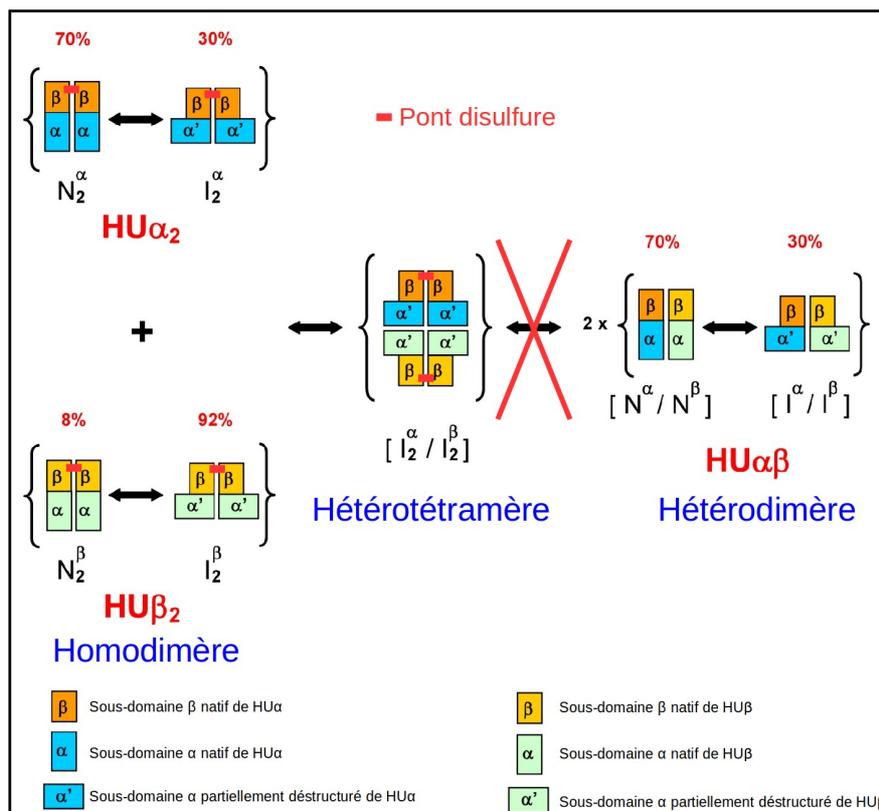


Figure 95: Modèle de stabilisation du tétramère par introduction d'un pont disulfure au niveau du sous-domaine β

Suivant ce même mécanisme, si on veut au contraire bloquer la formation de la conformation intermédiaire I_2 , alors on choisira de placer les ponts disulfures au niveau des sous-domaines α de HU.

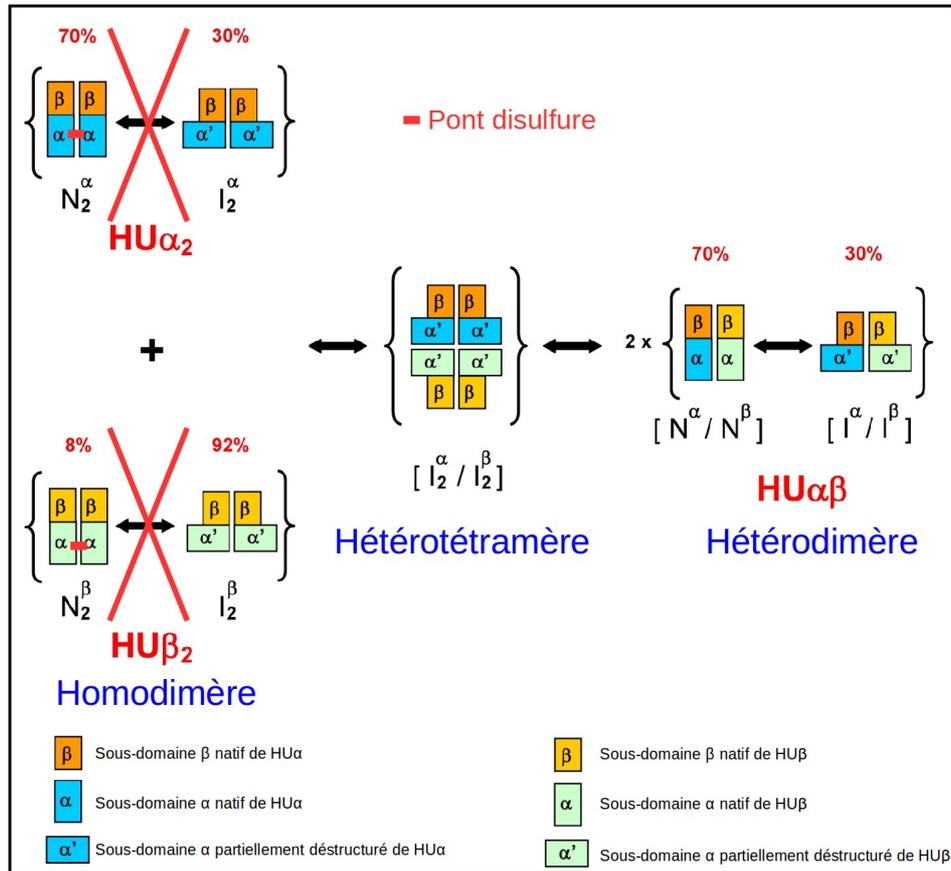


Figure 96: Modèle de stabilisation des dimères natifs par introduction d'un pont disulfure au niveau des sous-domaines α

6.3.1. Mutagenèse dirigée et test *in silico*

Les protéines HU de type sauvage d'*E. coli* ne possèdent pas de cystéine dans leur séquence primaire. Les seules possibilités d'appariement de ponts disulfure seront donc entre les cystéines introduites par mutagenèse dirigée. Comme nous ciblons des homodimères, nous cherchons dans les structures des ipso-contacts (points de contacts au sein d'un homo-oligomère entre deux acides aminés ayant la même position dans la séquence). Ainsi, une seule mutation est nécessaire pour former le pont disulfure et il n'y a pas d'ambiguïté d'appariement des ponts entre des cystéines multiples.

L'analyse des structures a permis de mettre en évidence deux de ces points de contact : la phénylalanine 79 au niveau des feuillets β (brin β_3) et l'alanine 24 au début des hélices α_2 (**Figure 97**). Nous avons utilisé le logiciel Sybyl™ pour réaliser les mutagenèses dirigées *in silico* à partir de la structure de HU α_2 (code PDB : 1MUL), former les ponts disulfure et minimiser l'énergie des structures obtenues. Le RMSD entre les structures de type sauvage de départ et les structures de type mutant, pontées et minimisées sont de 0,185 Å et 0,195 Å pour les mutants F79C et A24C respectivement, ce qui témoigne de perturbations très limitées des structures occasionnées par ces mutations *in silico* (pas de perturbations majeure des structures secondaires ou du repliement) (**Figure 97**).

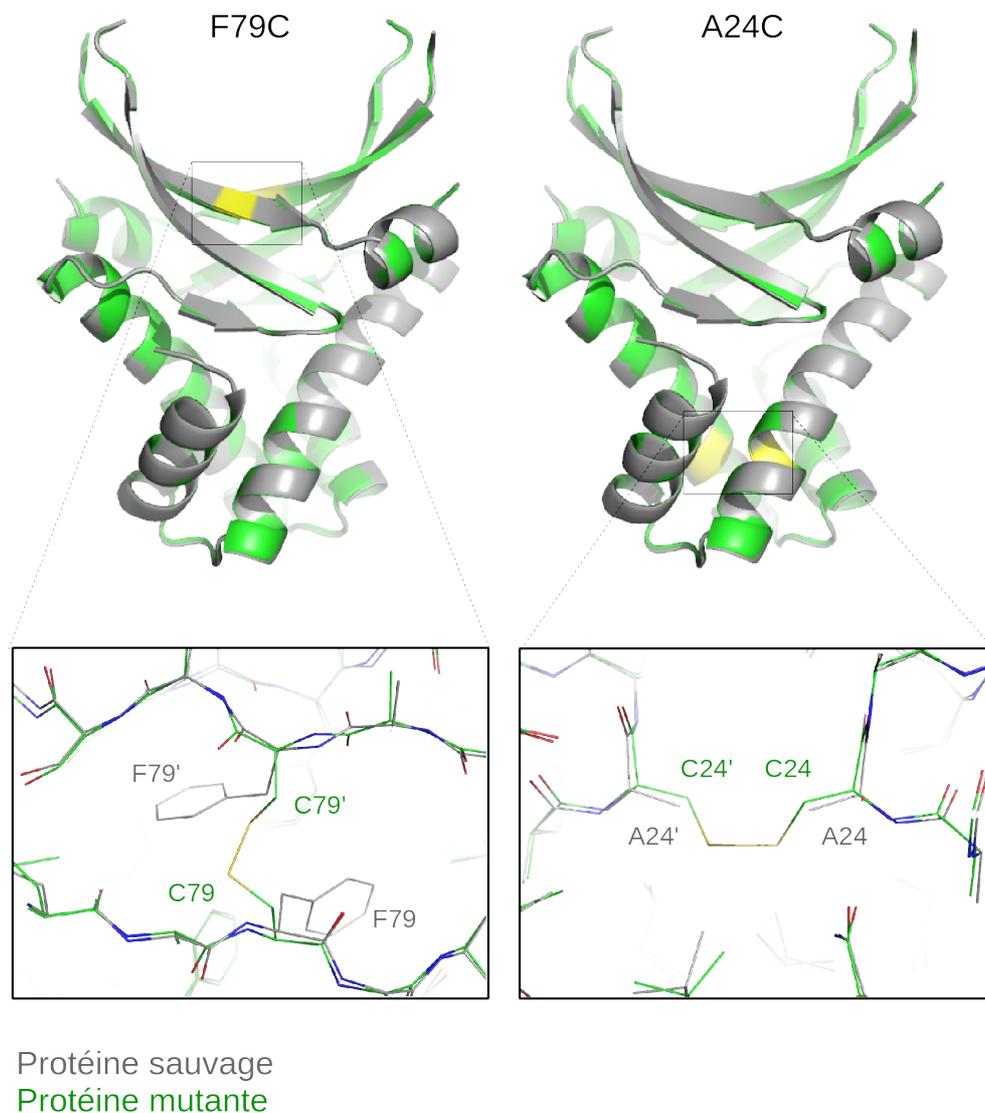


Figure 97: Mutations cystéines de HU *in silico*

6.3.2. Production et Purification du mutant F79C de HU α_2

La production du mutant F79C fusionné à la MBP est réalisée comme indiqué dans le chapitre 2 (matériels et méthodes p 62). La purification est réalisée selon le protocole retenu dans le chapitre 3 (fusion MBP). Pour maintenir l'état réduit du mutant F79C, la purification a été réalisée en supplémentant les tampons avec 10 mM final DTT (sauf dans le cas du clivage protéolytique à l'entérokinase car cette elle y est sensible au DTT). Les gels SDS-PAGE en conditions réductrices et non réductrices, résumant la purification est présenté **Figure 98**.

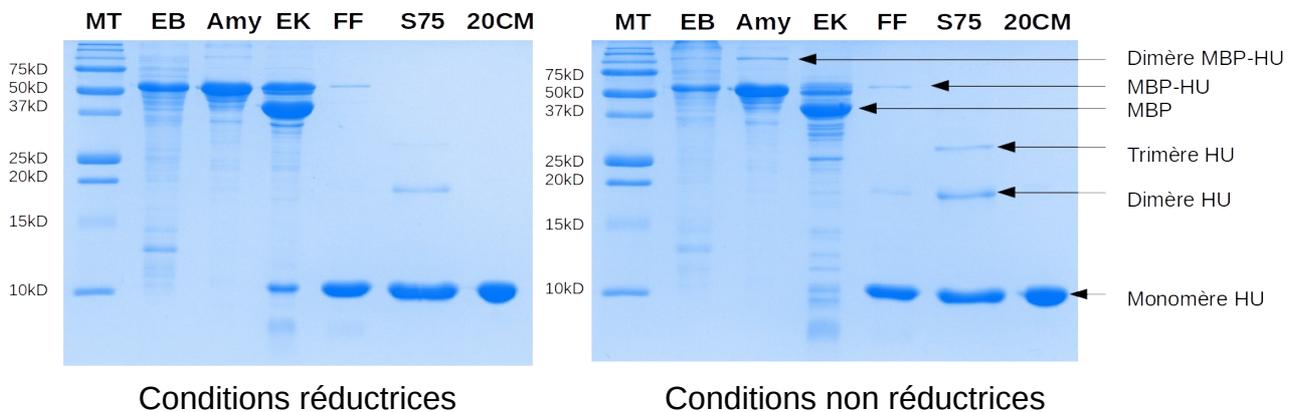


Figure 98: Gels SDS-PAGE bilans de la purification de HU α_2 F79C réduite MT = marqueur de taille, EB = extrait brut, Amy= fraction de la colonne d'affinité, EK = protéolysat, FF = fraction de la chromatographie échangeuse d'ion SP-FF, S75 = tamisage moléculaire sur support S75, 20CM = chromatographie échangeuse de cations faibles.

Malgré la présence de DTT en forte concentration, on peut observer différents oligomères de HU à l'issue de l'étape de gel filtration (S75). Même sur le SDS-PAGE en conditions réductrices, on observe une espèce qui a une masse moléculaire apparente compatible avec un dimère. Il semble qu'il y ait un pontage non spécifique qui ait lieu à cet étape, peut-être exacerbé par la haute concentration dans laquelle on se place pour pouvoir charger la colonne avec un volume d'injection suffisamment faible à cette étape (<1mL).

Cependant, on élimine ces espèces multimériques à la dernière étape de purification et c'est bien la protéine de type mutante, pure et réduite que l'on obtient. La prochaine étape consiste au pontage contrôlé des cystéines introduites pour former le dimère covalent.

6.3.3. Pontage des cystéines de la version mutante F79C de HU α_2

Pour réaliser le pont disulfure, la première approche a été de simplement enlever l'agent réducteur. Le tampon de purification de la protéine (contenant 10 mM DTT) a été échangé contre un tampon minimum (20 mM NaP pH 8 ; 200 mM NaCl). Le pH plutôt alcalin permet de s'approcher du pKa de la cystéine (8,14) et de favoriser l'espèce thiolate nécessaire pour la formation des ponts disulfure. On n'augmente pas le pH davantage car le pI de HU α_2 est de 9,57 et on risque de se confronter à des problèmes de solubilité. La protéine (concentrée à 1 mg/mL, 50 μ M) a ensuite été mise à buller avec l'oxygène de l'air en utilisant une pompe péristaltique circulant à vide.

Des échantillons sont prélevés au cours du temps (t= 0 ; 36h ; 15j , stockage à -20°C) puis sont analysés par SDS-PAGE (**Figure 99**).

A t=0, après élimination de l'agent réducteur, une partie de la protéine a été immédiatement oxydée. Près de 50 % de protéine est pontée à ce stade. Après 36h de barbotage, l'évolution est faible et on observe à peine plus de protéine pontée qu'à t= 0. Au bout de 15j de barbotage, une majorité de la protéine est oxydée mais il reste de la protéine monomérique. Le pontage n'est toujours pas total. On remarque également des bandes de dégradation entre la bande du dimère et celle du monomère. Des bandes de poids moléculaires apparents plus importants sont apparues au dessus de celle du dimère et pourraient correspondre à des trimères et à des tétramères de HU. Pourtant une seule cystéine est disponible, par chaîne, pour former un unique pont disulfure. La cinétique rapide d'oxydation (à t=0) puis lente (jusqu'à 15j) laisse penser que la solution protéique a rapidement épuisé les espèces oxydantes disponibles et que la dissolution de l'oxygène de l'air est insuffisante pour atteindre une oxydation totale.

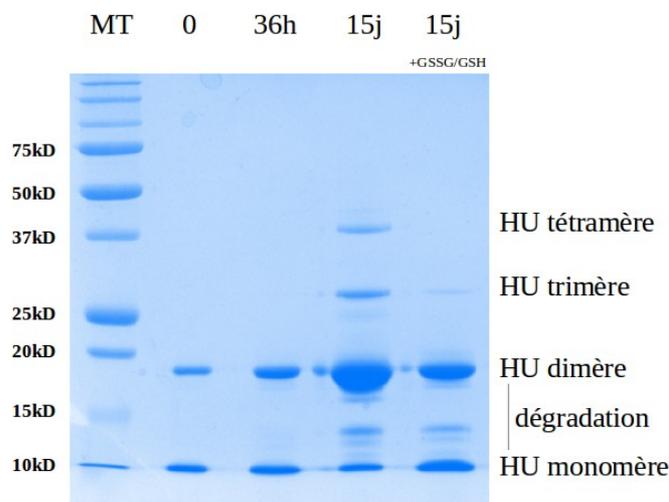


Figure 99: Tests préliminaires d'oxydation à l'oxygène de l'air. Analyse par SDS-PAGE de l'oxydation du mutant F79C au cours du temps (t=0, t=36h et t=15j) en faisant barboter à l'air la solution protéique. La dernière piste correspond à 15 jours de barbotage suivi de l'ajout du couple redox (glutathion oxydé/glutathion réduit) (GSSG/GSH=100/10).

Pour pousser plus loin la réaction d'oxydation, le couple redox glutathion oxydé (GSSG, 100 eq) et réduit (GSH, 10 eq) est ajouté à la solution protéique. Contre toute attente, l'analyse SDS-PAGE a révélé une diminution du dimère oxydé au profit du monomère réduit (**Figure 99**). L'excès du partenaire redox GSSG aurait dû favoriser l'oxydation du dimère de HU selon un mécanisme d'échange thiolate/disulfide.²³⁷ Cette diminution de dimères oxydés s'accompagne d'une disparition des espèces de plus hauts poids moléculaires. Il semble donc que l'apparente oligomérisation de HU soit effectivement une conséquence du pont disulfure du dimère, ce qui conforte notre stratégie de mutagenèse pour la stabilisation du tétramère.

L'oxydation passive, ou uniquement par l'oxygène de l'air, se révèle très lente pour ce système, au point que la dégradation de l'échantillon survient avant la formation complète du pont disulfure. Une oxydation active en utilisant des agents oxydants plus forts et/ou plus concentrés est nécessaire pour obtenir le dimère oxydé dans des périodes de temps plus courtes.

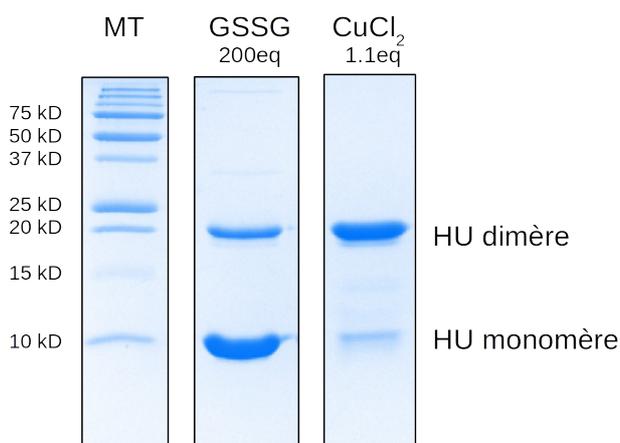


Figure 100: Oxydations en présence d'agents oxydants La protéine est incubée 20h avec soit 200 équivalents de glutathion oxydé (GSSG 200eq) soit 1,1 équivalent de chlorure de cuivre (CuCl_2 1.1eq). Les produits d'oxydation sont analysés par SDS-PAGE.

Pour accélérer l'oxydation des dimères de HU, des tests d'oxydation ont été réalisés soit avec un très large excès de GSSG seul (200 équivalents, soit 1 mM final) soit avec du chlorure de cuivre ajouté (1,1 équivalent, soit 55 μM final). L'analyse SDS-PAGE des produits d'oxydation est présentée **Figure 100**. Même en très large excès de GSSG et en absence de GSH, moins de 50 % de la

protéine est oxydée, ce qui est encore moins efficace que par barbotage avec de l'air. Par contre, l'effet du chlorure de cuivre a permis d'obtenir un excellent rendement, supérieur à 80 % au bout de 20h. On discerne cependant de faibles bandes de dégradation entre la bande du dimère et celle du monomère. Le Cuivre II est un oxydant bien plus fort que le glutathion avec un potentiel redox de 0,337 V contre -0,244 V pour le glutathion.^{238,239} Le potentiel redox des cystéines dans les protéines est fortement dépendant de la séquence locale et de l'environnement. Certaines protéines peuvent ainsi être plus ou moins efficacement oxydées par le couple GSSG/GSH. Dans notre cas, l'apparition de GSH suite à la formation de dimère covalent semble s'opposer à une formation quantitative du pont disulfure.

6.3.4. Suivi de la formation du pont disulfure de EcHU α_2 F79C par spectrométrie de masse

Un inconvénient majeur de l'analyse du pontage des cystéines par SDS-PAGE est l'étape de dénaturation thermique des échantillons. La haute température et les conditions dénaturantes sont favorables à l'oxydation et peuvent conduire à la surestimation de la vitesse et du taux de formation du pont disulfure. Pour pouvoir suivre l'apparition des dimères covalents et caractériser par ailleurs les éventuels dommages oxydatifs, nous avons employé la spectrométrie de masse en conditions dénaturantes. Des cinétiques d'oxydation ont été réalisées dans différentes conditions de température (37°C ou 0°C) et de pH (pH 5 ou pH 8). Dans notre cas, l'apparition de GSH suite à la formation de dimères covalents semble s'opposer à une formation quantitative du pont disulfure. Huit échantillons sont prélevés au cours du temps (entre 0 et 80h) et l'oxydation est bloquée par co-cristallisation et assèchement (dépôt en couche ultra mince, et analyse en MALDI-TOF)²⁴⁰. Les spectres de masse correspondant à ces cinétiques sont présentés **Figure 101**. Sur les états de charge paires, la différence de m/z entre la forme oxydée et réduite ne permet pas de les séparer à cette résolution. Ce problème est similaire à la séparation des oligomères de HU en conditions natives. Seuls les pics correspondant aux états de charges impairs du dimère oxydé sont non ambigus. Pour évaluer l'augmentation de l'espèce oxydée, le rapport entre la somme de ces pics (1+ 3+ 5+) divisée par l'ensemble des pics observables (monomère réduit + dimère oxydé) (R 1+ 3+ 5+) a été calculé. L'évolution de ce rapport (R 1+ 3+ 5+) au cours des cinétiques est présentée **Figure 102**.

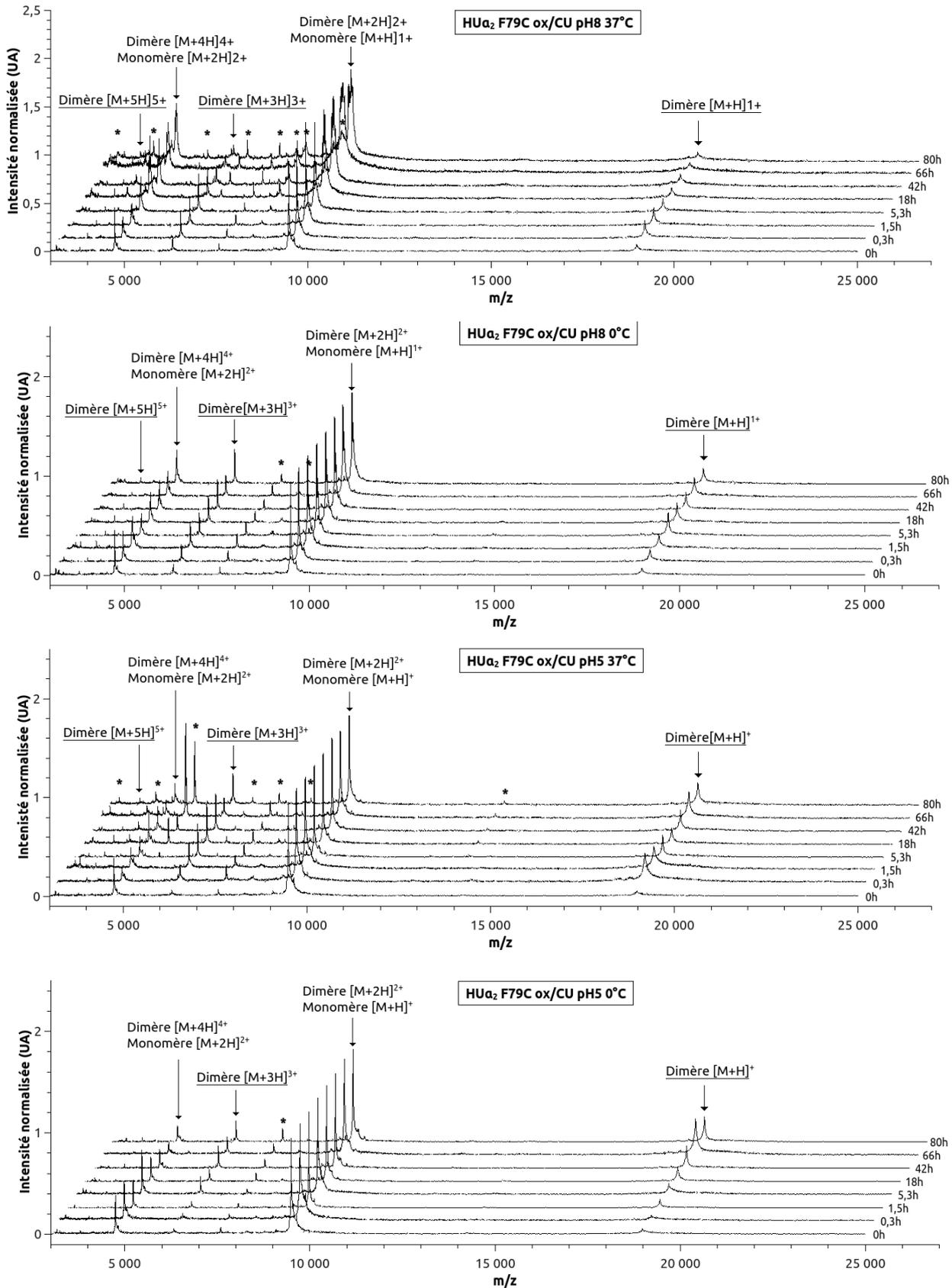


Figure 101: Cinétiques d'oxydation du dimère HU α_2 F79C par MS à différents pH et températures Les symboles '*' mettent en évidence des pics de dégradation.

A 37°C, les maxima d'oxydation sont atteints très rapidement (dès le second point à 20 min d'oxydation). Passé ce point, le rapport R diminue très rapidement (**Figure 102**) et on observe l'apparition de pics de dégradation sur les spectres de masse (*' **Figure 101**). A 0°C, les cinétiques paraissent globalement plus lente (en particulier à pH 5) mais on observe beaucoup moins de pics de dégradation. À pH acide (pH 5), les cinétiques d'oxydation sont plus lentes mais atteignent des maxima d'espèces pontées plus importantes qu'à pH alcalin (pH 8) (**Figure 102 à droite**).

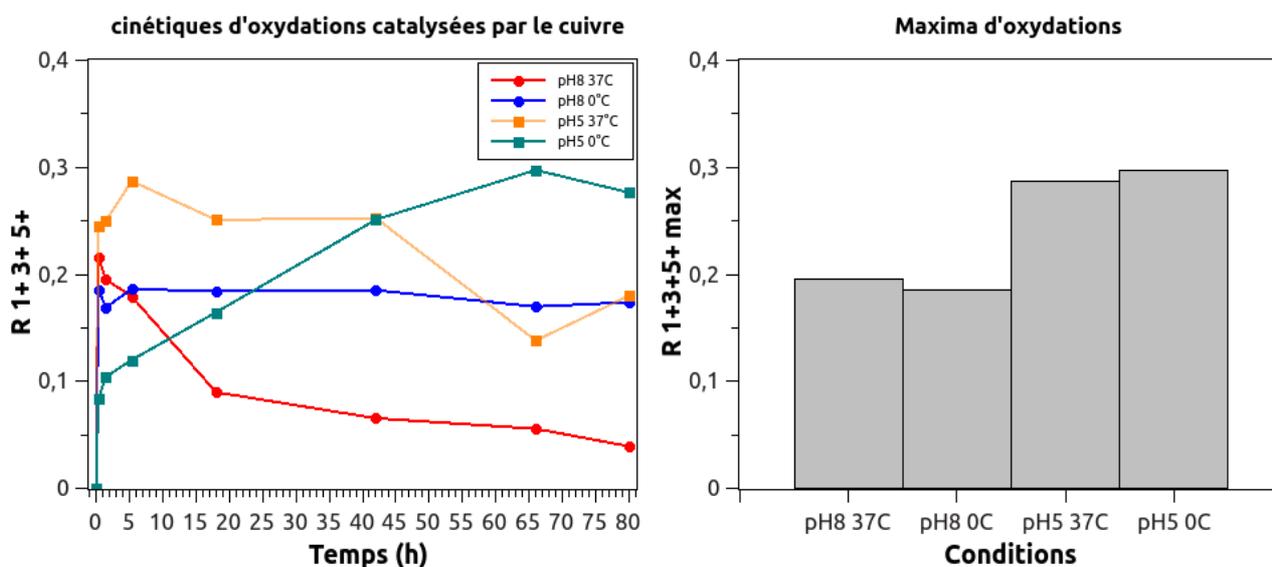


Figure 102: Evaluation de l'efficacité de pontage du mutant F79C Le rapport R 1+ 3+ 5+ représente le rapport des pics spécifiques du dimère oxydé (états de charges 1+ 3+ et 5+) sur la somme des intensités des pics observables.

En conclusion, la meilleure condition d'oxydation en terme de vitesse et de rendement a été la suivante : pH 5, 6 h, 37°C (pour laquelle on a encore peu de dommages oxydatifs). Si l'on souhaite conserver la température basse (dans la glace à 0°C), on réalisera l'oxydation sur un weekend (~72h) pour obtenir un maximum de dimères pontés sans générer trop de dommages oxydatifs.

6.3.5. Analyse des dommages oxydatifs du mutant EcHU α_2 F79C

Deux grands types de dommages ont été générés sur le mutants EcHU α_2 F79C au cours des expériences d'oxydation. Des produits de clivages de la chaîne protéique qui apparaissent dans une large gamme de m/z sur les spectres ('*' sur la Figure 101) et des modifications chimiques des chaînes latérales qui se manifestent de part et d'autre des pics des protéines non modifiées (Figure 103).

Différentes études ont permis de mettre en évidence de nombreuses modifications de chaînes latérales induites par l'oxydation catalysée par des métaux.²⁴¹⁻²⁴⁶ Les modifications chimiques observées lors de l'oxydation de F79C (Figure 103) peuvent s'expliquer par différentes combinaisons de dommages. Le +32 peut être dû à une modification de proline et les +16 à différents types d'oxydations supplémentaires.

Si l'identification des modifications chimiques reste difficile à partir de ces seules expériences, certains clivages protéolytiques peuvent être attribués et permettre d'identifier les points de coupures liés à l'oxydation. L'identification de ces sites de clivage a été réalisée à l'aide du logiciel libre Mmass^{198,199} et est présentée Figure 104. La protéine F79C est clivée du côté C-terminale comme cela est schématisé sur la Figure 104 B. Les sites de clivage identifiés sont majoritairement localisés au niveau des bras en amont du site de mutation. Il est probable que la dynamique importante et l'accessibilité des bras de HU soient responsables de la plus haute sensibilité aux clivages oxydatifs.

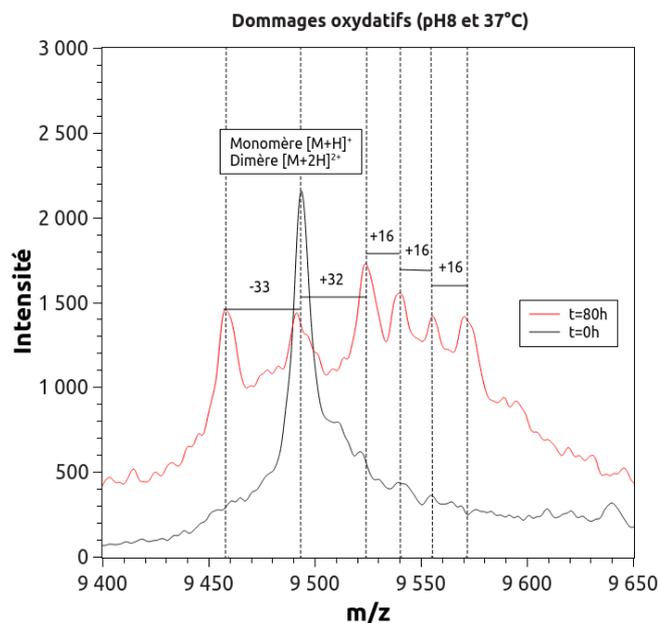


Figure 103: Modifications de chaînes latérales par oxydation Régions 9400-9650 (m/z) comparée entre les spectres enregistré à t=0 et t=80h (37°C et pH 8).

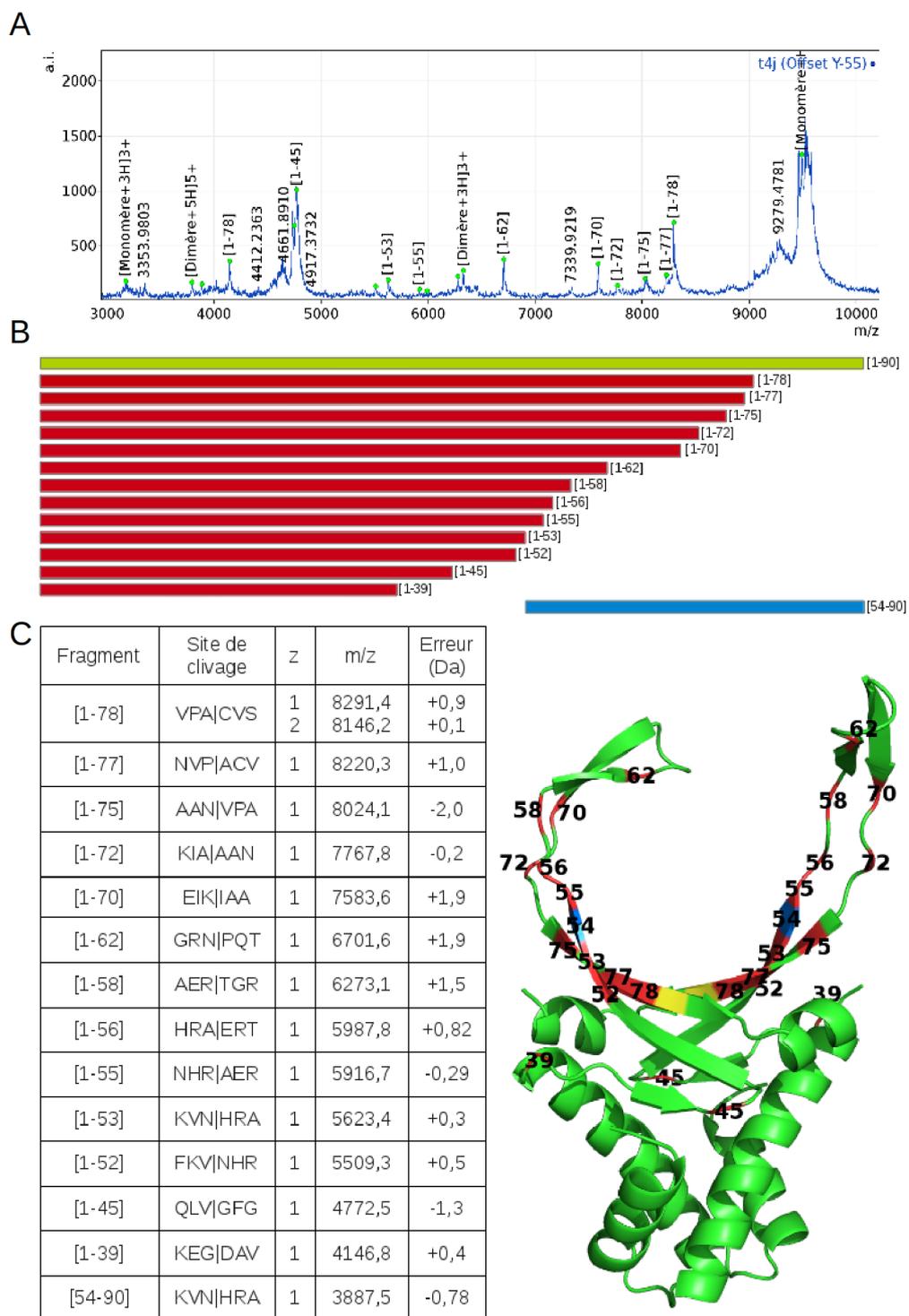


Figure 104: Identification des clivages protéolytiques observés au terme de l'oxydation à pH 8 et 37°C A- Spectre de masse. B- Représentation schématique des fragments identifiés. C- Identification et positionnement sur la structure des sites de clivage protéolytiques. Sont représentés en rouge les clivages de type C-terminaux et en bleu le clivage de type N-terminal. Le résidu muté, F79C, est repéré en jaune.

6.3.6. Purification et caractérisation du mutant F79C

Sur la base des conditions testées et suivies par spectrométrie de masse, une oxydation théoriquement quasi-optimale du mutant EchU α_2 F79C a été réalisée (pH 5, 37°C, 20 min soit un R attendu d'environ 0,25). Le produit d'oxydation a été purifié par chromatographie échangeuse d'ion (20CM). Le chromatogramme et l'analyse SDS-PAGE associée sont présentés **Figure 105**.

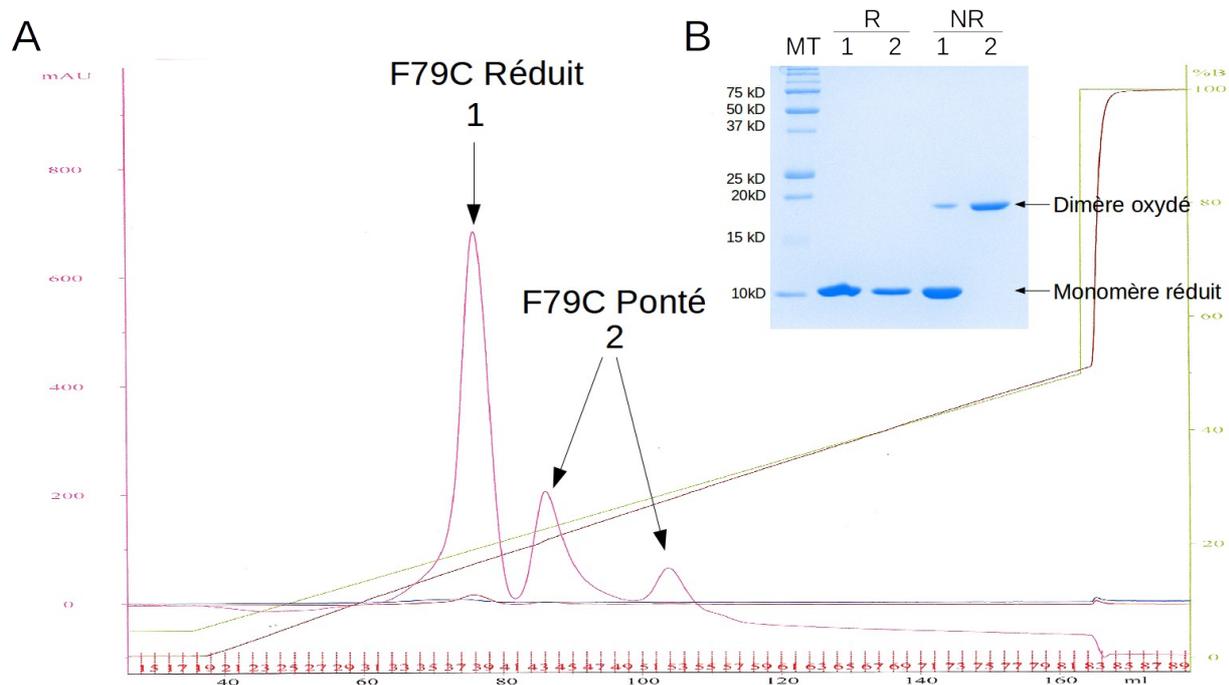


Figure 105: Chromatogramme (A) et analyse SDS-PAGE (B) de la purification sur 20CM de F79C après oxydation A- En magenta, en rouge et en bleu l'absorbance à 230, 260 et 280 nm respectivement. En marron la conductimétrie et en vert le gradient linéaire jusqu'à 0,5 M NaCl (50%). B- Analyse SDS -PAGE des pics 1 et 2 en conditions réductrices (R) ou non réductrices (NR).

Trois espèces sont séparées au cours de la purification sur le support 20CM. Le premier, majoritaire, correspond à la fraction de F79C réduite (élution à environ 190 mM NaCl) et contient très peu de dimère oxydé (**Figure 105 B**). Il semble que les conditions utilisées ici ne soient pas aussi efficaces que lors des tests d'oxydation suivis par spectrométrie de masse. Le pic 1 est collecté à part et réservé pour une oxydation ultérieure. Les pics suivants numérotés 2 (élus à 230 et 290mM NaCl) sont en équilibre. Si l'on réinjecte l'un ou l'autre de ces pics sur la colonnes, les deux pics sont de

nouveaux visibles à l'élution. Ces pics correspondent au F79C oxydé (**Figure 105 B**). Il semble que, contrairement au mutant réduit, le mutant ponté présente une hétérogénéité conformationnelle qui possèdent des propriétés chromatographiques différentes. On peut faire l'hypothèse que l'un des deux pics correspond au dimère ponté et que l'autre correspond à l'homo-tétramère. Nous n'avons pas été capable de séparer ces espèces par chromatographie d'exclusion (S75), mais la différence de poids moléculaire entre le dimère et le tétramère n'étant que de 20kDa, nous sommes à la limite de séparation de ce support.

Les analyses de spectrométrie de masses en conditions natives mettent en évidence des tétramères pour ce mutant sous sa forme oxydée (**Figure 106**). La proportion de tétramères n'est que modérément augmentée par rapport à la protéine de type sauvage et n'apparaît pas majoritaire dans les conditions utilisées (environ 2 à 3 % au lieu de 1% pour la protéine sauvage). Toutefois, les concentrations en protéine employées en spectrométrie de masse sont très faibles (<10 μ M). Comme l'association des dimères en tétramère est une étape bimoléculaire, elle doit fortement dépendre de la concentration en protéine. Une faible concentration est donc défavorable à la formation du tétramère.

L'autre hypothèse serait que la colonne 20CM sépare deux formes dimériques alternatives de ce mutant (N_2 et I_2). La séparation des deux conformations dimériques de HU par chromatographie n'a jamais été mis en évidence pour les protéines de type sauvage. Mais, comme la purification est menée à 4°C, on ne doit observer que des dimères natifs. Cependant, si la mutation a entraîné une diminution significative du Tm_1 , on peut imaginer que les deux formes soient abondantes à 4°C.

Nous avons tenté de cristalliser ce mutant sous sa forme oxydée et réduite en utilisant toutes les conditions du kit JCSG+ (Tableau 21 et 22 en annexe p 307 et 308), mais aucune tentative n'a abouti. Il est probable que l'hétérogénéité visible en chromatographie ne soit pas favorable au processus de cristallogenèse.

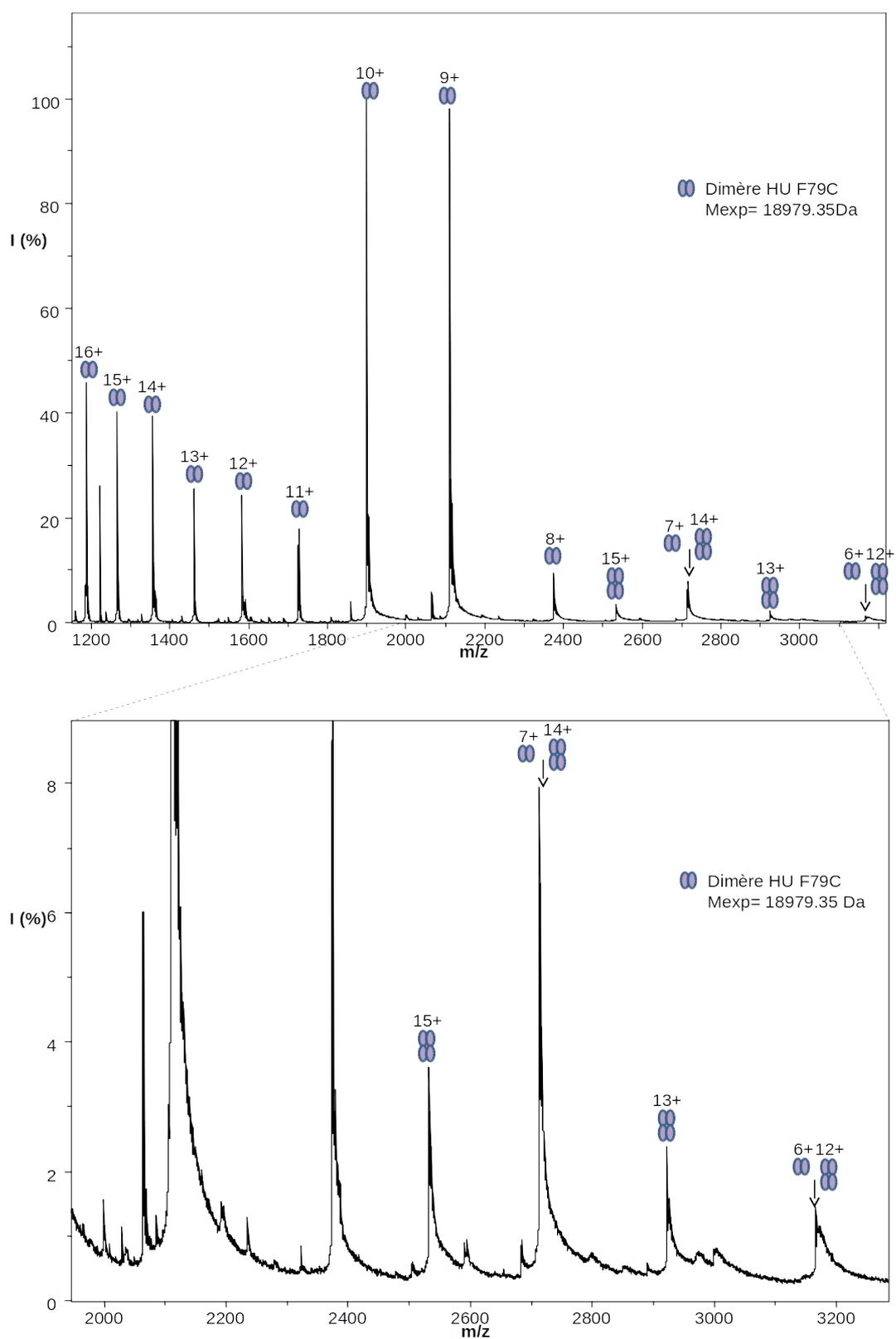


Figure 106: Spectre de masse en conditions natives du mutant ECHU α_2 F79C oxydé

6.4. Modèles putatifs de la structure du tétramère

Bien que nous ayons réussi à mettre en évidence la présence de conformations tétramériques des HU d' *E. coli*, nous n'avons pas pu obtenir leurs structures à ce jour. Nous ne pouvons que spéculer sur la structure des tétramères de HU et leur implication dans l'échange de chaînes. En effet, on peut imaginer la formation de deux types de tétramères : une association de dimères sous leurs formes N_2 ou une association de dimères en conformations intermédiaires I_2 .

Un mode d'interaction entre deux dimères N_2 peut être envisagé sur la base des empilements cristallins des structures de HU (**Figure 107**). Différents types d'arrangements entre ses dimères émergent. On observe des associations où les dimères se trouvent l'un au dessus de l'autre en établissant des contacts selle-HTH. C'est le cas pour les HU de *M. tuberculosis*, *B. stearothermophilus* et *B. anthracis*. D'autres présentent des protéines côte à côte en formant des interactions feuillet-HTH ou feuillet-feuillet comme les HU d'*E. coli*. Enfin, les protéines HU de *T. maritima* et de *L. lactis* s'associent « en travers » par des interactions entre les hélices notamment. Parmi toutes ces structures cristallographiques, celle de *L. lactis* présente la surface d'interaction inter-dimère la plus étendue. Les interactions entre les hélices des HTH et les hélices C-terminales sont à mettre en parallèle avec la caractérisation structurale de la conformation intermédiaire des homodimères. On pourrait envisager un mécanisme d'échange alternatif où les dimères s'associent sous forme N_2 puis subissent un changement de conformation de type I_2 au sein de ce tétramère pour échanger leur chaînes et former des hétérodimères.

Nous privilégions l'hypothèse où la structure tétramérique repose sur l'association des dimères dans leur conformation I_2 . En suivant l'hypothèse formulée à la fin du chapitre 5 sur la structuration de I_2 , on peut imaginer l'association de deux de ces dimères « ouverts » pour former un tétramère (**Figure 108**). Dans ce modèle, la position des hélices ouvertes permet de reformer deux domaines HTH (de type N_2 hétérodimérique). Ce serait donc les mêmes interactions qui maintenaient les dimères sous forme N_2 , qui permettent de former les tétramères à partir de I_2 . Ce mécanisme de formation d'oligomères correspond à de l'échange de domaines (« domain swapping ») et a déjà été mis en évidence pour de nombreuses protéines. Des revues sur ce mécanisme d'oligomérisation ont été publiées par Liu Y. et collaborateurs²⁴⁷ et Bennet M. J. et collaborateurs²⁴⁸. L'échange de domaine a notamment été proposé comme mécanisme d'évolution du repliement des histones, en particulier au niveau de leurs interactions inter-dimères par les domaines HTH.

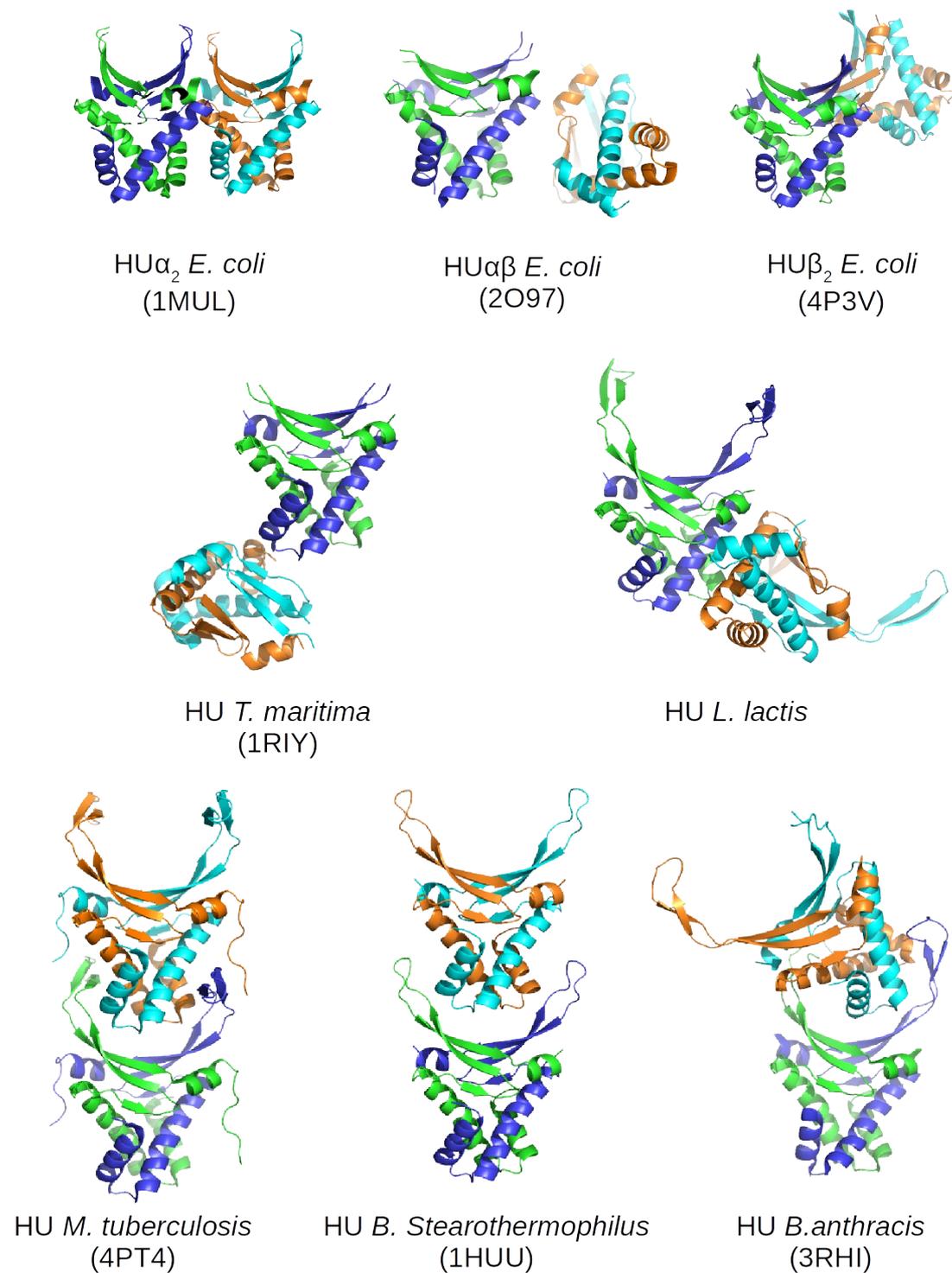


Figure 107: Contacts inter-dimères représentatifs au sein des mailles cristallines des protéines HU cristallisées

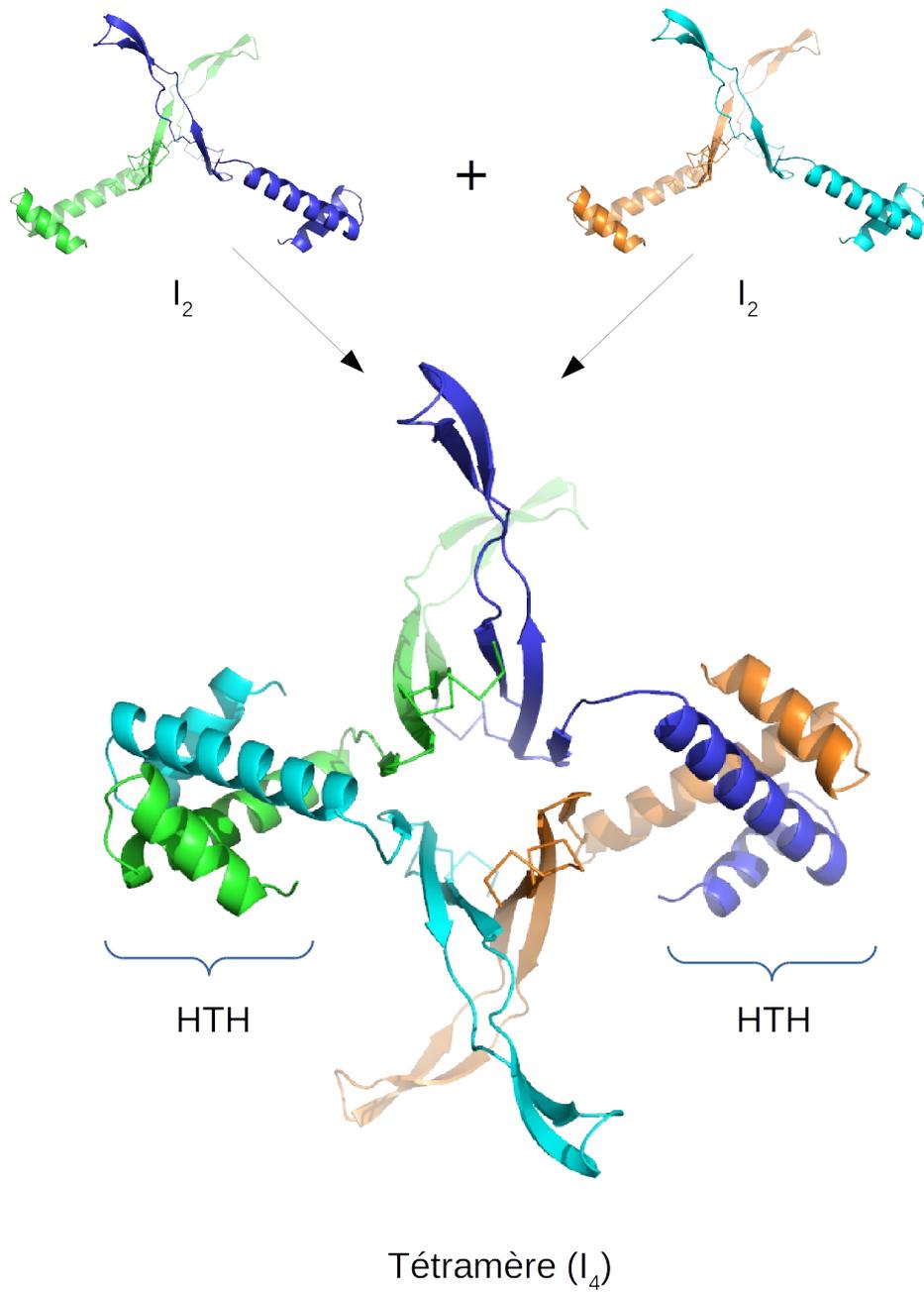


Figure 108: Modèle de formation d'un tétramère à partir de HU en conformations I_2

6.5. À retenir

Objectif du chapitre : Mettre en évidence le tétramère
<ul style="list-style-type: none">- Pontage chimique des dimères $HU\alpha_2$ $HU\beta_2$ et $HU\alpha\beta$→ non satisfaisant car non spécifique
<ul style="list-style-type: none">- Détection du tétramère par Spectrométrie de masse (sans agent pontant)→ Détection des homo- et hétéro-tétramères (instables et transitoires)→ Chaque dimère est capable d'échanger ses chaînes avec un autre dimère et potentiellement avec lui-même
<ul style="list-style-type: none">- Blocage du tétramère par introduction d'un pont disulfure<ul style="list-style-type: none">. Choix des points de mutations (ipsocontacts) : F79C et A24C. Production/purification du mutant $HU\alpha_2F79C$ sous forme réduite. Optimisation des conditions d'oxydation pour former le pont et suivi par SM. purification du mutant $HU\alpha_2F79C$→ Hétérogénéité de la forme oxydée : présence d'un équilibre dimère \leftrightarrow tétramère ou dimère $N_2 \leftrightarrow$ dimère I_2→ Détection du tétramère $HU\alpha_2F79C$ par SM (2-3%)
<p>Conclusion : dans ce chapitre, j'ai mis en évidence la présence de la forme tétramérique, forme transitoire instable, et je propose un modèle structural en accord avec l'ensemble des données actuelles.</p>

Chapitre 7 : Étude préliminaire de la Formation de l'hétérodimère

Nous avons vu dans le chapitre précédent que les dimères de HU peuvent être discriminés par spectrométrie de masse en conditions natives. Dans ce chapitre, nous exploitons cette technique pour suivre l'échange de chaînes au cours du temps et à différentes températures. Les informations extraites de ces suivis cinétiques sont utilisées afin d'éprouver notre modèle d'échange.

7.1.1. Mode opératoire

Pour pouvoir suivre la formation de l'hétérodimère au cours du temps, dans différentes conditions de température et de force ionique, des échantillons d'homodimères préincubés à la température souhaitée sont mélangés volume à volume à $t=0$. Des échantillons sont prélevés du mélange réactionnel au cours du temps et la réaction d'échange est stoppée par congélation rapide dans l'azote liquide. Les échantillons congelés sont stockés à -20°C en attendant l'analyse par spectrométrie de masse native. Au moment de l'analyse, les échantillons sont décongelés lentement dans la glace (à 0°C). A cette température toutes les HU sont dans leurs conformations N_2 et stables. L'échange de chaînes est bloqué. Au moment de l'analyse par spectrométrie de masse, les échantillons sont prélevés à l'aide d'une seringue Hamilton froide et injectés dans un capillaire jusqu'à la source d'ionisation (Electrospray). Pour empêcher la poursuite de l'échange sur ce trajet, où la grande surface de contact avec le capillaire pourrait rapidement réchauffer l'échantillon, la seringue est maintenue froide par un sac de glace et le capillaire est refroidi par un système de circulation d'eau glacée l'entourant tout le long du trajet de l'échantillon jusqu'à la source. On évite ainsi au maximum une poursuite de l'échange non souhaitée au cours de l'analyse.

7.1.2. Cinétique de formation de l'hétérodimère suivie par spectrométrie de masse en conditions natives

Nous suivons l'évolution des ions $9+$ de chaque dimère au cours du temps (**Figure 109**). Les intensités brutes ne sont pas comparables d'une expérience à l'autre et ne peuvent pas être utilisées directement pour déterminer l'évolution de la concentration des différentes espèces au cours du temps. Pour s'affranchir de cette limitation, on calcule, à chaque temps, le rapport R d'intensité du pic de l'hétérodimère sur la somme des trois pics $9+$ des dimères. En faisant l'hypothèse que la

dissociation en phase gazeuse est semblable pour chaque dimère, ce rapport est directement proportionnel à la concentration en hétérodimère dans l'échantillon. Ainsi, on peut suivre l'évolution de ce rapport au cours du temps et observer l'effet de la température sur les cinétiques. Les résultats sont présentés sous forme graphique **Figure 110**. L'apparition de l'hétérodimère est exponentielle, ce qui nous indique que la réaction globale est d'ordre 1 ou de pseudo-ordre 1. Trois équations ont été envisagées pour l'ajustement des courbes :

$$y = A \cdot (1 - e^{(-k_{obs} \cdot x)}) \quad (1)$$

$$y = A \cdot (1 - e^{(-k_{obs} \cdot x)}) + C_0 \quad (2)$$

$$y = A \cdot (1 - e^{(-k_{obs} \cdot (x+B))}) \quad (3)$$

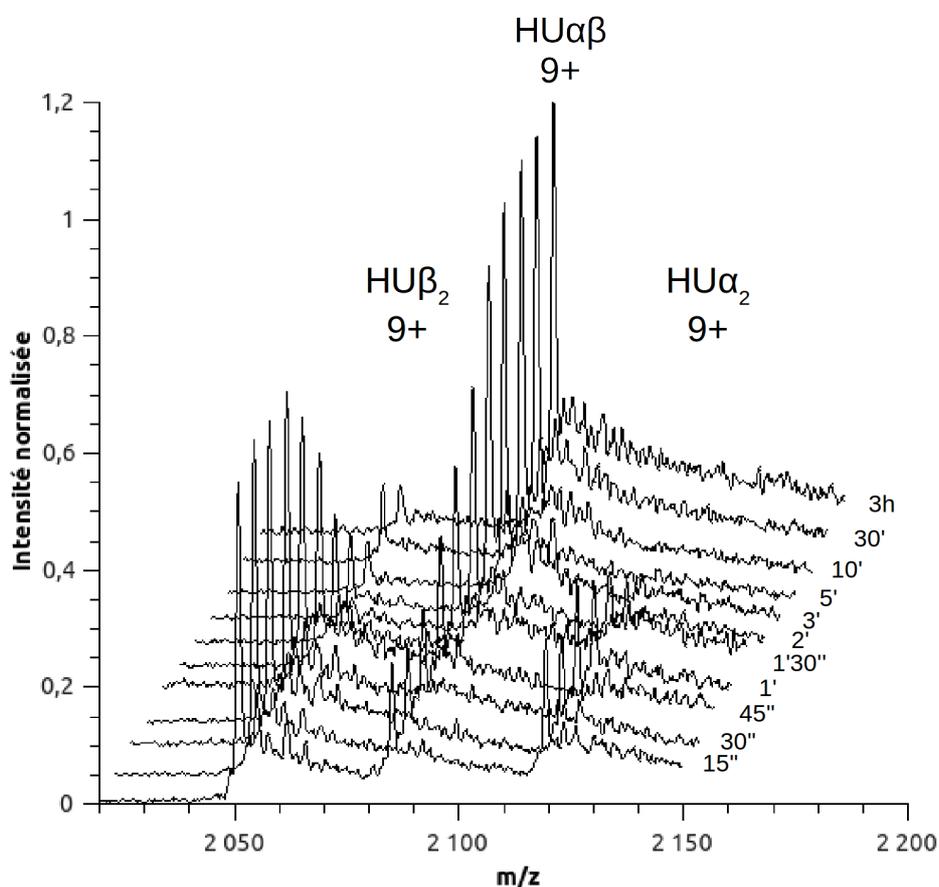
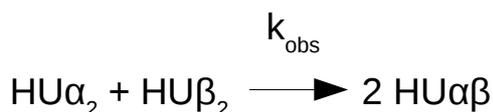


Figure 109: Suivi de la formation de l'hétérodimère au cours du temps par spectrométrie de masse en conditions natives Échange de chaînes réalisé entre les homodimères à 37°C avec 350 mM Acétate d'ammonium pH 5. Les valeurs de m/z des ions 9+ de HUβ₂, HUαβ et HUα₂ sont de 2051,1, 2085,8 et 2119,9, respectivement.

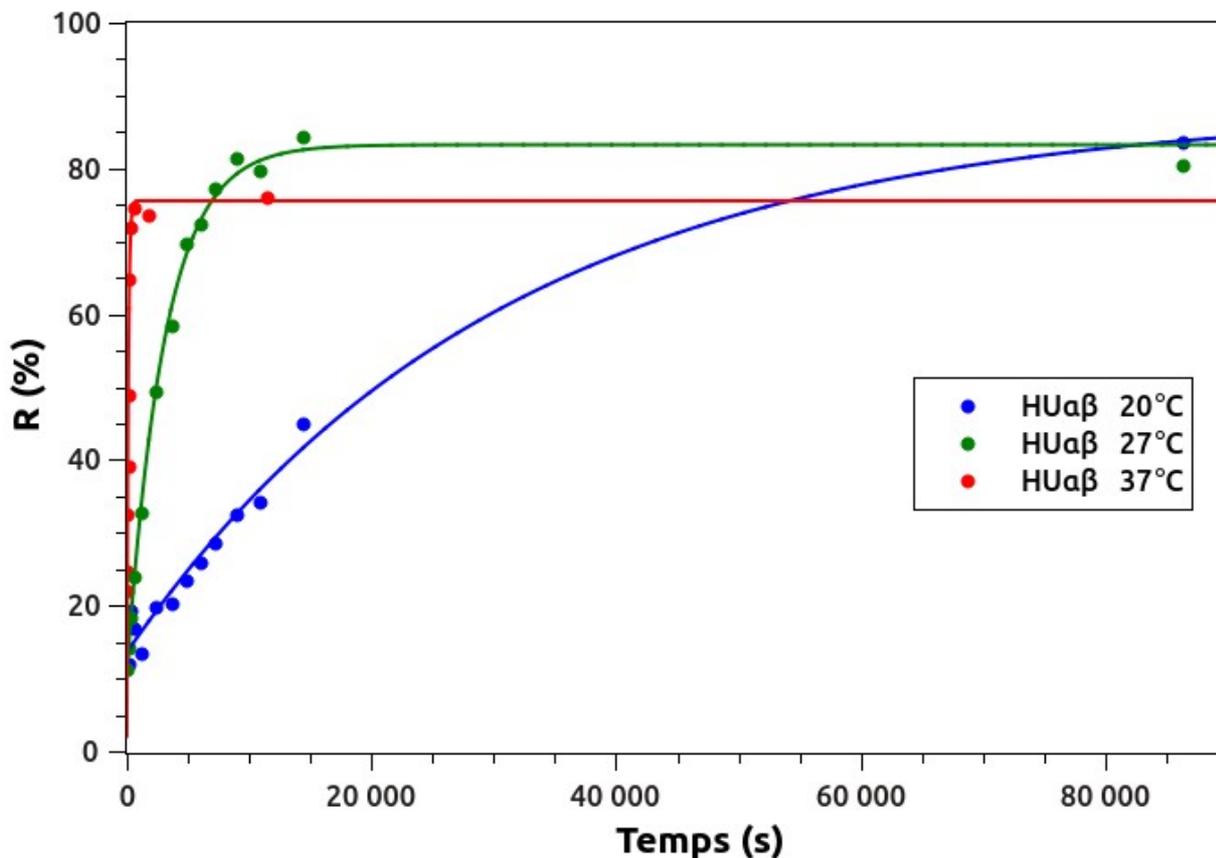


Figure 110: Cinétique de formation de l'hétérodimère, HUαβ, à différentes températures

L'équation 1 représente l'équation standard pour extraire les constantes de vitesse k_{obs} de ce type de réaction. Cependant, dans notre cas, les courbes ne passent pas par l'origine et on a systématiquement entre 15 et 20 % d'hétérodimère aux premiers points des cinétiques (**Figure 110**). Deux hypothèses ont été formulées pour expliquer cette valeur :

- (i) les échantillons d'homodimères de HU sont impurs et sont contaminés par 15 à 20 % d'hétérodimère dès le départ. Dans ce cas, un terme C_0 (concentration d'hétérodimère déjà présent dans l'échantillon) doit être introduit dans l'équation qui décrit la cinétique (équation 2).
- (ii) le blocage de la réaction par le froid n'est que partiellement efficace entre les expériences de cinétique et l'analyse des échantillons en spectrométrie de masse. Dans ce cas la durée d'échange est supérieure au temps théorique, il faut corriger la valeur de x et on utilise l'équation 3.

Étant donné que l'analyse par spectrométrie de masse (Maldi-Tof) des échantillons d'homodimères n'a révélé aucune contamination par l'hétérodimère mesurable, nous avons rejeté l'hypothèse (i).

Les échantillons ont été maintenus à froid en permanence et le trajet jusqu'à la source a été efficacement refroidi. Cependant nous ne pouvons pas maintenir ces conditions à l'étape ultime de l'analyse, à savoir l'ionisation. En effet, la haute température au niveau de l'aiguille de l'électrospray, nécessaire pour le processus d'ionisation, peut provoquer le niveau basal d'échange que l'on observe. Un temps très court à haute température est équivalent à un temps moyen à température intermédiaire ou à un temps très long à faible température en terme de taux d'échange. Ainsi, on a des valeurs du facteur B de correction sur le temps qui augmente lorsque la température diminue. Les statistiques des ajustements (effectuées avec le logiciel QtiPlot) et les paramètres cinétiques des réactions, pour trois températures, sont récapitulés dans le **Tableau 18**. Les k_{obs} obtenus varient fortement sous l'effet de la température avec des différences d'un ordre de grandeur entre chaque températures : $3,27 \cdot 10^{-5} \text{ s}^{-1}$, $3,22 \cdot 10^{-4} \text{ s}^{-1}$ et $8,89 \cdot 10^{-3} \text{ s}^{-1}$ pour les expériences menées à 20, 27 et 37°C, respectivement.

Tableau 18: Paramètres cinétiques de la formation de l'hétérodimère HU $\alpha\beta$ à partir des homodimères HU α_2 et HU β_2 Ajustés avec l'équation 3 : $y = A \cdot (1 - e^{(-k_{obs} \cdot (x+B))})$

Température	k_{obs} (s^{-1})	A (%)	B (s)	R ²
20°C	$3,27 \cdot 10^{-5} \pm 4,0 \cdot 10^{-6}$	88,2	5103	0,9857
27°C	$3,22 \cdot 10^{-4} \pm 1,5 \cdot 10^{-5}$	83,2	431	0,9975
37°C	$8,89 \cdot 10^{-3} \pm 9,3 \cdot 10^{-4}$	75,6	3	0,9888

7.1.3. Échange de chaînes entre homodimères identiques

Dans le chapitre précédent, nous avons observé que des homodimères identiques pouvaient s'associer entre eux pour former des homo-tétramères (EchU α_4 et EchU β_4). La présence de ces espèces suggère qu'un homodimère est capable d'échanger ses chaînes avec un homodimère du même type. Bien que ce type d'échange n'ait, *a priori*, pas de sens physiologique, nous avons voulu vérifier son existence dans le but de mieux comprendre le mécanisme moléculaire d'échange de chaînes. Nous avons réalisé des expériences d'échange de chaînes entre des homodimères marqués pour la RMN (¹⁵N) et non marqués (¹⁴N) et avons suivi la cinétique d'apparition du dimère hémimarké. Les masses moyennes sont 19069,9, 19308,2 et 19189,0 Da pour les dimères ¹⁴N/¹⁴N,

$^{15}\text{N}/^{15}\text{N}$ et $^{14}\text{N}/^{15}\text{N}$ de $\text{HU}\alpha_2$, respectivement et 18451,1, 18687,3 et 18569,2 Da pour les dimères $^{14}\text{N}/^{14}\text{N}$, $^{15}\text{N}/^{15}\text{N}$ et $^{14}\text{N}/^{15}\text{N}$ de $\text{HU}\beta_2$, respectivement. L'écart entre ces masses est suffisant pour pouvoir discriminer chaque espèce en spectrométrie de masse en conditions natives et calculer le rapport R (l'intensité du pic de l'ion 9+ de l'homodimère héli-marqué divisée par la somme des intensité des pics 9+ des trois homodimères) au cours du temps. Les cinétiques d'échanges ont été réalisées à différentes températures et sont présentées **Figure 111**. Les statistiques d'ajustement et les paramètres cinétiques obtenus sont présentés dans le **Tableau 19**.

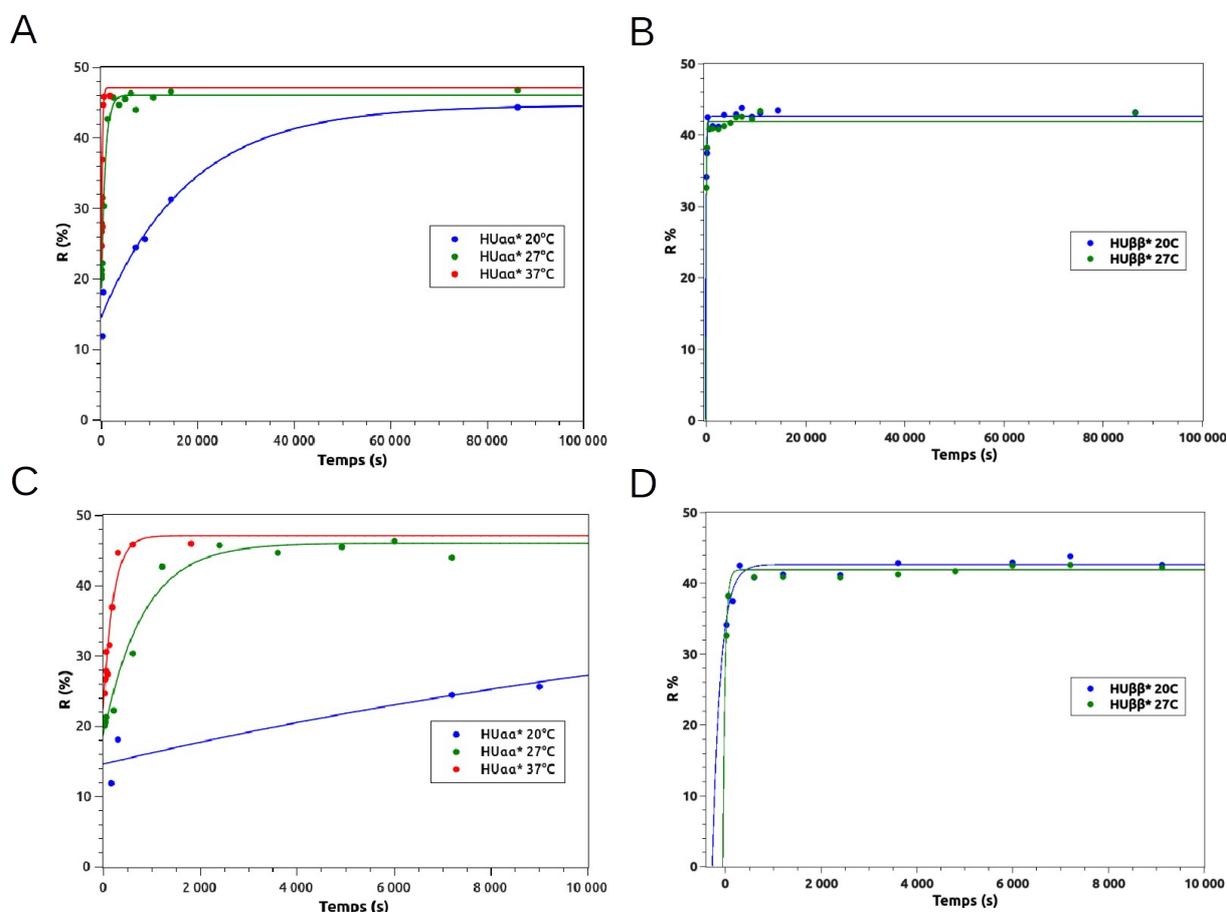


Figure 111: Cinétiques de formation des homodimères héli-marqués à différentes températures A/B- Formation des espèces héli-marquées de $\text{HU}\alpha_2$ et $\text{HU}\beta_2$, respectivement. C/D zoom dans la gamme de temps 0-10000 s pour $\text{HU}\alpha_2$ et $\text{HU}\beta_2$, respectivement. Les cinétiques ont été réalisées à trois températures pour $\text{HU}\alpha_2$ (20, 27 et 37°C en bleu, vert et rouge, respectivement) et à deux températures pour $\text{HU}\beta_2$ (20 et 27°C en bleu et vert, respectivement).

Tableau 19: Paramètres cinétiques d'échange d'un homodimère avec lui-même à différentes températures.

Dimère	Température	k_{obs} (s^{-1})	A (%)	B (s)	R^2
HU $\alpha\alpha^*$	20°C	$5,48 \cdot 10^{-5} \pm 1,3 \cdot 10^{-5}$	44,6	7213	0,9708
	27°C	$1,21 \cdot 10^{-3} \pm 1,7 \cdot 10^{-4}$	46,0	432	0,9820
	37°C	$4,89 \cdot 10^{-3} \pm 1,2 \cdot 10^{-4}$	47,1	133	0,9457
HU $\beta\beta^*$	20°C	$5,78 \cdot 10^{-3} \pm 2,1 \cdot 10^{-3}$	42,6	277	0,8386
	27°C	$2,07 \cdot 10^{-2} \pm 6,4 \cdot 10^{-3}$	41,9	58	0,9118

Les cinétiques de formation des espèces hémi-marquées atteignent des plateaux situés entre 40 et 50 % de dimères hémi-marqués (**Figure 111**). L'affinité des chaînes pour elles-mêmes étant indifférentes au marquage, et les T_{m1} étant également les mêmes pour chaque dimères dans les mélanges réactionnels, il y a autant de chance d'échanger avec chaque type de dimère différent. A l'équilibre, on atteint une répartition statistique de l'ordre de 25 %, 50 % et 25 % pour les espèces marquées $^{15}\text{N}/^{15}\text{N}$, $^{15}\text{N}/^{14}\text{N}$ et $^{14}\text{N}/^{14}\text{N}$, respectivement.

Tableau 20: Comparaison des constantes de vitesse entre les différentes réactions d'échange

Température	k_{obs} HU $\alpha\beta$ (s^{-1})	k_{obs} HU $\alpha\alpha^*$ (s^{-1})	k_{obs} HU $\beta\beta^*$ (s^{-1})	$k_{\text{obs}} / k_{\text{obs}}$ $\alpha\alpha / \alpha\beta$	$k_{\text{obs}} / k_{\text{obs}}$ $\beta\beta^* / \alpha\beta$	$k_{\text{obs}} / k_{\text{obs}}$ $\alpha\alpha^* / \beta\beta^*$
20°C	$3,27 \cdot 10^{-5}$ $\pm 4,0 \cdot 10^{-6}$	$5,48 \cdot 10^{-5}$ $\pm 1,3 \cdot 10^{-5}$	$5,78 \cdot 10^{-3}$ $\pm 2,1 \cdot 10^{-3}$	1,7	176	105
27°C	$3,22 \cdot 10^{-4}$ $\pm 1,5 \cdot 10^{-5}$	$1,21 \cdot 10^{-3}$ $\pm 1,7 \cdot 10^{-4}$	$2,07 \cdot 10^{-2}$ $\pm 6,4 \cdot 10^{-3}$	3,8	64	17
37°C	$8,89 \cdot 10^{-3}$ $\pm 9,3 \cdot 10^{-4}$	$4,89 \cdot 10^{-3}$ $\pm 1,2 \cdot 10^{-4}$	ND	0,6	ND	ND

Si l'on compare les k_{obs} de formation des homodimères hémi-marqués à ceux de formation de l'hétérodimère (**Tableau 20**), on obtient des valeurs du même ordre de grandeur entre HU $\alpha\alpha^*$ et HU $\alpha\beta$, avec des rapports de constante de vitesse entre 0,6 et 4. Par contre, dans le cas de HU $\beta\beta^*$, la vitesse d'échange est d'un à deux ordres de grandeur plus grande que celle des autres dimères aux mêmes températures.

7.1.4. Accord entre les mécanismes et constantes cinétiques

Description des mécanismes d'échange envisagés

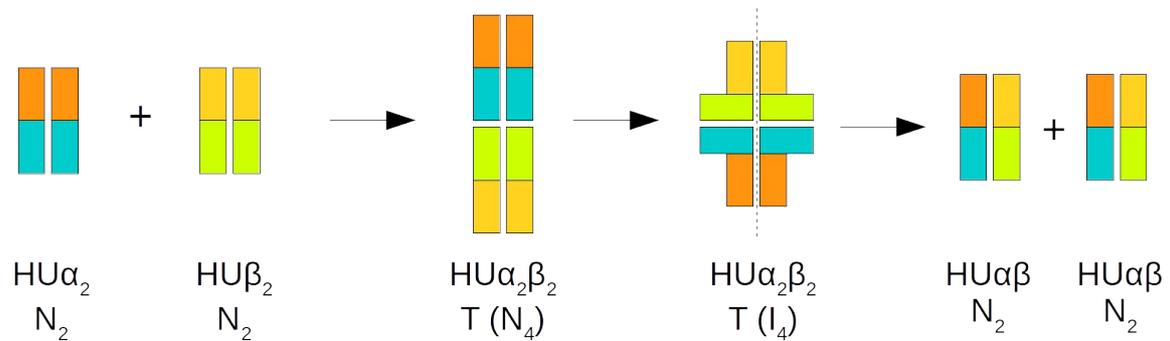
D'après nos hypothèses, le mécanisme d'échange serait basé sur les conformations I_2 des homodimères. Cependant, nous ne pouvons pas totalement exclure des modèles dits « alternatifs », avec des conformations de départ différentes (à partir des conformations natives, N_2 , ou dénaturées, $2D$). Ces modèles de mécanisme d'échange sont représentés schématiquement sur la **Figure 112**. La question est maintenant de savoir en quoi ces différents mécanismes d'échange sont cohérents avec les observations des cinétiques d'échange. Les trois modèles d'échanges envisagés proposent un mécanisme partant de l'une des trois conformations possibles des dimères de HU.

Dans le cas du modèle 1 (**Figure 112**), les homodimères natifs s'associent dans un tétramère N_4 (dimère de dimères natifs N_2). Des modèles putatifs d'association de dimères N_2 ont été proposés **Figure 107 p 218** sur la base des empilements cristallins de HU de différents organismes. Les dimères au sein de ce tétramère changent ensuite de conformation (passage sous forme I_2) pour former un tétramère I_4 (dimère de dimère I_2 , modèle **Figure 108 p 219**) qui pourra ensuite se dissocier pour former les deux hétérodimères.

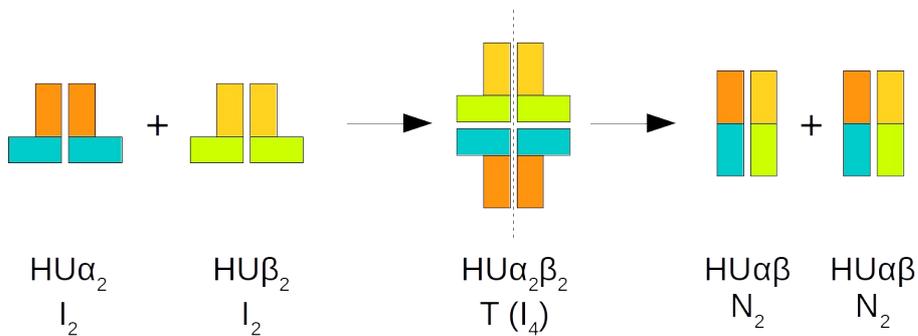
Le modèle d'échange 2, celui que nous privilégions, est celui basé sur les conformations intermédiaires I_2 . Dans ce modèle, les homodimères adoptent une conformation ouverte, I_2 , qui s'associent sous la forme d'un tétramère I_4 en interagissant par leurs domaines HTH (**Figure 108 p 219**). Puis ce tétramère est ensuite dissocié selon l'autre axe (au niveau des brins β_3) pour former les deux hétérodimères.

Enfin, le modèle d'échange 3 est simplement basé sur l'association des monomères dénaturés (D) pour former l'hétérodimère.

Modèle d'échange 1 : (N_2)



Modèle d'échange 2 : (I_2)



Modèle d'échange 3 : (D)

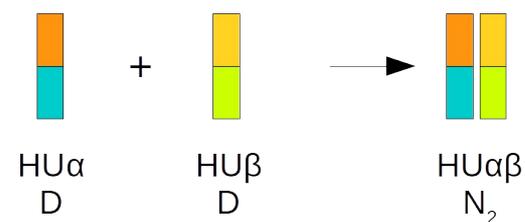


Figure 112: Représentation schématique des modèles de mécanismes d'échange selon trois conformations de départ N_2 : état natif dimérique, I_2 état intermédiaire dimérique, D état dénaturé monomérique, $T(I_4)$ Tétramère (composé de deux dimères I_2) et $T(N_4)$ Tétramère (composé de deux dimères N_2).

Déterminations des étapes limitantes

Pour comparer les 3 modèles, on considère que la vitesse de transition de conformation $N_2 \rightarrow I_2 \rightarrow 2D$ est très rapide (de l'ordre de la milliseconde à la seconde) et est non limitante par rapport à la vitesse globale de réaction (de l'ordre de la minute, à l'heure). Ainsi, à chaque instant de l'échange, les proportions de chaque conformation respectent l'équilibre thermodynamique du dimère considéré à une température donnée. C'est l'approximation de l'équilibre rapide (AER).

Le point commun de ces trois modèles est qu'ils commencent tous par une étape d'association bimoléculaire (association de dimères en tétramères ou de monomères en dimères). Ces étapes sont dépendantes de la concentration des espèces à associer. Plus la concentration de ces espèces est élevée, plus les cinétiques seront rapides.

Dans le cas des modèles 1 et 2, l'étape de dissociation du tétramère est également une étape bimoléculaire dépendante de la concentration. Dans ce cas, plus la concentration en protéine est élevée, plus la cinétique de dissociation sera lente. Comme la concentration en tétramère reste constamment faible au cours de la réaction d'échange, on en déduit que l'association des dimères en tétramère est plus lente que la dissociation du tétramère en hétérodimères. C'est l'approximation de l'état quasi-stationnaire.

Pour chaque modèle, c'est donc l'étape d'association qui constitue l'étape limitante du mécanisme d'échange.

Confrontation des modèles

On cherche donc à comparer les constantes de vitesse (k_{obs}) des réactions d'échanges des homodimères ($\alpha\alpha^*$, $\beta\beta^*$) avec les rapports de concentration théoriques des espèces de départ (différences de concentration entre les conformations des homodimères $[N_2]$, $[I_2]$, $[2D]$) en se basant sur les T_m des homodimères obtenus par DSC dans le travail de Ramstein J. et collaborateurs¹⁶⁵.

Dans le cas du modèle 1, l'étape d'association bimoléculaire concerne les conformations N_2 . Pour $HU\alpha_2$ le T_{m1} est de 41°C alors qu'il est de 27°C pour $HU\beta_2$. Pour toutes les températures, la proportion en forme N_2 est donc plus importante pour $HU\alpha_2$ que pour $HU\beta_2$. Pourtant, la vitesse d'échange est beaucoup plus rapide pour $HU\beta_2$ que pour $HU\alpha_2$. On élimine donc ce mécanisme putatif qui ne respecte pas les observations expérimentales.

Pour le modèle 2, l'étape d'association bimoléculaire part des conformations I_2 . Or le dimère ayant la proportion de I_2 la plus élevée est $HU\beta_2$. C'est également l'échange entre les homodimères de $HU\beta_2$ qui est le plus rapide. On a une corrélation entre l'abondance relative de I_2 entre les homodimères $HU\alpha_2$ et $HU\beta_2$ et la vitesse d'échange de ces homodimères avec des homodimères du même type. Ce modèle est en accord avec les observations expérimentales de cinétique.

Dans le cas du modèle 3, la formation de l'hétérodimère à partir des monomères n'est pas cohérente avec les vitesses de formation des homodimères héli-marqués. En effet, les T_m des homodimères de $HU\alpha_2$ et $HU\beta_2$ sont 50 et 56°C respectivement. C'est donc l'homodimère $HU\beta_2$ qui présente le moins de forme dénaturée monomérique et qui pourtant présente la cinétique d'échange la plus rapide. Ce modèle d'échange n'est pas en accord avec les données expérimentales.

7.2. Conclusion

Dans ce chapitre, nous avons montré que l'on pouvait suivre l'échange de chaînes entre les dimères de HU par spectrométrie de masse en conditions natives. Non seulement les homodimères peuvent échanger leurs chaînes pour former des espèces hétérodimériques, mais les homodimères sont également capables d'échanger avec eux-mêmes. Les cinétiques d'échanges sont fortement modulées par la température comme le montrent les **Figures 110 et 111**. Cet effet de température sur la cinétique d'échange est visible sur les k_{obs} qui varient d'un ordre de grandeur entre chaque température testée.

Parmi les mécanismes d'échange envisagés à partir des trois conformations en équilibre des dimères de HU, seul le modèle 2 (**Figure 112**), basé sur l'association des dimères dans leurs conformations intermédiaires est compatible avec les variations de vitesse d'échange entre les homodimères.

À retenir

<p style="text-align: center;">Objectif du chapitre : Suivre l'échange de chaînes et la formation de l'hétérodimère $HU_{\alpha\beta}$</p>
<p><i>Rq : ce chapitre a nécessité que je produise les formes simplement marquées(^{15}N) ou non marquée de chaque chaîne</i></p>
<ul style="list-style-type: none">- Suivre la cinétique de formation de l'hétérodimère $HU_{\alpha\beta}$ par SM en conditions natives <p>→ Estimation des paramètres cinétiques (kobs) à 3 températures</p>
<ul style="list-style-type: none">- Suivre la cinétique de formation des homodimères $HU_{\alpha\alpha}$ et $HU_{\beta\beta}$ (nécessite l'hémimarquage) <p>→ Estimation des paramètres cinétiques (kobs) à plusieurs températures</p> <p>Conclusion 1 : La vitesse d'échange pour $HU_{\beta\beta}$ est 1 à 2 ordres de grandeurs plus grande que pour $HU_{\alpha\alpha}$ et $HU_{\alpha\beta}$</p>
<ul style="list-style-type: none">- Proposer un modèle de mécanisme d'échange. 3 mécanismes d'échange sont envisagés, à partir des confirmations dimériques N_2, dimérique I_2 ou monomérique D <p>→ Quel que soit le modèle, l'étape d'association est l'étape limitante du mécanisme d'échange</p> <p>Conclusion 2 : seul le modèle basé sur l'association des dimères dans leurs conformations intermédiaires I_2 est compatible avec les données expérimentales, en particulier avec les variations de vitesse d'échange entre les homodimères</p>

Conclusions & Perspectives

L'objectif principal de mon travail de thèse était de caractériser le mécanisme d'échange de chaînes entre les homodimères de HU d'*E. Coli*, $HU\alpha_2$ et $HU\beta_2$, essentiel à la formation de l'hétérodimère, $HU\alpha\beta$. Pour ce travail, je me suis basé sur le modèle d'échange proposé par Ramstein et collaborateurs¹⁶⁶, présenté sur la **Figure 113**. J'ai cherché à caractériser chaque étape du mécanisme d'un point de vue structural et cinétique en utilisant différentes techniques de caractérisation biophysique et biochimique.

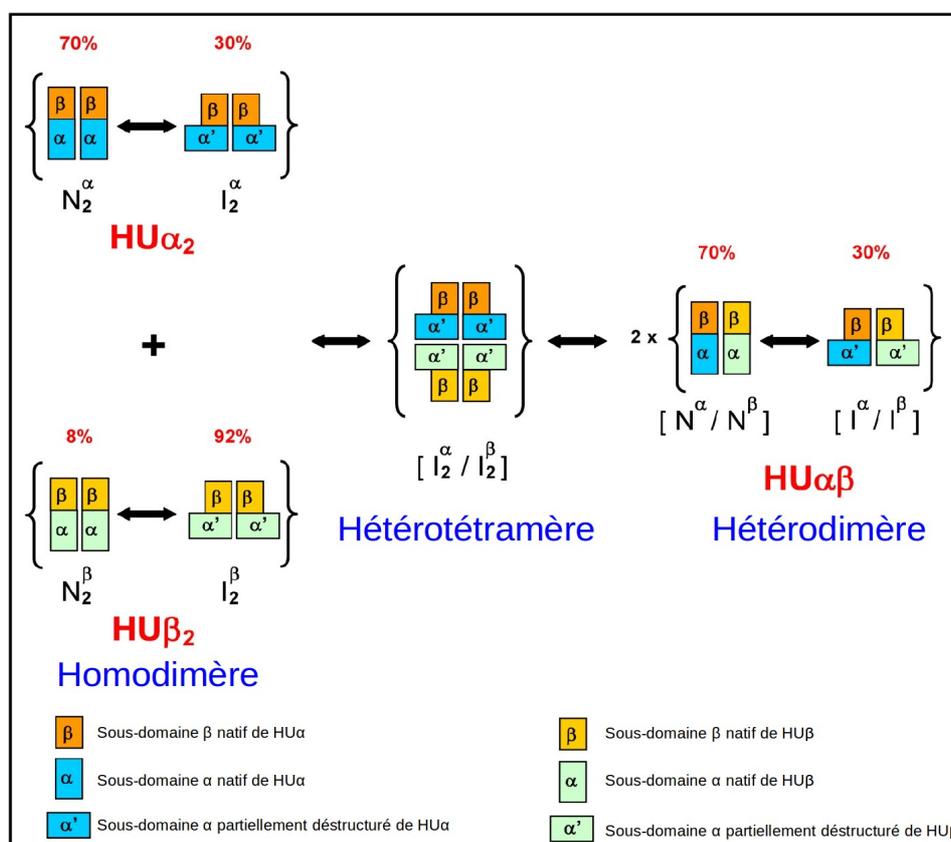


Figure 113: Modèle du mécanisme d'échange de chaînes entre les dimères de HU d'*E. coli* (Adapté de Ramstein et collaborateurs)¹⁶⁶

Au début de ma thèse, seules les structures des dimères $HU\alpha_2$ (1MUL, R=2,30Å) et $HU\alpha\beta$ (2O97, R=2,45Å) étaient disponibles chez *E. coli*. Des structures de HU de 5 autres organismes ainsi que celles d'homologues structuraux (IHF, TF1, Hbb, BVU_2165) sont également disponibles dans la PDB.

Nous avons résolu deux structures cristallographiques originales de HU : (i) $HU\beta_2$ d'*E. coli* (4P3V, R=1,25Å) et HU de *L. lactis* (R= 2,19Å). La structure de $HU\beta_2$ complète la caractérisation structurale des HU d'*E. coli* dans leur conformation native N_2 (**Figure 110 A**). La résolution obtenue est la meilleure rapportée pour les protéines de cette famille. L'analyse des structures de HU obtenues au cours de cette thèse et de celles disponibles dans la PDB a permis de mettre en évidence une corrélation inverse entre la taille d'une cavité et le Tm_1 . Ainsi le remplissage de cette cavité, par mutagenèse ou par fixation d'une petite molécule, pourrait être un moyen de stabiliser la forme N_2 et de bloquer la transition vers I_2 et la formation de l'hétérodimère (**mutation L29F Figure111, ou Figure 112**).

La caractérisation de la structure de l'intermédiaire I_2 , partiellement dénaturée, n'était pas adaptée à la cristallographie des rayons X. Nous avons opté pour une étude par RMN et par dynamique moléculaire de la transition vers I_2 . La caractérisation initiale de la conformation N_2 par RMN a permis de mettre en évidence l'instabilité des liaisons hydrogène dans la partie C-terminale (hélice α_3 et une partie du brin β_3) à partir de leurs coefficients de température. Au cours de l'enregistrement de ces expériences ^{15}N -HSQC réalisées à différentes températures, nous avons pu mettre en évidence l'apparition de signaux spécifiques de la conformation I_2 en échange lent avec la conformation N_2 (à l'échelle de la RMN). Les expériences ZZ- ^{15}N -HSQC ont permis de calculer les constantes cinétiques d'échange $N_2 \rightarrow I_2$ pour $HU\alpha_2$ et d'estimer qu'elles étaient dans la gamme de temps millisecondes à secondes pour $HU\beta_2$ et $HU\alpha\beta$ (**Figure 110 A**).

L'attribution et la caractérisation préliminaire de la conformation I_2 des homodimères (CSP, prédiction de structures secondaires, NOEs) croisées avec les résultats des expériences de simulation de dynamique moléculaire en « gros grain » ont permis d'élaborer un modèle de la conformation I_2 (**Figure 110 A**) où la partie HTH est ouverte mais conserve ses structures secondaires et s'accompagne d'une déstructuration de la partie C-terminale (hélice α_3 et une partie du brin β_3). Cette dernière observation est cohérente avec les expériences réalisées précédemment par Ramstein et collaborateurs¹⁶⁶ et Garnier et collaborateurs¹⁶⁷, qui avaient mis en évidence une perte de structuration secondaire en hélice α pour la conformation I_2 (CD et MD).

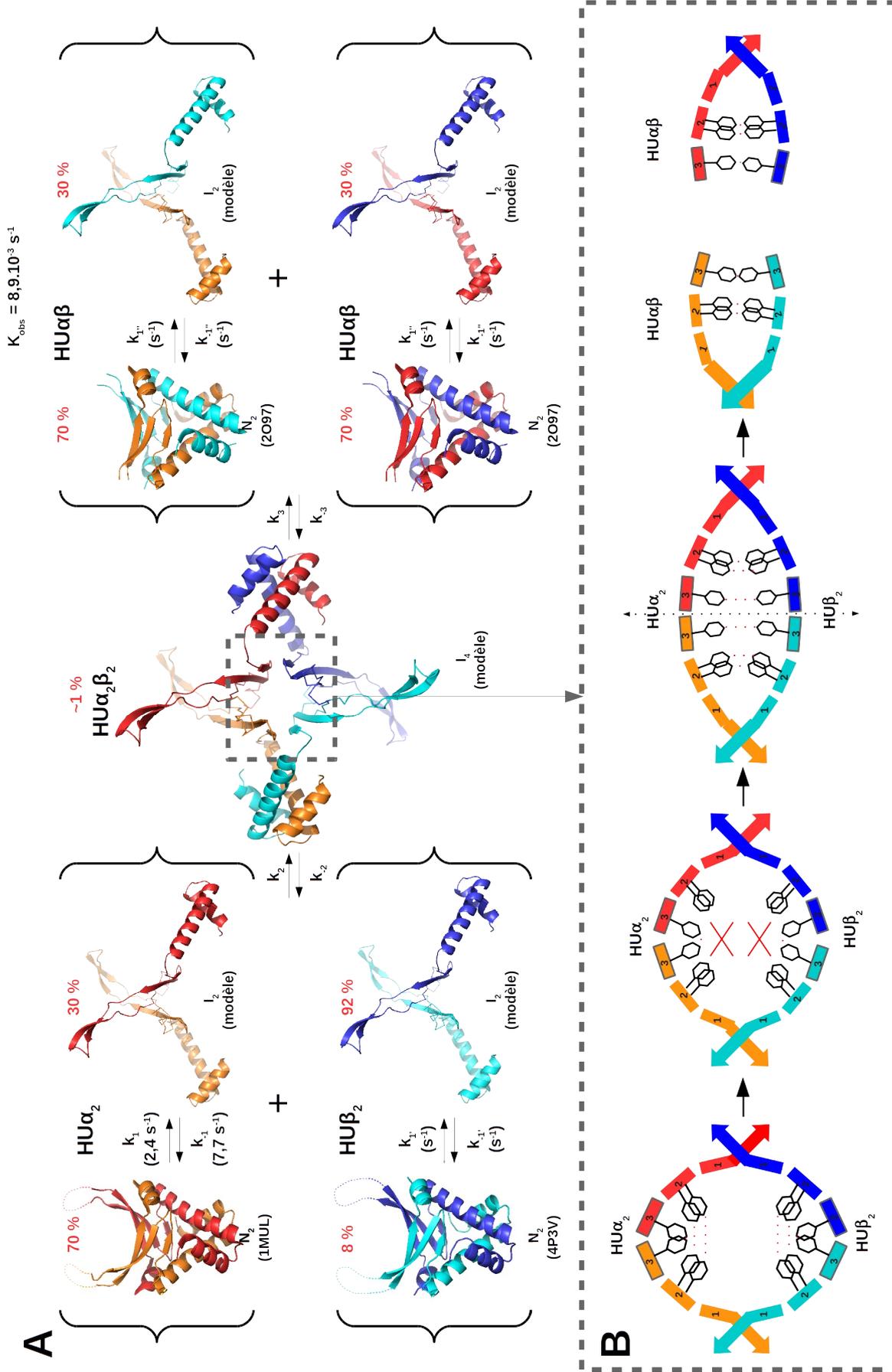


Figure 110: Bilan structural et cinétique du mécanisme d'échange de chaînes pour la formation de l'hétérodimère à 310K (A) et mécanisme de l'effondrement de la poche hydrophobe du tétramère jusqu'à la dissociation en hétérodimères (B)

L'existence d'un état tétramérique des protéines HU a été confirmée au cours de cette thèse, d'une part, par des expériences de pontage chimique et d'autre part, par des analyses en spectrométrie de masse en conditions natives sur les dimères de type sauvage d'*E. Coli*. La faible proportion apparente de tétramère en solution (~1%), transitoire et instable, constitue un obstacle à sa caractérisation structurale. Nous avons conçu et produit F79C, dans le but de conduire l'échange dans un cul-de-sac de réaction en faveur de l'accumulation du tétramère. La nécessité de produire et purifier HU α_2 sous sa forme F79C nous a conduit à mettre en place le sous-clonage dans un vecteur possédant une étiquette d'affinité (MBP), la mise au point de la production/purification de la protéine de fusion et la mise au point du protocole d'oxydation permettant la formation d'un pont disulfure dans le dimère natif (oxydation au cuivre Cu²⁺). Bien que cette mutation n'a que peu enrichi la proportion de tétramère (passage de 1 à 2-3 %), tous les outils ont été mis en place pour l'élaboration d'autres mutants ou d'une combinaison de mutations qui pourraient permettre d'obtenir une conformation tétramérique stabilisée et adaptée à la caractérisation structurale.

Sur la base des modèles de la conformation I₂ des homodimères, un modèle de tétramère a été proposé. Dans ce modèle, l'association de deux homodimères ouverts au niveau du HTH permet de reformer les domaines HTH inter-dimère (de type hétérodimérique) par échange de domaines (« domain swapping »). L'association des deux cœurs hydrophobes engendrerait la formation d'une très grande cavité au sein du tétramère. Le collapse hydrophobe au niveau de cette cavité permettrait le transfert des interactions hydrophobes, notamment des résidus aromatiques, d'un dimère à l'autre (**Figure 110 B**). La dissociation (le long de l'axe vertical représenté par une double flèche en pointillé) de ce tétramère conduirait ensuite à la formation des hétérodimères. Par ailleurs, la région de contact au niveau du brin β_3 (**encadré en gris Figure 110 B**) est prédite partiellement déstructurée au niveau des structures secondaires dans la conformation I₂ (prédiction d'angle dièdre en RMN) et pourrait aider à la dissociation du tétramère en hétérodimères.

Le mécanisme présenté ici reste putatif et des études supplémentaires seront nécessaires pour sa caractérisation complète. En particulier, les structures des conformations intermédiaires et tétramériques de HU n'ont pas encore été résolues et demeurent des modèles. Ni la RMN ni la dynamique moléculaire n'ont pu aboutir à des modèles satisfaisant de la conformation intermédiaire I₂. Le nombre de NOEs et la difficulté de leur attribution sans ambiguïté n'a pas permis de calculer une structure expérimentale. De plus l'échelle de temps de la transition N₂ → I₂ est trop longue pour être simulée directement en dynamique moléculaire. Une approche prometteuse réside dans le

couplage des deux méthodes : la dynamique moléculaire sous contraintes RMN. En utilisant les déplacements chimiques et éventuellement d'autres données (NOEs, RDC), on peut contraindre la simulation de dynamique moléculaire par des données expérimentales et parvenir à un modèle réaliste de la conformation I_2 .

La résolution de la structure de I_2 permettrait de proposer un meilleur modèle du tétramère de HU (par amarrage moléculaire ou « docking ») et de proposer des points de mutagenèse inter-dimères pour la stabilisation covalente de cette espèce et sa caractérisation structurale. Tous les outils sont désormais mis en place pour l'élaboration de ces mutants (production avec étiquette MBP, purification, oxydation du ponts disulfure). On peut aussi utiliser la mutagenèse dirigée pour moduler la transition $N_2 \rightarrow I_2$. Si la taille de la poche influe effectivement sur le Tm_1 , on pourrait la réduire par mutagenèse en introduisant des résidus des chaînes latérales hydrophobes plus longues : par exemple L29F, **Figure 111**. De plus, si le mouvement des hélices du HTH est également nécessaire à la transition vers I_2 , comme le suggère les expériences de dynamique moléculaire, l'introduction de ponts disulfure dans le HTH pourrait permettre de restreindre leurs mouvements et d'empêcher la transition (mutants A24C ou N2C+Q43C). A l'inverse, certaines mutations permettraient de déstabiliser N_2 et favoriser la transition vers I_2 . Par exemple, si la déstructuration de l'hélice α_3 C-terminale est une étape nécessaire à la transition vers I_2 , un mutant exempt de cette hélice (ΔC -ter [82-90] **Figure 111**) doit être capable de transiter plus facilement vers la conformation intermédiaire qu'une protéine de type sauvage.

Le modèle de mécanisme d'échange présenté **Figure 110** est en accord avec toutes les données structurales et cinétiques recueillies au cours de cette thèse ou disponibles dans la littérature. L'échange de chaînes par un état dimérique intermédiaire représente une fonction complètement originale d'une conformation partiellement désordonnée. L'intérêt majeur de ce mode d'échange, par rapport à un mécanisme passant par une étape monomérique, est la diminution de la barrière énergétique à franchir pour la formation de l'hétérodimère. En effet, la dissociation des dimères de HU requiert plus d'énergie que la transition vers une conformation intermédiaire ($Tm_2 > Tm_1$). De plus, si la conformation intermédiaire (I_2) est plus sensible à la dégradation par les protéases, telles que Lon, que la conformation Native (N_2), on peut facilement imaginer que l'espèce monomérique le soit encore davantage (2D). Il est donc probable que le peu de HU monomérique dénaturée, si cette espèce est présente *in vivo*, soit très rapidement dégradée. Le mécanisme d'échange de chaînes par des espèces dimériques représenterait ainsi une voie alternative et efficace pour la formation de l'hétérodimère, $HU\alpha\beta$.

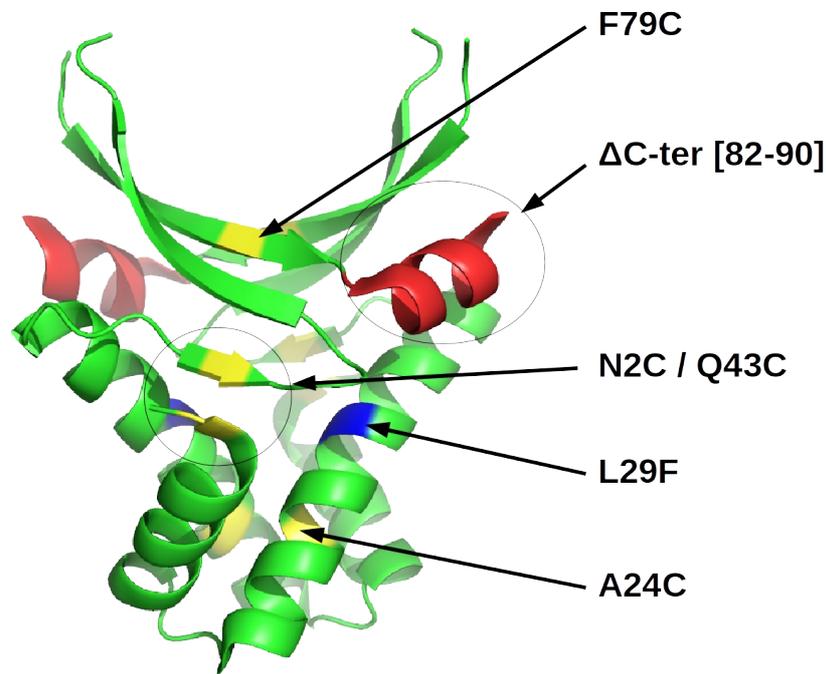


Figure 111 : Cibles de mutagénèse dirigée basées sur les données structurales En jaune les mutations en cystéine pour la formation de ponts disulfure, en bleu une substitution en phénylalanine pour combler la cavité 1, en rouge la délétion de l'hélice C-terminale pour favoriser la transition $N_2 \rightarrow I_2$.

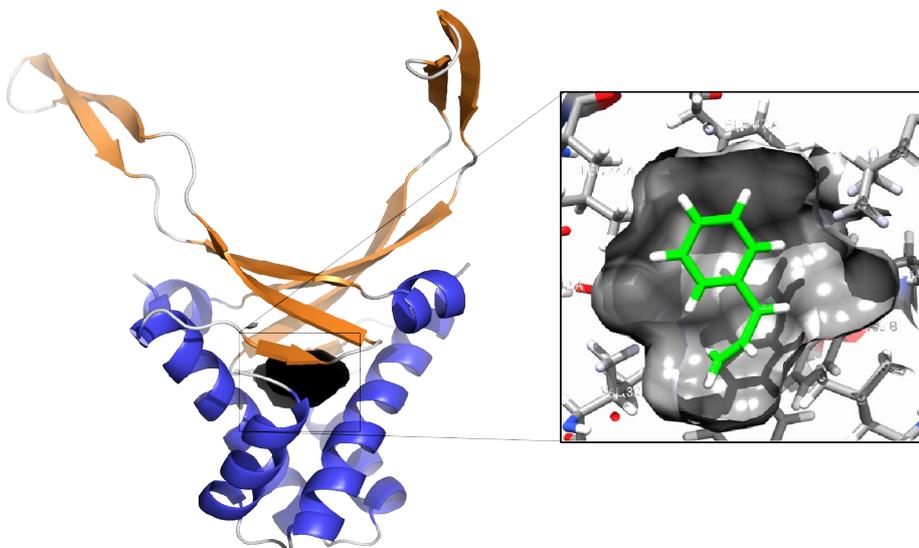


Figure 112 : Ciblage de la poche hydrophobe par des ligands criblage du modèle de $HU\beta_2$ par Dockblaster (<http://blaster.docking.org/>)

La possibilité de bloquer le mécanisme d'échange de chaînes permettrait de mieux pouvoir étudier le rôle fonctionnelle de HU $\alpha\beta$ *in vivo*. En effet, on ne peut pas s'affranchir de l'espèce hétérodimérique ($\Delta hupAB$) sans perdre les espèces homodimériques de HU. Il est donc difficile de déterminer dans ces conditions si le phénotype associé à la suppression des deux gènes est le résultat de l'absence d'hétérodimère ou d'un effet additionnel de l'absence de chacune des isoformes de HU. En introduisant une mutation adaptée bloquant l'échange de chaînes, on pourrait étudier le phénotype associé à l'absence spécifique de l'hétérodimère en conservant des homodimères fonctionnels.

Par ailleurs, si la mutagenèse dirigée sur HU permet d'inhiber ou d'activer significativement la transition $N_2 \rightarrow I_2$, au point d'obtenir 100 % de l'une ou l'autre conformation dans des conditions physiologiques, ces mutants seront également intéressants à étudier du point de vue fonctionnel. En effet, dans le cadre de ma thèse, nous nous sommes intéressés au rôle de I_2 dans le mécanisme d'échange de chaînes vers la formation de l'hétérodimère, mais il n'est pas exclu que cette conformation I_2 puisse être associée à d'autres fonctions biologiques. Par exemple, elle pourrait permettre d'établir des interactions avec des partenaires protéiques alternatifs par rapport à la conformation native. En effet, la nature partiellement désordonnée de la conformation I_2 peut être mise en parallèle avec les protéines HMGA. Ces protéines d'architecture de l'ADN sont intrinsèquement désordonnées et constituent des réseaux d'interactions très large avec des partenaires impliqués dans des fonctions associées au métabolisme des acides nucléiques (transcription, réparation, remodelage de la chromatine, maturation des ARN).²³ On peut également imaginer que la dynamique conformationnelle de cette protéine joue un rôle dans son large panel de reconnaissance de différentes structures de l'ADN (jonction 4 voies, lacunes, hernies, etc...)

Il est intéressant de noter que d'autres protéines histones ou de type histones présentent des conformations intermédiaires : H2A-H2B, Fis, hUBF HMG Box-5 et Alba2.^{43,88,90,249-251} Bien que ces protéines possèdent des fonctions similaires, elles sont très distantes du point de vue phylogénétique. Il est possible que la conformation intermédiaire ait un rôle fonctionnel commun à ces protéines (échange de chaînes ou autre) et que l'existence de cette forme structurale soit le résultat d'une convergence évolutive.

Outre les fonctions endogènes des protéines architecturales de l'ADN, certaines protéines telles que HU sont retrouvées dans le milieu extracellulaire (lyse et/ou excrétion).²⁵² Le travail de Goodman et collaborateurs^{253,254} a permis de mettre en évidence que les protéines de la famille DNABII, dont fait partie HU, avaient un effet structurant sur les biofilms bactériens, en se liant à l'ADN extracellulaire présent dans les biofilms, et pouvaient être ciblées par des anticorps polyclonaux contre des protéines de la famille. Plus spécifiquement, des anticorps dirigés contre IHF, l'homologue structural de HU, ont été utilisés avec succès pour la dispersion de biofilms de différentes bactéries, même dans le cas de souches bactériennes dépourvues d'IHF. Nous pensons que l'effet observé de déstructuration des biofilms provient en réalité du ciblage de HU par les anticorps (réaction croisée). De plus, l'ADN extracellulaire contenu dans ces biofilms présente des topologies semblables à des ADN cruciformes. Or, HU est la seule protéine de la famille DNABII qui ait une affinité spécifique pour ce type d'ADN. Pour vérifier le rôle de HU dans la structuration des biofilms bactériens, j'ai produit les quantités de protéines HU nécessaires pour faire produire les anticorps anti-HU. Les premiers essais sur des biofilms bactériens sont actuellement en cours (*Collaboration T. Jouenne Rouen*). Si les anticorps anti-HU s'avèrent efficaces pour la déstructuration des biofilms, ils pourraient ouvrir une nouvelle voie thérapeutique dans la lutte contre les infections associées à des biofilms.

Chapitre Connexe : La protéine HU d'*Escherichia coli* est capable de moduler le système de Réparation de l'ADN par Excision de Bases

Ce chapitre est consacré à une étude fonctionnelle sur HU suggérant un rôle direct de la protéine dans la réparation de l'ADN chez *Escherichia coli*. Il est présenté sous la forme d'un article à soumettre.

7.3. Introduction

L'un des thèmes majeurs de recherche développé dans l'équipe de Bertrand Castaing concerne l'étude des relations structure-fonction des ADN glycosylases, des enzymes de réparation de l'ADN qui initient le système par excision de base (BER).²⁵⁵ Ces enzymes sont capables d'éliminer sélectivement des bases modifiées (ou normales mais mal appariées) dont l'apparition dans l'ADN peut être physiologique (alkylation en C5 de la cytosine par exemple) ou induite par des stress endogènes ou exogènes (oxydation, alkylation, déamination et dégradation). Ces enzymes sont ainsi nécessaires au maintien de la stabilité génétique (réparation), au contrôle épigénétique de l'expression des gènes (déméthylation active de l'ADN en participant à l'excision de la 5-méthylcytosine) et à l'hypermutation somatique (maturation de l'antigénicité des immunoglobulines par mutation contrôlée dans la région hypervariable des anticorps couplée à la sélection clonale).

L'un des axes de cette recherche sur ce thème a pour objectif d'identifier des facteurs cellulaires ou de petites molécules naturelles ou synthétiques (pour le ciblage cosmétique ou pharmacologique) capables de moduler l'activité de ces enzymes. L'existence de partenaires cellulaires des ADN glycosylases a été suggérée par l'observation paradoxale que beaucoup de ces enzymes isolées présentent une efficacité enzymatique médiocre alors qu'elles sont très efficaces in vivo. Ce faible rendement enzymatique est généralement dû à un très mauvais turnover de ces enzymes (recyclage après l'acte catalytique) car elles peuvent garder une très forte affinité pour leurs produits de réaction (le site abasique résultant de l'excision de la base endommagée ou la cassure simple-brin résultant de l'excision de la base et du sucre). Ainsi, il a été montré que des protéines ayant des affinités fortes pour les produits de réaction des ADN glycosylases pouvaient stimuler leur activité

in vitro. Chez les procaryotes comme chez les eucaryotes, les AP endonucléases du système BER telles que l'exonucléase III et l'endonucléase IV de *E. coli* ou leur homologue humain APE1 sont capables de stimuler *in vitro* les ADN glycosylases (2-5).²⁵⁶⁻²⁵⁹ Les protéines XRCC1 et PARP1 chez les eucaryotes supérieurs (protéines impliquées dans le BER et capables de reconnaître les cassures simple-brin dans l'ADN) sont également capables de stimuler les ADN glycosylases.

L'observation que la protéine HU est capable de reconnaître avec une haute affinité les cassures simple-brin dans l'ADN a suggéré qu'elle pourrait intervenir dans le système BER en stimulant par exemple le « turnover » des ADN glycosylases. Pour éprouver cette hypothèse, nous avons évalué l'effet de HU sur l'activité de la protéine Fpg, une ADN glycosylase bactérienne impliquée dans l'élimination des purines oxydées telles que la 8-oxoguanine, une lésion extrêmement mutagène.

7.4. Article : La protéine HU est capable de stimuler la réparation de la 8-oxoguanine médiée par la Formamidopyrimidine ADN Glycosylase.

The nucleoid-associated protein HU enhances 8-oxoguanine base excision by the Formamidopyrimidine-DNA glycosylase

Rémy Le Meur^{1,2}, Françoise Culard¹, Virginie Nadan¹, Stéphane Goffinont¹, Franck Coste¹, Martine Guerin¹, Karine Loth¹, Céline Landon¹ and Bertrand Castaing^{1*}

¹Centre de Biophysique Moléculaire (UPR4301), CNRS, rue Charles Sadron, 45 071 Orléans cedex 02, France

²Graduate School of “Santé, Sciences Biologiques et Chimie du Vivant (ED 549), University of Orléans, France

*Corresponding author: Phone: +33 (0)2 38 25 78 43; E-mail: castaing@cnrs-orleans.fr

Key words: nucleoid-associated protein HU, BER, Fpg (MutM), DNA repair stimulation, SSB

Running title: Stimulation of Fpg (MutM) by the histone-like HU protein

Abstract

Among nucleoid-associated proteins, the HU protein is involved in numerous DNA transactions and thus is essential in DNA maintenance and bacterial survival. The high affinity of HU for single-strand breaks (SSB) has suggested its involvement in DNA protection, repair and recombination. DNA-containing SSB are major intermediates transiently generated by bifunctional DNA *N*-glycosylases that initiate the base excision repair pathway (BER). Enzyme kinetics and DNA binding experiments demonstrate that HU enhances the 8-oxoguanine DNA glycosylase activity of Fpg by facilitating the release of the enzyme from its end-product. We propose that the displacement of Fpg from its end-DNA product (one nucleoside gap) by HU is an active mechanism in which HU recognizes the product when it is still bound by Fpg. Through DNA binding, the two proteins interplay to form a transient ternary complex Fpg/DNA/HU which results in the release of Fpg and the molecular entrapment of SSB in a stable complex with HU. These results suggest that HU could play a role in BER *in vivo*.

Introduction

The effectiveness of DNA transactions such as replication, recombination, transcription and repair depends on the structural functionality and plasticity of DNA chromosomes. Thus, the DNA compaction level *in vivo* determines the physical accessibility of the genetic material for molecular effectors in prokaryote and eukaryote cells (1). Like the eukaryote histones, about fifteen small and basic proteins (nucleoid-associated proteins, NAPs) contribute in bacteria to organize the circular chromosome in a functional and dynamic high-order structured nucleoprotein complex, called the nucleoid (2,3). Among these proteins, the small, basic and usually homodimeric (~2x9.5 kDa) HU protein (“H” for histone-like and “U” for the U93 strain used to isolate the *E. coli* nucleoid) is one of the most abundant NAP found in bacteria (over 30 000 molecules per bacterium) (4). HU is also present in archaea (5), plant chloroplasts (6) and can also be encoded by bacteriophage genomes (7). More than just a histone-like protein involved in DNA compaction, HU plays a pleiotropic role in bacteria and through DNA binding (bending and/or wrapping) modulates several DNA-dependent activities (4,8). In enteric bacteria such as *E. coli*, the homologous α and β subunits of HU are encoded by the *hupA* and *hupB* genes, respectively. HU subunits can fold into three dimeric forms (HU α_2 , HU β_2 and HU $\alpha\beta$) whose relative abundance is modulated during growth phases (9). HU α_2 predominates during the exponential phase and HU $\alpha\beta$ accumulates during the stationary phase. HU β_2 is produced in the late exponential phase but does not accumulate and readily recombines with HU α_2 , already present in bacteria, to produce the heterodimer HU $\alpha\beta$. Thus, HU α_2 is rather required in active dividing cells whereas HU $\alpha\beta$ is necessary to protect the genetic material by contributing to the hyper-condensation of the bacterial chromosome in quiescent cells. It is interesting to note that HU β_2 is the most abundant dimer during the adaptive phase following a cold shock (10,11). Clearly, each *E. coli* dimer does not play the same role. Although showing very similar primary and 3D structures (12-16), the three HU dimers display different DNA binding properties (17) probably connected to their different thermodynamic characteristics (14,18).

Among the different roles played by HU, several reports demonstrate its involvement in DNA repair (19-22). Especially, the *E. coli* strains lacking HU are extremely sensitive to γ and UV irradiation known to induce mainly oxidized base lesions and single-strand breaks in DNA. At the molecular level, the ability of HU to bind specifically to a wide variety of repair and recombination intermediates (often containing single-strand breaks) is proposed to be responsible for its role in

DNA repair (17,23-25). Such DNA strand interruptions, usually seen as DNA damage, can also be considered as transient DNA repair products. During the damage base removal of the base excision repair (BER) pathway, single-strand breaks varying by the nature of their ends are generated from DNA glycosylases coupled to AP lyase activity and from AP endonucleases (26). For example, oxidized DNA base lesions induced by reactive oxygen species are mainly removed by DNA glycosylases, the enzymes which initiate the BER pathway (cleavage of the N-glycosidic bond between the damaged base and its associated sugar). Bifunctional DNA glycosylases such as Fpg (Ogg1 bacterial homologue) and Nei (NEIL1, 2 and 3 homologues) have an AP lyase activity consisting in the cleavage of the phosphodiester bond at the 3' (or 3' and 5') side(s) of the abasic (AP) site by β (or $\beta\delta$)-elimination (the AP site resulting from their DNA glycosylase activity) (26,27). Thus, bifunctional DNA glycosylases can generate single-strand breaks with 5'-phosphate (p) and 3'-deoxyribosephosphate (dRp) ends or with 5'-p and 3'-p ends (one nucleoside gap). Interestingly, several studies report that most DNA glycosylases present high affinity for their final products (AP site or single-strand break) which strongly and negatively impacts their turnover *in vitro* (28-35). However, the effectiveness of these enzymes *in vivo* suggests that there are cellular partners capable of removing them from their final reaction products. In previous work, we showed that HU recognizes with high affinity single-strand breaks (nicks or gaps) whatever the nature of their ends and can be considered as the major single-strand break binding protein in crude extract of *E. coli* (23). It was, therefore, of interest to study the effect of HU on the BER efficiency at the molecular level. In this work, we examine the effect of HU on the Formamidopyrimidine-DNA glycosylase (Fpg or MutM), an *E. coli* bifunctional DNA glycosylase involved in the removal of oxidized purines such as the mutagenic 8-oxoguanine (8-oxoG) and the imidazole-ring open purine (FapyG) base lesions and which is strongly entrapped by its final reaction product, a one nucleoside gap (35-37).

Materials and Methods

DNA and proteins

All synthetic oligodeoxyribonucleotides used in this study were purchased from Eurogentec (Fig.S1). Lesion(X)-containing strand (24-mer X) or its complementary strand (24-mer C) were labeled at their 5' terminus with [γ^{32} -P] ATP using polynucleotide kinase prior to annealing with the complementary strand(s) to generate the 24-mer [X:C] DNA duplex. Abasic (AP) site-containing substrate (24-mer [AP:C]) was obtained by incubating for 30 min at 37°C 25 nM of 24-mer [U:C] (5'-radiolabeled on the U-strand) with 0.3 nM of UDG (uracil-DNA glycosylase from New England BioLabs) in the Fpg incubation buffer (see below). Quantitative conversion of U into the AP site was assessed by alkaline treatment (0.2 M NaOH, 50°C for 5 min). Due to its extreme lability, AP site-containing DNA was prepared immediately before use. One nucleoside gap containing DNA (24-mer [Gap:C]) was obtained by annealing the 5'-[32 P]-24-mer C strand with both the 11-mer 3'p and 12-mer 5'p strands (Fig.S1).

The *E. coli* Fpg protein was overproduced as described previously (38) and purified according to the procedure already described for its *Lactococcus lactis* homologue (39). The *E. coli* HU proteins (HU α_2 , HU β_2 and HU $\alpha\beta$) were produced and purified as described by Ramstein *et al.* (14). The open reading frame of the wild-type yOGG1 was isolated by PCR using the plasmid pYSB160 as a template (40) and the oligonucleotide primers AAATGTCTTATAAATTCGG and CTCGAAGCTTCCCTAATCTATTTTTGCTTC. After EcoRI/HindIII digestion, the resulting PCR fragment was inserted by ligation into the linear EcoRI/HindIII pET30a expression vector (Novagen) and yOgg1 was overexpressed in the *E. coli* BL21(DE3) strain. After immobilizing by metal-affinity chromatography (IMAC) using Talon resin (Clontech), the homogeneous His-tagged protein was obtained after AcA54 (IBF-LKB) gel filtration chromatography and HS Poros perfusion chromatography (PerSeptiveBiosystem).

Fpg and Ogg1 assays

In the case of Fpg, the assay consisted in following the removal of 8-oxoG-coupled with the excision of the AP site (8-oxoG-DNA Glycosylase/AP lyase activity). All AP sites resulting from the excision of 8-oxoG by Fpg were converted into one nucleoside gaps (β,δ -elimination). Under the same conditions, yOgg1 essentially removed 8-oxoG from DNA and has a very weak AP lyase activity. AP sites resulting from the excision of 8-oxoG by yOgg1 were cleaved with NaOH which

performed the β,δ -elimination reaction (8-oxoG-DNA Glycosylase activity).

The activity of each enzyme was measured under multiple-turnover conditions by incubating 25 nM of the [5'-³²P]-labeled 24-mer [8-oxoG:C] (labelling on the 8-oxoG-containing strand) with 2.5 nM of Fpg from *E. coli* (3 nM of yOgg1) at 37° in the Fpg incubation buffer (25 mM Hepes pH 7.4, 75 mM NaCl, 1 mM MgCl₂, 100 μ M BSA (~0.66 %), 4.25% glycerol and 1 mM DTT). Incubations were performed with Fpg (or yOgg1) alone or in the presence of HU protein for fixed or varying time intervals up to 10 min. The reactions were stopped by addition of formamide loading dye buffer (final 48 % formamide, 5 mM EDTA). In the case of yOgg1, the reaction was stopped by 5 min incubation with 0.2 M NaOH and with the formamide loading buffer. Substrate and products were separated by electrophoresis on a 20% denaturing (6M Urea) polyacrylamide gel (Urea-PAGE) and quantified with Typhoon Molecular Imager and ImageQuant software. For single-turnover conditions, the same protocol was applied except that the DNA and Fpg concentrations were 2.5 and 25 nM, respectively.

According to the enzyme kinetic scheme proposed for Fpg (yOgg1) (Scheme 1), relevant rate constants in multiple-turnover conditions (k_2 , the rate constant of the catalytic step and k_3 , the rate constant of the product release) were determined with the same method described previously for the *E. coli* MutY DNA glycosylase (41). To determine kinetic parameters, plots of Fpg end-product formation (24-mer [Gap:C]) versus time were analyzed using equation 1

$$[\text{Gap:C}]_t = A_0 [1 - \exp(-k_2 t)] + k_3' t \quad (\text{eq.1})$$

where [Gap:C] is the end-product, A_0 and k_2 are the amplitude and the rate constant of the burst phase respectively, and k_3' the slope of the linear steady step phase. The rate constant of the release of the Fpg product k_3 can be extracted from equation 2

$$k_3 = k_3' / [\text{Fpg}]_0 \quad (\text{eq.2})$$

where $[\text{Fpg}]_0$ is the active Fpg concentration in the assay. Raw data obtained for yOgg1 in multiple-turnover conditions were analyzed with the same approximations.

To examine a putative effect of HU only on the catalytic step mediated by Fpg, kinetic experiments were performed in single-turnover conditions. Under these conditions, raw data can be analyzed using equation 3

$$[\text{Gap:C}]_t = A_0 [1 - \exp(-k_{\text{obs}} t)] \quad (\text{eq.3})$$

where A_0 and k_{obs} are the amplitude and the observed rate constant respectively of the observed single exponential phase. $k_{\text{obs}} \approx k_2$ if we consider that the rate of the Fpg/[8-oxoG:C] (ES complex in scheme 1) formation is not limiting ($k_{-1} \gg k_2$).

All kinetic data were fitted using the KaleidaGraph 4.1 software.

Electromobility shift assays (EMSA)

100 pM of DNA were incubated with various amounts of HU and/or Fpg proteins in the same buffer as for glycosylase assays (expect the amount of glycerol 10%) for 30 min at 4°. The mixtures were loaded onto a 10 % polyacrylamide gel (29:1 acrylamide:bisacrylamide) in TBE buffer (89 mM Tris pH 8.3, 89 mM boric acid, 1 mM EDTA). Electrophoresis runs were conducted at 14 V /cm for 3 h at 8°C. Quantifications were performed using the Typhoon imager and ImageQuant software.

EMSA titration experiments were used to extract the apparent dissociation constant K_D^{app} . For that and after quantification, the binding curves were fitted to a single binding site model using the equation: $Y = [\text{protein}] / ([\text{protein}] + K_D^{app})$ where Y is the fraction of bound DNA and [protein] is the concentration of the free protein that is reasonably approximated by the total protein concentration ($[\text{protein}]_0 \gg [\text{DNA}]_0$).

For kinetic binding experiments, the Fpg/[Gap:C] or HU/[Gap:C] complexes were first preformed at 4°C. Then, 100 nM of cold 24-mer [THF:C] or 25 nM of HU was added. After the indicated time, dissociation was stopped by loading the incubation mixture onto a running non-denaturing polyacrylamide gel (identical to that used in equilibrium EMSA). Dissociation rates (k_{off}) and half-life ($T_{1/2}$) were determined as described elsewhere (42).

Cyanoborohydride-trapped complex preparation (rSB)

During the catalytic process mediated by Fpg, an imino-enzyme DNA intermediate (Schiff base, SB) is transiently formed between the N-terminal proline of the enzyme (P1) and the C1' of the intermediate AP site product. SB can be stabilized by its irreversible reduction with sodium cyanoborohydride (NaBH_3CN) in the reduced Schiff base intermediate (rSB, Fig.6a). The formation of rSB was monitored by SDS-PAGE. Quantitative conversion of SB to rSB was achieved by incubating (2 hours at 25°C) 22 nM of the 5'-[^{32}P]-labeled 24-mer [8-oxoG/C], 500 nM of Fpg and 58 mM NaBH_3CN in the Fpg incubation buffer. The reducer was then eliminated by exclusion chromatography onto a mini Bio-Gel P-6 column (Bio-Rad) equilibrated in the Fpg incubation buffer.

Results and discussion

The three dimeric forms of the HU protein modulate the activity of Fpg

To explore whether the enzymatic efficiency of *E. coli* Fpg can be modulated by HU, we followed the 8-oxoG-DNA glycosylase/AP lyase activity of the enzyme in the presence of each of the three *E. coli* HU dimers. Since the two activities of Fpg are concerted, all abasic (AP) sites resulting from the removal of 8-oxoG (*N*-glycosydic bond cleavage of 8-oxoG, Glycosylase activity) are quantitatively converted into a one nucleoside gap (Gap, Fig.S1) by the cleavage at both the 3' and 5' sides of the AP site according to a β,δ -elimination reaction (AP lyase activity). If the 8-oxoG-containing strand of DNA substrate is 5'-[³²P]-radiolabelled, the final reaction product can be easily separated from the substrate by Urea-PAGE as a radiolabelled DNA fragment (P) shorter than the intact damage strand-containing DNA substrate (S) and visualized and quantified from gel autoradiography. Representative experiments are shown in Figure 1a when incubations were performed with increasing concentrations of HU α_2 or HU β_2 (lanes 2 to 6, *upper* and *lower* panels, respectively, Fig.1a) and a constant amount of Fpg (lanes 1 to 6, Fig.1a). Whereas no DNA cleavage was observed with the HU homodimers alone (lanes 7, Fig.1a) indicating that they cannot process 8-oxoG, a strong increase in the amount of DNA product was generated until the [HU:Fpg] molar ratio reached 50 and up to 100 for HU α_2 and HU β_2 , respectively (lanes 4 and 6, respectively, Fig.1a). Although both HU homodimers are able to significantly stimulate the glycosylase/lyase activity of Fpg, HU α_2 appeared more efficient than HU β_2 (compared for the same [HU:Fpg] ratio, lanes 3 for example, Fig.1a). Quantification of such experiments is shown in Figure 1b for the three HU dimers. Overall, the *E. coli* HU dimers are able to modulate the Fpg glycosylase/lyase activity. The HU concentration-dependence of the Fpg activity displays a similar bell-shape for all dimers. HU α_2 and HU $\alpha\beta$ display a maximum of an 8-fold increase in activity around a 50 molar excess of HU. Higher concentrations of HU β_2 are needed to observe an Fpg stimulation similar to that of the other HU dimers. The maximal stimulation effect (6-fold increase in activity) is observed at around a 100 molar excess of HU β_2 (Fig.1b). This singular behavior of HU β_2 as compared to the other *E. coli* HU dimers can be due to its lower DNA binding affinity whatever the structural nature of DNA (linear, branched or containing single-strand breaks) (14,17). However, HU β_2 preferentially recognizes single-strand breaks with lower affinity than HU α_2 and HU $\alpha\beta$. This can explain the shift in the concentration range needed to observe the same effect. 10 times more HU β_2 is needed to

observe stimulation similar to that induced by HU α_2 (compare lane 2 for HU α_2 and lane 4 for HU β_2 , Fig.1a). This result is in good agreement with the observation that HU α_2 (HU $\alpha\beta$) binds single-stranded breaks with an affinity 10 times greater than that of HU β_2 . Beyond the [HU:Fpg] ratio corresponding to the maximal stimulation (50 or 100 depending on the dimer considered, Fig.1b), a decrease in the enzyme activity was progressively observed until a total inhibition over a [HU:Fpg] ratio of about 1500 ([HU:DNA = 150]).

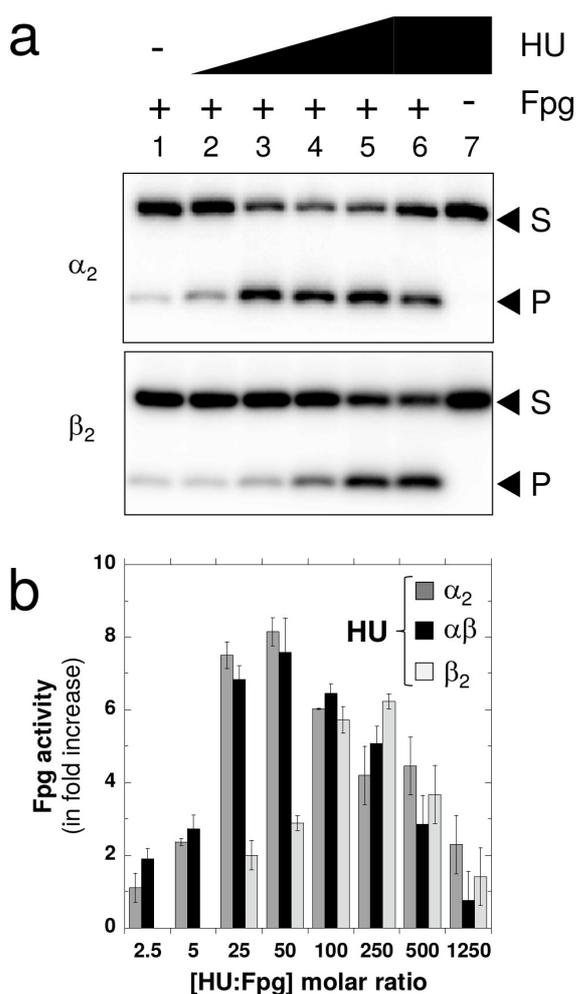


Figure 1: Modulation of the 8-oxoG DNA glycosylase activity of Fpg by the HU proteins

25nM of a radiolabeled 24-mer-[8-oxoG:C] duplex (Fig.S1) was incubated for 5 min at 37°C with HU alone (lanes 7) or with a limiting amount of Fpg (2.5 nM) (lanes 1) in the presence of a molar excess of HU:Fpg of 5, 25, 50, 100, 250 (lanes 2 to 6, respectively). Reactions were stopped by the addition of the formamide loading buffer and the reaction product was separated from the substrate by denaturing 20% Urea-PAGE as described in *Materials and Methods*. (a) Representative gel autoradiography with HU α_2 and HU β_2 as indicated (S) and (P) are for Fpg substrate and product, respectively. (b) Quantifications of autoradiography obtained with each *E. coli* HU dimer. The Fpg activity stimulation factor (fold increase) corresponds to the ratio between the relative quantity of product generated by Fpg in the presence of HU and that without HU. The observed fold increase is reported as a function of the indicated [HU:Fpg] molar ratio. The values correspond to the average (\pm SD) of four independent experiments.

Indubitably, the effect of HU is connected with its ability to bind DNA and can result from at least two opposite molecular mechanisms: (i) one operating at low HU concentration and resulting in the stimulation of the enzyme and (ii) a second one occurring at higher HU concentration leading to a loss of the stimulation ranging to an apparent inhibition of the enzyme (under a large molar excess of HU as compared to DNA). Such opposite processes (stimulation-inhibition) depending on the

concentration have already been observed for the stimulation of topoisomerase activity by HU from *Mycobacterium tuberculosis* (43). A fine analysis of the HU effect indicated that the stimulation of Fpg at low HU α_2 and HU $\alpha\beta$ concentration ($[\text{HU:Fpg}] < 25$ and $[\text{HU/DNA}] < 2.5$) was linear as a function of HU concentration (*circles* and *squares*, Fig.S2). This suggested that at low HU concentration only the stimulation mechanism occurred (a similar observation can be made for $25 < [\text{HU}\beta_2:\text{Fpg}] \text{ ratio} < 100$; *triangles*, Fig.S2). Over the $[\text{HU:Fpg}]$ ratio of 20 (HU α_2 and HU $\alpha\beta$), the linearity of the stimulation was progressively lost until the observed maximal stimulation effect, suggesting that the HU inhibition mechanism starts operating (Fig.1b). At high concentration, Fpg inactivation could result from the entrapment of either the enzyme in an unproductive Fpg/HU complex or the DNA substrate in a HU/DNA complex in which the peculiar DNA binding mode of HU hides the damaged base from the enzyme. Despite our attempts, we were unable to demonstrate the formation of a stable complex between Fpg and HU. The HU concentration range shift between HU α_2 (HU $\alpha\beta$) and HU β_2 indicates that both mechanisms (stimulation and inhibition) relied on the relative affinity of HU isoforms for DNA. These observations strongly suggest that HU modulates Fpg activity through DNA binding. According to this assumption, the transition from one to the other modulation mechanism could result in a DNA binding mode transition of HU.

The precise HU binding mode could depend on the $[\text{HU:DNA}]$ molar ratio which is partly connected to the precise stoichiometry of the HU/DNA complexes formed (44). As a function of the HU binding mode (depending on HU origin, local HU concentration, salt concentration, incubation time, structural nature of the DNA probe...), the size of the HU binding site can fluctuate from 6 to 42 base pairs/dimer (44-47). It can be suggested that at high concentration, HU molecules stack and wrap around DNA as already proposed from biochemical, microscopy and crystal structure studies of HU bound to intact and linear DNA (respecting one dimer/ \sim 6-10 base pairs, “*small site size binding mode*”) (12,44). Probably, this results in sequestering the Fpg DNA-substrate in a non-accessible form for the enzyme. A similar situation has already been described for the human 8-oxoG-DNA glycosylase, hOgg1 (the human Fpg functional homolog): the removal of 8-oxoG by hOgg1 in positioned nucleosomal and internucleosomal structures is highly impeded and requires the assistance of chromatin remodeling factors (48-50). On the other hand, at low concentration, the ability of HU to stimulate Fpg is probably connected with their ability to preferentially bind single-strand breaks, according to the high affinity binding mode (“*large site size binding mode*”). It was previously proposed that HU bound with high affinity single stranded-breaks by inducing a strong bent in DNA (named U-turn DNA) characterized by two $80\sim 90^\circ$ kinks separated by 9-10

nucleotides similar to the one observed in the crystal structures of HU/IHF bound to DNA (15,16,23,25). Through this high affinity binding mode, the size of the binding site of HU is larger and can oscillate from 17 to 42 base pairs depending on the HU molecular species considered, the HU:DNA ratio and the nature of DNA (size and the presence or not of special structures such as single-stranded loops, branched, bulged or nicked structures). Both the flexible arms and the core of the dimer contribute to the interaction with DNA and stabilize the HU/DNA complex (50). At low HU concentration, the precise molecular mechanism by which HU stimulates Fpg remains to be deciphered.

HU stimulates Fpg turnover by stimulating release of the enzyme from its final DNA product

The stimulation of Fpg by the HU protein can have several origins: through binding, HU may (i) modify the intrinsic structure of the substrate, resulting for example in a better binding of the enzyme to its DNA substrate (stabilization of the enzyme/substrate complex), (ii) stimulate the catalytic step (biochemical conversion of the substrate into product and/or formation of a complex between the two proteins that is more efficient than the enzyme alone) or (iii) promote the recycling of the enzyme after each catalytic act. To decipher the stimulation mechanism at the molecular level, we investigated the effect of HU on the kinetic of the enzyme excising 8-oxoG in DNA (glycosylase) and generating one nucleoside gap as final DNA product from the resulting AP site (AP lyase). Because DNA repair operates essentially in dividing cells, all the experiments described in the rest of this work were performed only with HU α_2 (called “HU” for convenience), the *E. coli* HU form which predominates during the growth exponential phase (9).

Under multiple-turnover conditions ($[Fpg] \ll [DNA\text{-}substrate]$), Fpg exhibits biphasic kinetics with a rapid burst phase (pre-steady state kinetics phase) followed by a slower steady state phase (*black circles*, Fig.2a). The simplest kinetic scheme to analyze the enzyme kinetic behavior of Fpg is as follows:



with E, S and P for Fpg, substrate (24-mer 8-oxoG-containing DNA duplex, [8-oxoG:C], Fig.S1) and final product (one nucleotide gap, [Gap:C], Fig.S1) and where the rate constant of the burst

phase is proportional to the rate constant of the catalytic step (k_2) and the rate constant associated with the linear steady state is dominated by the rate constant of the product release step (k_3). Kinetic data were fitted with eq.1 as described in *Materials and Methods*. Under the conditions used, the product release is the rate-limiting step in Fpg kinetics.

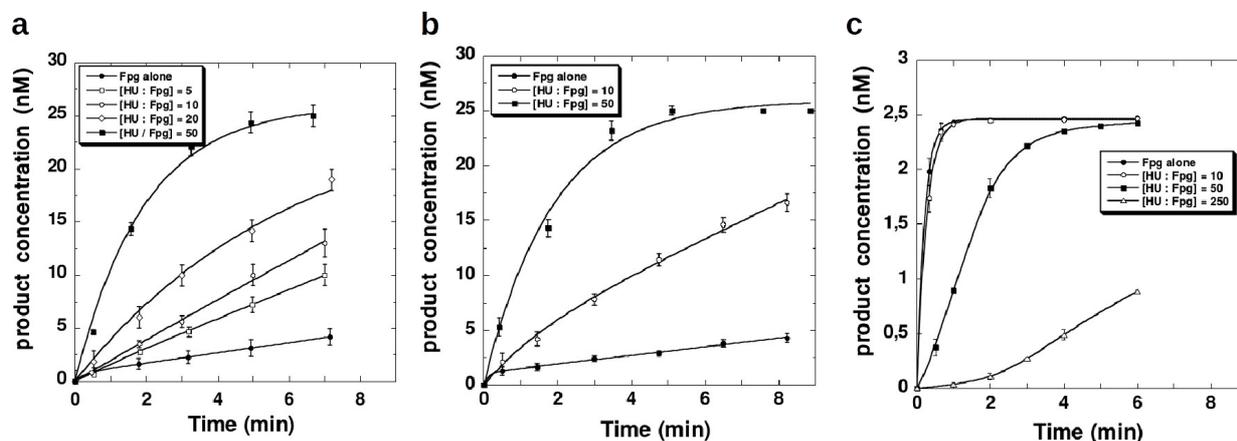


Figure 2: Enzyme kinetics of the excision of 8-oxoG and AP site cleavage by Fpg in the presence of HU

(a) and (b) Plot of product formation as a function of time under multiple-turnover conditions 2.5 nM of Fpg was incubated at 37° with 25 nM of radiolabeled 24-mer [8-oxoG:C] (a) or 24-mer [AP:C] (b) alone or with the indicated ([HU:Fpg] molar ratio). (c) Plot of product formation as a function of time under single-turnover conditions for the concerted 8-oxoG-glycosylase/AP lyase activity 25 nM of Fpg was incubated at 37°C with 2.5 nM of radiolabeled 24-mer [8-oxoG:C] DNA duplex with or without HU as indicated.

For the three types of experiments (a, b and c), samples of reaction mixtures were analyzed as a function of time by Urea-PAGE. After gel autoradiography, autoradiograms were quantified as described in the *Materials and Methods* section. For each sample, the reaction product of Fpg (in all cases, a DNA duplex containing a one nucleoside gap, 24-mer [Gap:C], Fig1a) is plotted as a function of the incubation time. The indicated kinetic points correspond to the mean values (\pm SD) of four independent experiments. Experimental points were fitted using eq.1 or eq.2 as described in the text and the *Materials and Methods* section.

The effect of increasing the concentration of HU on the enzyme kinetics was then investigated. The presence of HU in a [HU:Fpg] molar ratio of 5 and 10 induced a remarkable increase in the steady state rate whereas the initial burst velocity seems relatively poorly affected (*open* squares and circles, respectively, Fig.2a). For a [HU:Fpg] ratio of 20-50, the release of the product is no longer rate-limiting and the kinetic plot of the product formation displays a monophasic shape and experimental data can be fitted by a single exponential curve as described in *Materials and Methods* (eq.3) (*black* squares, Fig.2a). The substrate is quantitatively converted into the final product after 8 min for a [HU:Fpg] ratio of 50, while only 20% of product is generated after the same time period without HU. The constant rate k_3 is increased by about 4-fold with 10 molar

excess of HU (table 1). These data suggest first that Fpg keeps a high affinity for its final DNA product and second that HU is able to displace Fpg from the one nucleotide gap in DNA, thus facilitating its molecular recycling (stimulation of the turnover). If this proposition is true, we should expect that a similar stimulation also exists for the uncoupled AP lyase activity of Fpg (direct conversion of preformed AP site-containing DNA into a one nucleotide gap). Similarly to the concerted 8-oxoG-DNA glycosylase/AP lyase reaction, the time course of the conversion of the AP site into a one nucleoside gap by Fpg displays a biphasic shape (*black circles*, Fig.2b). As expected, this result suggests that the turnover of the Fpg AP lyase activity is also strongly limited by the dissociation rate of the [enzyme/final product] complex (EP, scheme 1) in free enzyme and product (E+P, scheme 1). In the presence of a 10 molar excess of HU, the linear steady state rate corresponding to the product release (k_3) is significantly increased (*open circles*, Fig.2b; table 1) without apparently impacting the preceding steps (pre-steady steps). Less than 5 nM of product (about 20%) was produced in 8 min without HU (*black circles*, Fig.2b) whereas the substrate was quantitatively converted into the end-product with a 50 molar excess of HU during the same time period (*black squares*, Fig.2b). Kinetics can be fitted by eq.3 indicating that at the [HU:Fpg] ratio of 50, the product release step is no longer limiting. Once again, HU seems to stimulate the Fpg AP lyase activity by facilitating the final product release and thus the recycling of the enzyme after each catalytic act.

Due to the high velocity of the pre-steady steps (the rapid burst phase), it was difficult to assess whether HU can also modulate the kinetic steps preceding the product release under multiple-turnover conditions. To explore that, the 8-oxoG-DNA glycosylase/AP lyase activity of Fpg was investigated under single-turnover conditions ($[Fpg] \gg [DNA]$ substrate) in the presence of various HU concentrations. The time course of the product formation can be fitted by a single exponential curve using eq.3 (*black circles*, Fig.2c). Under these conditions and in the presence of HU, the observed rate constant remains unchanged up to a [HU:Fpg] molar ratio of 10 (*open circles*, Fig.2c) suggesting that HU minimally affects both the recognition of the substrate (8-oxoG or AP site) in DNA (k_1 and k_{-1} , scheme 1), and the intrinsic chemical process leading to the damage removal (k_2 , scheme 1). In this HU concentration range, the substrate is quantitatively converted to product with the observed rate constants (k_{obs}) of $4.9 \pm 0.2 \text{ min}^{-1}$. The combined results suggest that up to a [HU:Fpg] ratio of 10, the stimulation effect observed under multiple-turnover conditions is limited to the steps governing the steady-state rate, i.e. the rate of the product release (k_3 , scheme 1). Beyond a [HU:Fpg] ratio of 50, Fpg is progressively less efficient, indicating an inhibitor effect of

HU at high concentration (*black squares and open triangles, Fig.2c*). This phenomenon can result either from a direct inactivation of the enzyme (Fpg/HU inactive complex) or from the hijack of the DNA substrate by HU (abortive HU/DNA substrate) as suggested below. Assuming that HU is unable to interact physically with Fpg (all our attempts to demonstrate the possible formation of a stable [Fpg/HU] without DNA were unsuccessful), the simplest hypothesis to interpret this observation is to consider that, through non-specific DNA binding (“*small site size binding mode*”), a large quantity of HU progressively entraps the Fpg DNA substrate in a non Fpg processing form. The apparent stimulation of the enzyme observed in multiple-turnover up to the [HU:Fpg] ratio of 50 (Fig.1 and Fig2a and 2b) results from two opposite mechanisms: a major positive effect on the product release and a minor negative effect on the enzyme-substrate complex formation (k_1 , scheme 1). Beyond a [HU:Fpg] ratio of 50, the negative effect of HU on the enzyme substrate binding progressively predominates until a total apparent inactivation due to the sequestration of the intact substrate (as seen in Fig.1b and Fig.2c).

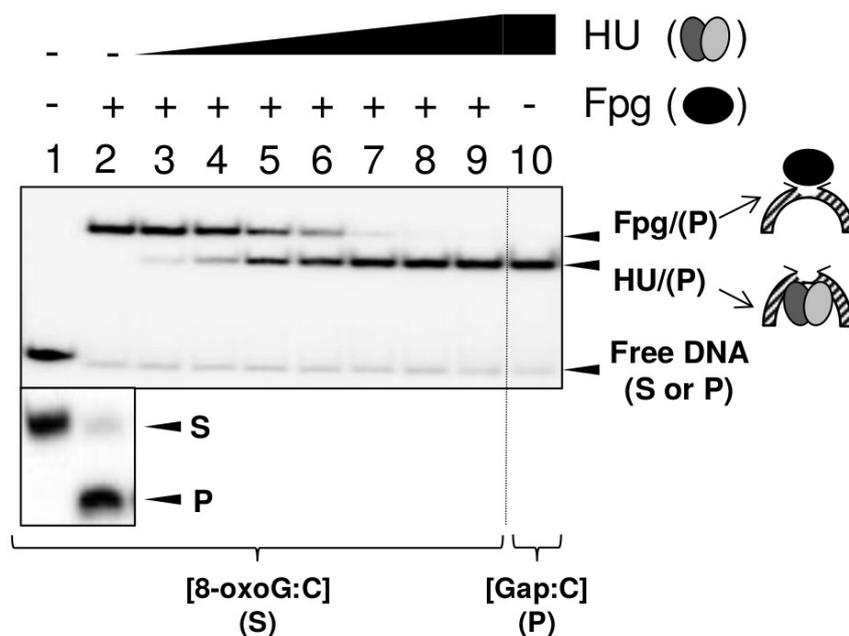


Figure 3: Effect of HU on Fpg/DNA-product complex

0.1 nM of 24-mer [8-oxoG:C] was pre-incubated 5 hours at 8°C alone (lane 1) or with 2.2 nM of Fpg (lane 2-9). After pre-incubation, 0, 0.08, 0.13, 0.25, 0.5, 1, 2 and 4 nM of HU was added (lanes 2-9, respectively). After 20 min additional incubation, incubation mixtures were analyzed by EMSA (*Upper autoradiogram*) or by Urea-PAGE (*Lower autoradiogram*) as described in the *Materials and Methods* section. The analysis of 24-mer [8-oxoG:C] by Urea-PAGE with and without Fpg was to show that under the incubation conditions used, the Fpg substrate [8-oxoG:C] was quantitatively converted into the Fpg final product [Gap:C]. The electromobility of the HU/[Gap:C] was identified in native polyacrylamide gel by analyzing the incubation mixture (20 min at 8°C) of 0.1 nM of 24-mer [Gap:C] with 4 nM of HU (lane 10).

HU selectively displaces Fpg from its final DNA product

The clear enhancement of the apparent rate of the product release (k_3 , scheme 1) at low HU concentrations suggested that HU may recognize the complex Fpg/DNA-product and displace Fpg from its end-DNA product (active process). Alternatively, considering the high affinity of HU for DNA containing single-strand breaks (14,17,23), both HU and Fpg may compete for the Fpg reaction product (passive process).

To decide between these two processes, we first analyzed the effect of HU on the preformed Fpg/end-product by using the Electro-Mobility Shift Assay (EMSA) which resolves the protein/DNA complex from the free DNA probe by electrophoresis in non-denaturing polyacrylamide gel (Fig.3). The pre-incubation of the [8-oxoG:C] duplex (lane 1, Fig.3) with 2.2 nM of Fpg results in the quantitative conversion of the substrate into product (Urea-PAGE of lower panel, Fig.3) that leads to the formation of a stable Fpg/DNA-product complex (Fpg/(P), lane 2, upper panel, Fig.3). This stable complex migrates more slowly than the complex HU/DNA-product (HU/(P), lane 10, Fig.3). The electro-mobility of each complex is consistent with the larger molecular weight of Fpg (~30 kDa) relative to that of HU (~20 kDa). It was therefore easy to distinguish between the two protein/DNA complexes by this method. The effect of HU on the preformed Fpg/DNA-product complex was investigated by adding increasing amounts of HU (lanes 3 to 9, Fig.3). The addition of HU resulted in the lower-shift of the bound DNA which corresponds to the conversion of the Fpg/DNA-product complex into the HU/DNA-product complex (Fpg/(P) and HU/(P) complex, respectively, Fig.3). A [HU:Fpg] molar ratio of less than 0.5 is sufficient to displace more than 50% of the Fpg/(P) complex (lane 7, Fig.3). For a two molar equivalent of HU, the preformed Fpg/(P) complex is quantitatively converted into the HU/(P) complex (lane 9, Fig.3). Unambiguously, HU was able to displace Fpg from its DNA product (P, also called [Gap:C]).

The specificity of this displacement was then investigated by the same method using as DNA probe a 24-mer [THF:C] DNA duplex containing a tetrahydrofuran (THF) site, an AP site substrate analogue (Fig.S1). THF-containing DNA is not processed by Fpg but constitutes a high affinity ligand for the enzyme. Through Fpg binding, [THF:C] behaves as a strong inhibitor of both the Glycosylase and AP lyase activities of the enzyme (38,51). As for the Fpg DNA-product (lane 10, Fig.S3), 2.2 nM of Fpg was sufficient to quantitatively titrate the [THF:C] duplex into a stable complex (lane 2, Fig.S3). However, a two molar excess of HU was unable to displace the preformed Fpg/[THF:C] complex (lane 3, Fig.S3) whereas for the same excess, the preformed Fpg/[Gap:C]

was quantitatively converted into HU/[Gap:C] (lane 9, Fig.3). Beyond an HU molar excess of 16 (lane 6, Fig.S3), the preformed complex of Fpg/[THF:C] begins to be disturbed to the benefit of free DNA (and of a small quantity of HU/DNA complex and of a supershifted molecular species compatible with a ternary Fpg/[THF:C]/HU complex; *open* and grey arrows, respectively, lanes 6-8, Fig.S3). Clearly, HU has low affinity for the free and Fpg-bound DNA duplex [THF:C] and the displacement is effective only on the Fpg DNA-product. To gain further insight into this molecular process, we investigated the effect of Fpg on the preformed HU/[Gap:C] complex as compared to that of HU on the preformed Fpg/[Gap:C] complex (Fig.4). Whereas the displacement of the Fpg from its DNA product was quantitative for a two molar excess of HU (*black* spheres, Fig.4), less than 5% of the preformed HU/[Gap:C] complex was displaced by the same molar excess of Fpg (*open* squares, Fig.4). Although Fpg and HU display a similar affinity for [Gap:C] (Fig.S4), the efficiency of the molecular displacement is not reciprocal and Fpg loses its high affinity for its product bound to HU. HU acts as an Fpg-product trap preventing the enzyme from rebinding it.

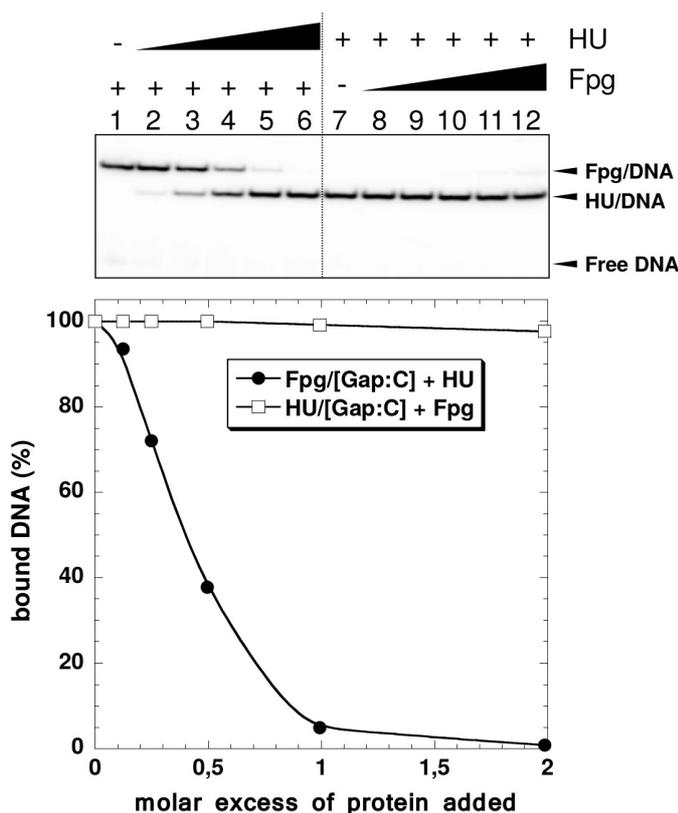


Figure 4: Effect of Fpg on the preformed complexes HU/[Gap:C]

0.1 nM of the 24-mer [Gap:C] was pre-incubated for 30 min at 8°C with 2.2 nM of Fpg (lanes 1 to 6) or HU (lanes 7 to 12). 0, 0.08, 0.13, 0.25, 1.2 and 12.4 nM of the competitive protein were then added (HU in lanes 1 to 6, respectively and Fpg in lanes 7 to 8, respectively) and the resulting mixtures were incubated for 10 additional min at 8°C. Each mixture was analyzed by EMSA as described in Fig.3. The autoradiography obtained (upper panel) and its quantification (lower panel) are presented.

An active molecular process for the displacement of Fpg by HU was investigated by analyzing the relative affinity of HU for the free and Fpg-bound product using as DNA probe the Fpg final DNA product, the 24-mer [Gap:C] DNA duplex (Fig.S5). In such an experiment, the Fpg concentration

was chosen to titrate 50% of the free product (lane 1, Fig.S5a). Separately, aliquots of this equilibrium mixture were then incubated with increasing concentrations of HU and analyzed by EMSA (lanes 2 to 6, Fig.S5a). As anticipated by previous experiments, both free and bound DNA probes were converted to a stable HU/DNA-product complex (i.e., HU/[Gap:C]). HU progressively and indifferently titrated the free and bound [Gap:C] DNA duplex as a unique DNA species as if Fpg was not present in the mixture (Fig.S5b). This unexpected result does not make it possible to discriminate between a passive and an active mechanism but it again shows the efficiency of the displacement of Fpg by HU. However, the contribution of each mechanism to the stimulation of the Fpg turnover and the possibility of forming a transient ternary complex during this process remain to be explored.

Evidence and structural determinants for the formation of a transient ternary complex during the active displacement of Fpg from its DNA end-product by HU

The active displacement mechanism implies that HU is able to recognize the end-product still engaged in the Fpg/[Gap:C] complex (formation of a transient ternary complex too unstable to be visible by EMSA). The transient nature of this ternary complex is necessary for efficient Fpg turnover stimulation. To further explore this issue, a comparative analysis of the lifetime of the Fpg/[Gap:C] complex in the presence of HU or DNA competitor was investigated (Fig.5). The dissociation of the preformed Fpg/[Gap:C] was followed by EMSA as a function of time in the presence of a fixed concentration of competitor (HU or cold DNA) chosen so that it does not affect the observed dissociation velocity. Under these conditions, HU induces the instant dissociation of Fpg/[Gap:C] and the concomitant formation of HU/[Gap:C] (lifetime < 30 sec, too fast to be determined by EMSA, Fig.5a). The dissociation rate associated with the addition of cold DNA competitor is much slower (with a half-life of 26 min and a lifetime > 2 hours, *black circles* Fig.5c). Clearly, the displacement of Fpg from its end-DNA product by HU is not a passive process as with DNA competitor but an active mechanism. This result constituted indirect evidence for the formation of a transient ternary complex Fpg/DNA/HU. Although the affinities of Fpg and HU for [Gap:C] are very similar (Fig.S4), the longer half-life (more than 4 times) of the HU/[Gap:C] ($T_{1/2} = 115$ min) as compared to that of Fpg/[Gap:C] ($T_{1/2} = 26$ min) (*black diamonds*, Fig.5c) and the affinity loss of Fpg for [Gap:C] when HU is bound to [Gap:C] (Fig.4) prevent Fpg from re-binding to its end-product. Altogether these parameters may contribute to the efficient Fpg stimulation by HU.

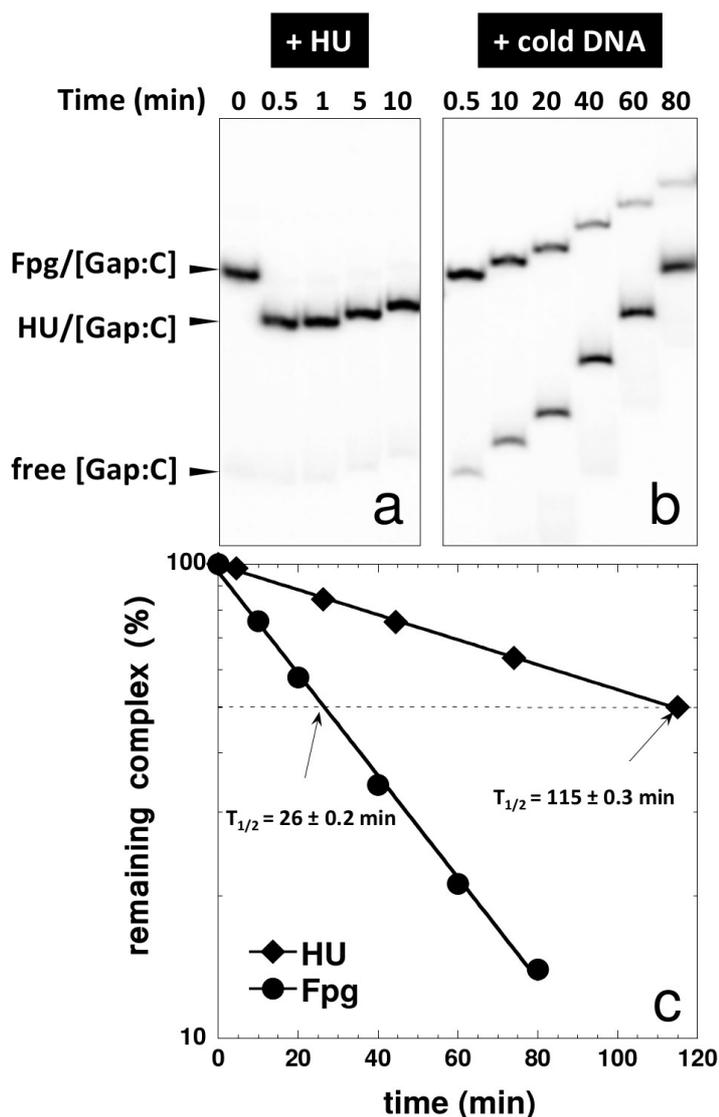


Figure 5: Lifetime of protein/[Gap:C] complexes in different conditions

The preformed Fpg/[Gap:C] complex was obtained by incubating for 30 min at 8°C 1 nM of Fpg with 0.1 nM of 24-mer [32 P]-[Gap:C] DNA duplex. The resulting equilibrium mixture was then incubated either with 25 nM of HU (+ HU) or 100 nM of cold 24-mer [THF:C] DNA duplex (+ cold DNA). Aliquots of each mixture were then analyzed as a function of time by EMSA. Representative gel autoradiographies of the experiments are shown in panels (a) and (b). The time course of the dissociation of HU/[Gap:C] was also performed (gel autoradiography is not shown). (c) Representative plots for the dissociation kinetics of the Fpg/[Gap:C] and HU/[Gap:C] complexes in the presence of cold DNA are shown (circles and diamonds, respectively). After autoradiography quantification, the remaining protein/DNA complex was plotted as a function of time. Indicated half-life times (time period required for observing 50% of complex, $T_{1/2}$) correspond to the mean values obtained from three independent experiments.

The transient nature of the ternary complex Fpg/DNA/HU does not permit its direct experimental observation. To counteract this difficulty, we entrapped Fpg on its intermediate AP site product by reducing (with NaBH_3CN) the transient covalent imino-enzyme-DNA intermediate (Schiff base intermediate, SB) in a stable protein-DNA complex (reduced Schiff base, rSB or covalent Fpg-[AP:C], Fig.6a) (52). The irreversible entrapment of Fpg with AP site-containing DNA prevents HU-induced Fpg DNA dissociation. The incubation of rSB in the presence of increasing HU concentrations results in the formation of a stable ternary Fpg-DNA/HU complex identified by EMSA as a discrete super-shifted band (rSB/HU, lanes 2 to 5, Fig.6b). Unlike the non-covalent Fpg/[THF:C] which was displaced by HU regenerating free DNA (Fig.S3), the covalent Fpg-[AP:C] was titrated by HU in a ternary complex. Although strongly suggested by the lifetimes of the different

complexes (Fig.5), this experiment constituted the first evidence for the ability of HU to physically recognize the preformed Fpg/DNA complex.

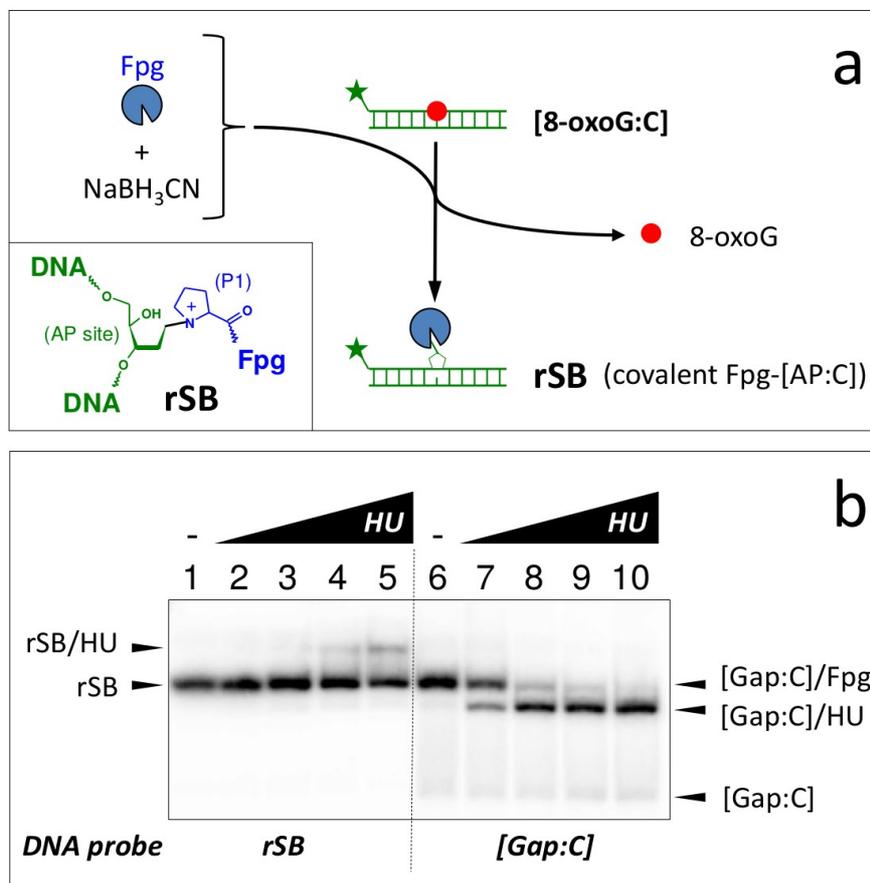


Figure 6: Titration of the borohydride-trapped Fpg-[AP:C] covalent complex by HU in a ternary Fpg-DNA/HU complex

(a) Schematic representation for the preparation of the borohydride-trapped complex. The transient imino-enzyme DNA intermediate concomitantly formed during the excision of 8-oxoG and the abasic (AP) site formation (Schiff base intermediate, SB = double-link established between the C1' of the resulting AP site and the amino group of the N-terminal proline, P1) was irreversibly stabilized by a strong reducer such as sodium cyanoborohydride in reduced Schiff base (rSB, see insert for its chemical nature). The biochemical procedure for the preparation of rSB is described in the *Materials & Methods* section. **(b)** Comparative titration of rSB and [Gap:C] by HU. 3 nM of rSB (lane 1) or the preformed Fpg/[Gap:C] complex (lane 6, obtained with 74 nM of Fpg) was incubated for 20 min at 8°C with 7.6, 38, 190 and 951 nM of HU (lanes 2-7, 3-8, 4-9 and 5-10, respectively). Equilibrium mixtures were then analyzed by EMSA as described in the caption of Fig. 3.

Since the covalent Fpg-DNA complex (rSB) is different from the non-covalent complex Fpg/DNA-end product, the possibility of forming a ternary complex depends not only on the precise chemical nature of the Fpg end-DNA product but also on the global structure of the DNA in the Fpg/DNA

complex. In all Fpg/DNA complexes studied to date (Fpg bound to 8-oxoG-, FapyG-, AP site-, end-product-containing DNA or borohydride-trapped complex with AP site, *i.e.* rSB), crystal structures reveal that the overall structure of DNA in these complexes appears strongly and similarly bent by the enzyme (about 60-70° of the DNA double-helix axis) (53-55). In all cases, the Fpg-induced DNA curvature is centered on the DNA modified site. Thus, the Fpg-induced DNA bent could be one of the features recognized by HU in the Fpg/DNA complex. However, the ability of HU to displace Fpg from its initial substrate ([8-oxoG:C]) or intermediate product ([AP:C]) is not kinetically efficient. During catalysis, the lifetime of these transient Fpg/DNA complexes is too short. If the lifetime of the reaction intermediate is artificially increased by using a substrate analogue such as the AP site analogue THF, the abortive complex Fpg/[THF:C] can be disturbed by HU which recognizes the Fpg-induced bent DNA in the complex even if HU has no preferential affinity for the free AP site in DNA (23). In this precise situation, HU however cannot prevent Fpg from re-associating with THF-containing DNA as in the case of [Gap:C] and, consequently, the displacement of Fpg from THF-DNA by HU is apparently less efficient (compare Fig.3 and Fig.S3). On the other hand, when the enzyme is bound to its end-product (the limiting step being the release of the enzyme from the one-nucleoside gap), the lifetime of this complex is long enough to allow HU recognition of the Fpg-induced bent DNA structure. In addition, the presence of the single-strand break could also increase the intrinsic DNA flexibility inside the Fpg/[Gap:C] complex, thereby facilitating its recognition by HU.

To assess whether the recognition of the Fpg/DNA complex depends on the presence of a flexible curvature or bent in the bound DNA, we investigated the effect of HU on the yeast 8-oxoG-DNA glycosylase (Ogg1) (56). Ogg1 is the 8-oxoguanine-DNA glycosylase in eukaryotes (26). Through binding, Fpg and Ogg1 proteins are known to similarly bend DNA (37,53,57). The excision of 8-oxoG from DNA by Ogg1 in the presence of HU was significantly stimulated (Fig.S6a). Under the same experimental conditions, a molar ratio of [HU:DNA glycosylase] of 50 induced a 9 and 1.5 fold increase in Fpg and Ogg1 activity, respectively (Fig.1 and Fig.S6a). As suggested above, the stimulation of both Fpg and Ogg1 by HU is incompatible with a direct recognition between the DNA glycosylase and HU. Although considered as functional homologues, the two proteins are not structurally related which excludes the HU recognition of common structural determinants present in both DNA glycosylases. Because the AP lyase activity of Ogg1 is very low (Ogg1 can be considered as monofunctional DNA glycosylase, (58)), the stimulation effect was observed only on the glycosylase activity of Ogg1 (the end-product in this case being the

AP site). Kinetic experiments under multiple turnover conditions indicated that the HU stimulation essentially relies on the ability of HU to stimulate the rate of the release of Ogg1 from the AP site, the limiting kinetic step in this case (Fig.S7). Here again, HU stimulates the turnover of the enzyme by facilitating its dissociation from its end-product. This observation is strongly consistent with the HU recognition of the Ogg1-induced AP-DNA bent. The nature of the DNA bent and/or its flexibility (AP site or single-strand break) in the glycosylase/end-product complex can however modulate the observed stimulation effect of HU. Fpg is more efficiently stimulated by HU than Ogg1, probably because AP site-containing DNA is not a high affinity ligand for HU as compared to a one nucleoside gap (Fig.S4).

Proposed model for HU-induced Fpg stimulation

The stimulation of the Fpg turnover by HU is an active process of the enzyme recycling which requires (i) recognition of the Fpg/[Gap:C] complex by HU resulting in the formation of a transient ternary complex Fpg/[Gap:C]/HU in which Fpg and HU do not directly interact (only the DNA connects the two proteins), (ii) the release of Fpg from its end-product (displacement induced by the HU binding to the Fpg/DNA end-product) leading to the formation of a stable [Gap:C]/HU complex and (iii) the inability of the released Fpg to efficiently rebind the DNA product sequestered by HU (non-reciprocity in the displacement). This molecular mechanism is illustrated by the cartoon in Figure 7. In such a mechanism, the Fpg DNA product is never free of protein interactions and appears continuously bound by Fpg and/or HU (channeling of the repair intermediate, as already proposed for the processing of AP sites, (59,60)). This mechanism is significantly different from that of the AP lyase activity stimulation of the DNA glycosylase Nth from *Bacillus subtilis* by HBSu, the HU homolog in this organism (61). In the latter case, the authors proposed that HU-induced DNA superhelicity changes stimulate the binding of Nth to its AP site substrate.

In our case, kinetic data suggest no effect of HU (at low concentration) on either the substrate recognition by Fpg or on the catalytic steps and were consistent with the formation of a transient ternary complex in which both proteins are simultaneously bound to single-stranded break-containing DNA. Crystal structures of Fpg and HU bound to DNA could support such a model. Indeed, Fpg binds essentially the damage-containing strand (substrate or product) on the convex side of the DNA curvature whereas most interactions of HU are established between the phosphodiester backbone on the concave side of the DNA curvature and the saddle-shape basic area of the dimer from which the HU arms emerged (1L2B and 1IHF, respectively: Fig.7) (13,15,16).

Consequently, HU can approach the Fpg/[Gap:C] complex by interacting with the opposite DNA surface to that contacted by Fpg. According to this scenario, the physical interaction between the two proteins is not direct but is transiently mediated by DNA (Fig.7). By using the crystal structure of IHF bound to nicked DNA (as a model for HU, IHF and HU displaying an identical global fold and a similar DNA binding mode, Fig.S8 & S9) (15) and the crystal structure of Fpg from *G. stearothermophilus* (GstFpg) bound to a one nucleoside gap (GstFpg displays 41% sequence identity with its *E. coli* homolog and a similar DNA binding mode, Fig.S9) (54,57), it was possible to visualize a realistic 3D model for the ternary complex (Fig.S10). By aligning as well as possible both structures at the single-strand break, some steric clashes (red box, Fig.S10) occur in the vicinity of the break between β -hairpin loops of the N-terminal domain (NTD) of Fpg (β 4- β 5 and β 7- β 8 loops) and the β -hairpin loop of one of the two arms of HU. This indicated that the fine interactions to the nicked site are unique and exclusive to each protein. According to this model, HU and Fpg could weakly interact by these dynamic loops (negative data to show that do not constitute conclusive proof). Consequently, to destabilize the Fpg/Gap-DNA complex, HU has to first break up tight interactions established by Fpg with the nicked strand (Fig.S10c & d). Thus, the HU-induced displacement of Fpg from its product can be achieved by bending DNA more strongly than Fpg (DNA is 90° kinked by HU whereas Fpg induced a bent of less than 70°). This strong curvature induced by HU relies on the two intercalated P64 from the arms (two kinks of 90° separated by 9-10 base pairs, Fig.S10d) and additional interactions of the dimer body with DNA (Fig.7). The strong curvature induced by HU could be unfavorable to the symmetrical DNA displacement of HU by Fpg as we observed (Fig.4). This suggests that the HU / Fpg end-product complex could be a physiological target for BER enzymes involved downstream of DNA glycosylases.

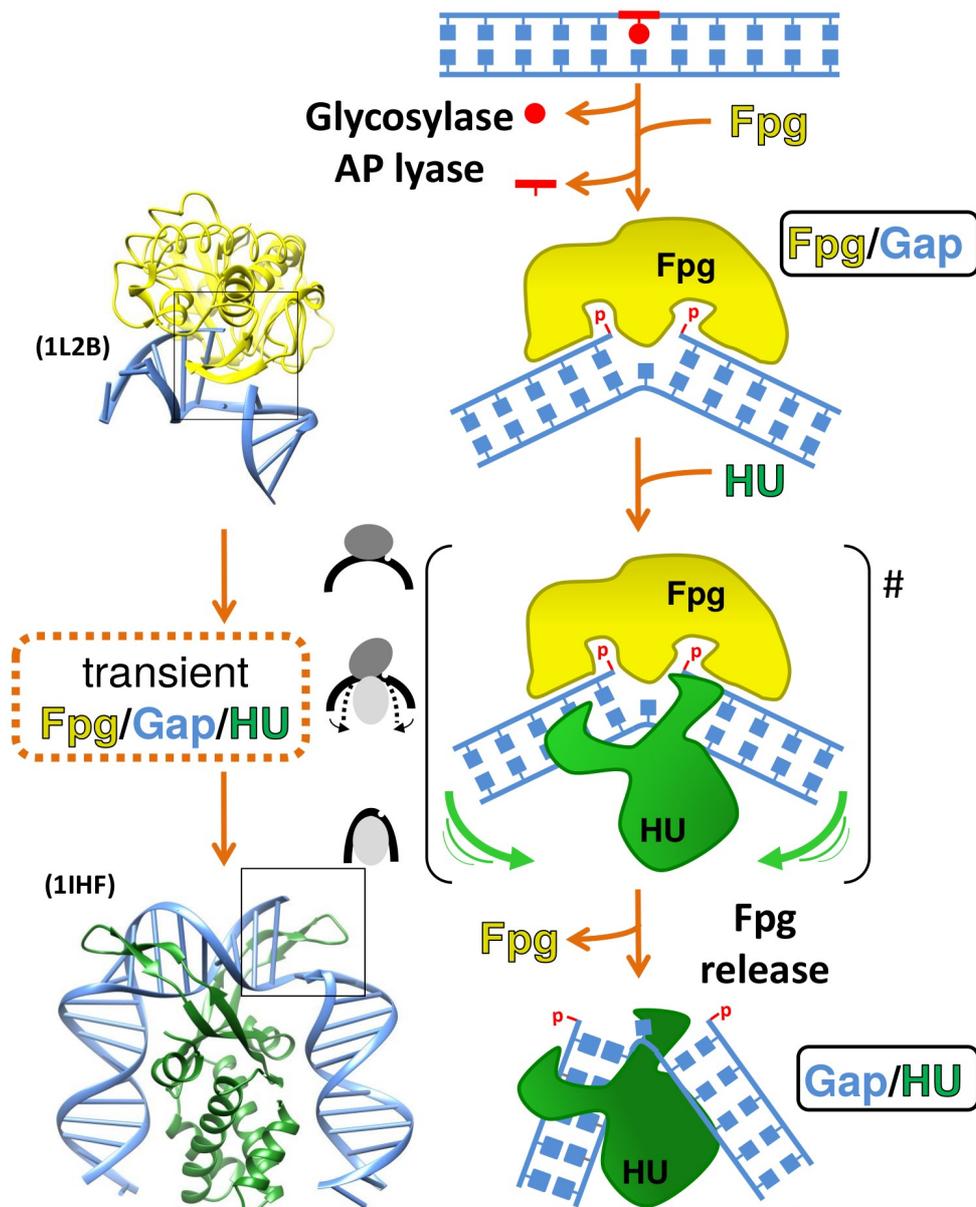


Figure 7

Figure 7: Proposed model for the stimulation of the Fpg turnover by HU according to an active displacement mechanism of Fpg from its final DNA product

Conclusion

The proposed interaction mode between HU and Fpg is similar to that proposed for the interaction of the N-terminal domain of the human XRCC1 protein with DNA polymerase beta (Pol β) bound to gap-containing DNA (62). Like DNA glycosylases, Pol β (a non-processive gap filling polymerase involved in the short patch BER sub-pathway) interacts with the convex side of the DNA curvature

whereas XRCC1 interacts on the concave side like HU. Unlike HU which binds indifferently free and Fpg bound Gap-DNA (Fig.4), XRCC1 can slightly interact with Pol β but single-strand break DNA bound by XRCC1 is a higher affinity ligand for Pol β than the free DNA. XRCC1, which is essential for genome maintenance (like HU), is involved in base excision and single-strand break repair (BER and SSBR). Like HU, XRCC1 binds to single-strand breaks and modulates DNA glycosylases (63). More than just a single-strand break binding protein, XRCC1 physically interacts with multiple DNA glycosylases and acts as a recruitment platform for other BER partners (63). It has therefore been proposed that XRCC1 coordinates BER and SSBR enzymes by preventing the potential toxicity of transient DNA repair intermediates such as single-strand breaks from coming forward (60,64). A similar role has been proposed for HU to prevent the formation of double-strand breaks during the BER action on closely opposed lesions including oxidized bases and single-strand breaks (24). In such a situation, the sharp DNA bent induced by the binding of HU to a single-strand break impairs the recognition and removal of the closely oxidized base on the opposite strand by DNA glycosylases. HU is believed to establish a sequential order of DNA lesion management by BER (coordination) and thus prevent the formation of double-strand breaks at the multi injured DNA site. Similarly, HU can also be seen as a bacterial poly(ADP-ribose) polymerase 1 (PARP1) homolog for its involvement in chromatin structural modeling, its strong interaction with SSB and its potential role in the recruitment and modulation of repair enzymes at the damaged DNA site (65). Among its many roles in DNA transactions, we can conclude that HU is a DNA repair actor required in bacteria for an efficient BER/SSBR.

Funding

This work was supported by the Centre national de la recherche scientifique (CNRS, France), the committees of “Charente-Maritime”, “Indre-et-Loire”, “Loiret” “Finistère”, “Cher” and “Deux-Sèvres” of the “Ligue contre le cancer Grand-Ouest” (CSIRGO), the Scientific Radioprotection Committee of “Electricité de France” (EDF), the ARC foundation and the Région Centre. R.L.M. is supported by a doctoral fellowship from the France government.

Acknowledgments

The authors wish to thank Mélanie Spothem-Maurizot and Serge Boiteux for their critical reading of the manuscript and helpful discussion.

Bibliography

1. Luijsterburg, M.S., White, M.F., van Driel, R. and Dame, R.T. (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Critical reviews in biochemistry and molecular biology*, **43**, 393-418.
2. Dillon, S.C. and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature reviews. Microbiology*, **8**, 185-195.
3. Rimsky, S. and Travers, A. (2011) Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. *Current opinion in microbiology*, **14**, 136-141.
4. Oberto, J., Drlica, K. and Rouviere-Yaniv, J. (1994) Histones, HMG, HU, IHF: Meme combat. *Biochimie*, **76**, 901-908.
5. DeLange, R.J., Williams, L.C. and Searcy, D.G. (1981) A histone-like protein (HTa) from *Thermoplasma acidophilum*. II. Complete amino acid sequence. *The Journal of biological chemistry*, **256**, 905-911.
6. Kobayashi, T., Takahara, M., Miyagishima, S.Y., Kuroiwa, H., Sasaki, N., Ohta, N., Matsuzaki, M. and Kuroiwa, T. (2002) Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *The Plant cell*, **14**, 1579-1589.
7. Johnson, G.G. and Geiduschek, E.P. (1972) Purification of the bacteriophage SP01 transcription factor 1. *The Journal of biological chemistry*, **247**, 3571-3578.
8. Oberto, J., Nabti, S., Jooste, V., Mignot, H. and Rouviere-Yaniv, J. (2009) The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PloS one*, **4**, e4367.
9. Claret, L. and Rouviere-Yaniv, J. (1996) Regulation of HU alpha and HU beta by CRP and FIS in *Escherichia coli*. *Journal of molecular biology*, **263**, 126-139.
10. Giangrossi, M., Giuliadori, A.M., Gualerzi, C.O. and Pon, C.L. (2002) Selective expression of the beta-subunit of nucleoid-associated protein HU during cold shock in *Escherichia coli*. *Molecular microbiology*, **44**, 205-216.
11. Gualerzi, C.O., Giuliadori, A.M. and Pon, C.L. (2003) Transcriptional and post-transcriptional control of cold-shock genes. *Journal of molecular biology*, **331**, 527-539.
12. Guo, F. and Adhya, S. (2007) Spiral structure of *Escherichia coli* HUalpha-beta provides foundation for DNA supercoiling. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 4309-4314.
13. Mouw, K.W. and Rice, P.A. (2007) Shaping the *Borrelia burgdorferi* genome: crystal structure and binding properties of the DNA-bending protein Hbb. *Molecular microbiology*, **63**, 1319-1330.
14. Ramstein, J., Hervouet, N., Coste, F., Zelwer, C., Oberto, J. and Castaing,

- B. (2003) Evidence of a thermal unfolding dimeric intermediate for the Escherichia coli histone-like HU proteins: thermodynamics and structure. *Journal of molecular biology*, **331**, 101-121.
15. Swinger, K.K. and Rice, P.A. (2004) IHF and HU: flexible architects of bent DNA. *Current opinion in structural biology*, **14**, 28-35.
 16. Swinger, K.K. and Rice, P.A. (2007) Structure-based analysis of HU-DNA binding. *Journal of molecular biology*, **365**, 1005-1016.
 17. Pinson, V., Takahashi, M. and Rouviere-Yaniv, J. (1999) Differential binding of the Escherichia coli HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *Journal of molecular biology*, **287**, 485-497.
 18. Garnier, N., Loth, K., Coste, F., Augustyniak, R., Nadan, V., Damblon, C. and Castaing, B. (2011) An alternative flexible conformation of the E. coli HUbeta(2) protein: structural, dynamics, and functional aspects. *European biophysics journal : EBJ*, **40**, 117-129.
 19. Boubrik, F. and Rouviere-Yaniv, J. (1995) Increased sensitivity to gamma irradiation in bacteria lacking protein HU. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 3958-3962.
 20. Joseph, N., Sawarkar, R. and Rao, D.N. (2004) DNA mismatch correction in Haemophilus influenzae: characterization of MutL, MutH and their interaction. *DNA repair*, **3**, 1561-1577.
 21. Li, S. and Waters, R. (1998) Escherichia coli strains lacking protein HU are UV sensitive due to a role for HU in homologous recombination. *Journal of bacteriology*, **180**, 3750-3756.
 22. Mukherjee, A., Bhattacharyya, G. and Grove, A. (2008) The C-terminal domain of HU-related histone-like protein Hlp from Mycobacterium smegmatis mediates DNA end-joining. *Biochemistry*, **47**, 8744-8753.
 23. Castaing, B., Zelwer, C., Laval, J. and Boiteux, S. (1995) HU protein of Escherichia coli binds specifically to DNA that contains single-strand breaks or gaps. *The Journal of biological chemistry*, **270**, 10291-10296.
 24. Hashimoto, M., Imhoff, B., Ali, M.M. and Kow, Y.W. (2003) HU protein of Escherichia coli has a role in the repair of closely opposed lesions in DNA. *The Journal of biological chemistry*, **278**, 28501-28507.
 25. Kamashev, D. and Rouviere-Yaniv, J. (2000) The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *The EMBO journal*, **19**, 6527-6535.
 26. Dalhus, B., Laerdahl, J.K., Backe, P.H. and Bjoras, M. (2009) DNA base repair--recognition and initiation of catalysis. *FEMS microbiology reviews*, **33**, 1044-1078.
 27. Wallace, S.S. (2013) DNA glycosylases search for and remove oxidized DNA bases. *Environmental and molecular mutagenesis*, **54**, 691-704.
 28. Hill, J.W., Hazra, T.K., Izumi, T. and Mitra, S. (2001) Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic acids research*, **29**, 430-438.
 29. Mokkapati, S.K., Wiederhold, L., Hazra, T.K. and Mitra, S. (2004) Stimulation of DNA glycosylase activity of OGG1 by NEIL1: functional collaboration between two human DNA glycosylases. *Biochemistry*, **43**,

- 11596-11604.
30. Parikh, S.S., Putnam, C.D. and Tainer, J.A. (2000) Lessons learned from structural results on uracil-DNA glycosylase. *Mutation research*, **460**, 183-199.
 31. Pope, M.A., Porello, S.L. and David, S.S. (2002) Escherichia coli apurinic-apyrimidinic endonucleases enhance the turnover of the adenine glycosylase MutY with G:A substrates. *The Journal of biological chemistry*, **277**, 22605-22615.
 32. Saitoh, T., Shinmura, K., Yamaguchi, S., Tani, M., Seki, S., Murakami, H., Nojima, Y. and Yokota, J. (2001) Enhancement of OGG1 protein AP lyase activity by increase of APEX protein. *Mutation research*, **486**, 31-40.
 33. Sidorenko, V.S., Nevinsky, G.A. and Zharkov, D.O. (2007) Mechanism of interaction between human 8-oxoguanine-DNA glycosylase and AP endonuclease. *DNA repair*, **6**, 317-328.
 34. Yang, H., Clendenin, W.M., Wong, D., Demple, B., Slupska, M.M., Chiang, J.H. and Miller, J.H. (2001) Enhanced activity of adenine-DNA glycosylase (Myh) by apurinic/apyrimidinic endonuclease (Ape1) in mammalian base excision repair of an A/GO mismatch. *Nucleic acids research*, **29**, 743-752.
 35. Rabow, L., Venkataraman, R. and Kow, Y.W. (2001) Mechanism of action of Escherichia coli formamidopyrimidine N-glycosylase: role of K155 in substrate binding and product release. *Progress in nucleic acid research and molecular biology*, **68**, 223-234.
 36. Castaing, B., Geiger, A., Seliger, H., Nehls, P., Laval, J., Zelwer, C. and Boiteux, S. (1993) Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant FPG proteins. *Nucleic acids research*, **21**, 2899-2905.
 37. Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C. and Castaing, B. (2004) Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. *The Journal of biological chemistry*, **279**, 44074-44083.
 38. Castaing, B., Fourrey, J.L., Hervouet, N., Thomas, M., Boiteux, S. and Zelwer, C. (1999) AP site structural determinants for Fpg specific recognition. *Nucleic acids research*, **27**, 608-615.
 39. Duwat, P., de Oliveira, R., Ehrlich, S.D. and Boiteux, S. (1995) Repair of oxidative DNA damage in gram-positive bacteria: the Lactococcus lactis Fpg protein. *Microbiology*, **141 (Pt 2)**, 411-417.
 40. Girard, P.M. and Boiteux, S. (1997) Repair of oxidized DNA bases in the yeast Saccharomyces cerevisiae. *Biochimie*, **79**, 559-566.
 41. Porello, S.L., Leyes, A.E. and David, S.S. (1998) Single-turnover and pre-steady-state kinetics of the reaction of the adenine glycosylase MutY with mismatch-containing DNA substrates. *Biochemistry*, **37**, 14756-14764.
 42. Riggs, A.D., Bourgeois, S. and Cohn, M. (1970) The lac repressor-operator interaction. 3. Kinetic studies. *Journal of molecular biology*, **53**, 401-417.
 43. Ghosh, S., Mallick, B. and Nagaraja, V. (2014) Direct regulation of topoisomerase activity by a nucleoid-associated protein. *Nucleic acids research*, **42**, 11156-11165.
 44. Koh, J., Saecker, R.M. and Record, M.T., Jr. (2008) DNA binding mode

- transitions of Escherichia coli HU(alpha-beta): evidence for formation of a bent DNA--protein complex on intact, linear duplex DNA. *Journal of molecular biology*, **383**, 324-346.
45. Kundukad, B., Cong, P., van der Maarel, J.R. and Doyle, P.S. (2013) Time-dependent bending rigidity and helical twist of DNA by rearrangement of bound HU protein. *Nucleic acids research*, **41**, 8280-8288.
 46. Wojtuszewski, K., Hawkins, M.E., Cole, J.L. and Mukerji, I. (2001) HU binding to DNA: evidence for multiple complex formation and DNA bending. *Biochemistry*, **40**, 2588-2598.
 47. Wojtuszewski, K. and Mukerji, I. (2003) HU binding to bent DNA: a fluorescence resonance energy transfer and anisotropy study. *Biochemistry*, **42**, 3096-3104.
 48. Menoni, H., Gasparutto, D., Hamiche, A., Cadet, J., Dimitrov, S., Bouvet, P. and Angelov, D. (2007) ATP-dependent chromatin remodeling is required for base excision repair in conventional but not in variant H2A.Bbd nucleosomes. *Molecular and cellular biology*, **27**, 5949-5956.
 49. Menoni, H., Shukla, M.S., Gerson, V., Dimitrov, S. and Angelov, D. (2012) Base excision repair of 8-oxoG in dinucleosomes. *Nucleic acids research*, **40**, 692-700.
 50. Swinger, K.K., Lemberg, K.M., Zhang, Y. and Rice, P.A. (2003) Flexible DNA bending in HU-DNA cocystal structures. *The EMBO journal*, **22**, 3749-3760.
 51. Pereira de Jesus, K., Serre, L., Zelwer, C. and Castaing, B. (2005) Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. *Nucleic acids research*, **33**, 5936-5944.
 52. Tchou, J. and Grollman, A.P. (1995) The catalytic mechanism of Fpg protein. Evidence for a Schiff base intermediate and amino terminus localization of the catalytic site. *The Journal of biological chemistry*, **270**, 11671-11677.
 53. Fromme, J.C. and Verdine, G.L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nature structural biology*, **9**, 544-552.
 54. Gilboa, R., Zharkov, D.O., Golan, G., Fernandes, A.S., Gerchman, S.E., Matz, E., Kycia, J.H., Grollman, A.P. and Shoham, G. (2002) Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *The Journal of biological chemistry*, **277**, 19811-19816.
 55. Serre, L., Pereira de Jesus, K., Boiteux, S., Zelwer, C. and Castaing, B. (2002) Crystal structure of the Lactococcus lactis formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA. *The EMBO journal*, **21**, 2854-2865.
 56. van der Kemp, P.A., Thomas, D., Barbey, R., de Oliveira, R. and Boiteux, S. (1996) Cloning and expression in Escherichia coli of the OGG1 gene of Saccharomyces cerevisiae, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 5197-5202.

57. Bruner, S.D., Norman, D.P. and Verdine, G.L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature*, **403**, 859-866.
58. Dalhus, B., Forsbring, M., Helle, I.H., Vik, E.S., Forstrom, R.J., Backe, P.H., Alseth, I. and Bjoras, M. (2011) Separation-of-function mutants unravel the dual-reaction mode of human 8-oxoguanine DNA glycosylase. *Structure*, **19**, 117-127.
59. Mol, C.D., Izumi, T., Mitra, S. and Tainer, J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature*, **403**, 451-456.
60. Wilson, S.H. and Kunkel, T.A. (2000) Passing the baton in base excision repair. *Nature structural biology*, **7**, 176-178.
61. Collier, C., Machon, C., Briggs, G.S., Smits, W.K. and Soutlanas, P. (2012) Untwisting of the DNA helix stimulates the endonuclease activity of *Bacillus subtilis* Nth at AP sites. *Nucleic acids research*, **40**, 739-750.
62. Marintchev, A., Mullen, M.A., Maciejewski, M.W., Pan, B., Gryk, M.R. and Mullen, G.P. (1999) Solution structure of the single-strand break repair protein XRCC1 N-terminal domain. *Nature structural biology*, **6**, 884-893.
63. Campalans, A., Marsin, S., Nakabeppu, Y., O'Connor T, R., Boiteux, S. and Radicella, J.P. (2005) XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair. *DNA repair*, **4**, 826-835.
64. Rice, P.A. (1999) Holding damaged DNA together. *Nature structural biology*, **6**, 805-806.
65. Robert, I., Karicheva, O., Reina San Martin, B., Schreiber, V. and Dantzer, F. (2013) Functional aspects of PARylation in induced and programmed DNA repair processes: preserving genome integrity and modulating physiological events. *Molecular aspects of medicine*, **34**, 1138-1152.

Supplementary materials

Figure S

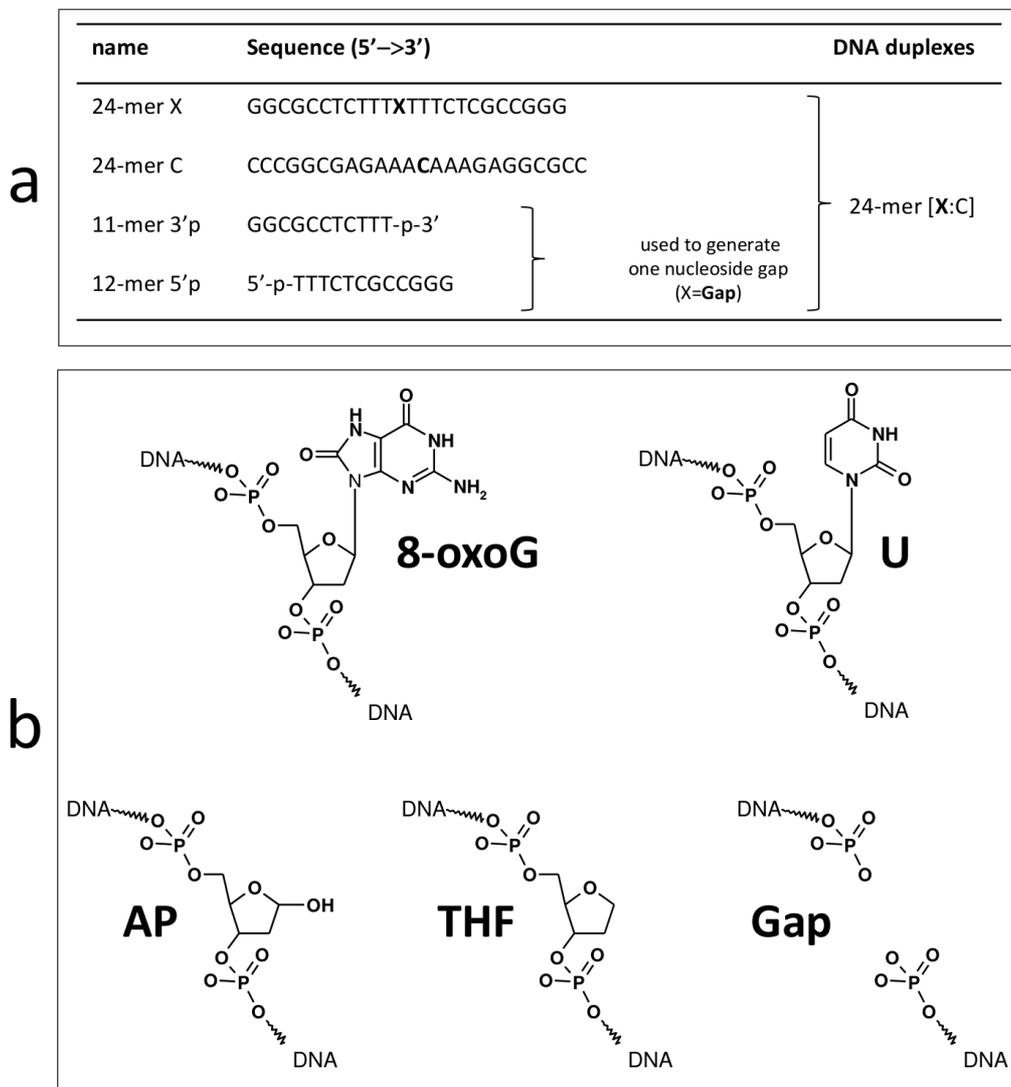


Figure S1: Oligonucleotides used in this study

(a) Sequences of single-strand oligonucleotides used to generate 24-mer [X:C] **(b)** Structures of the modified DNA sites (X) present in DNA duplexes.

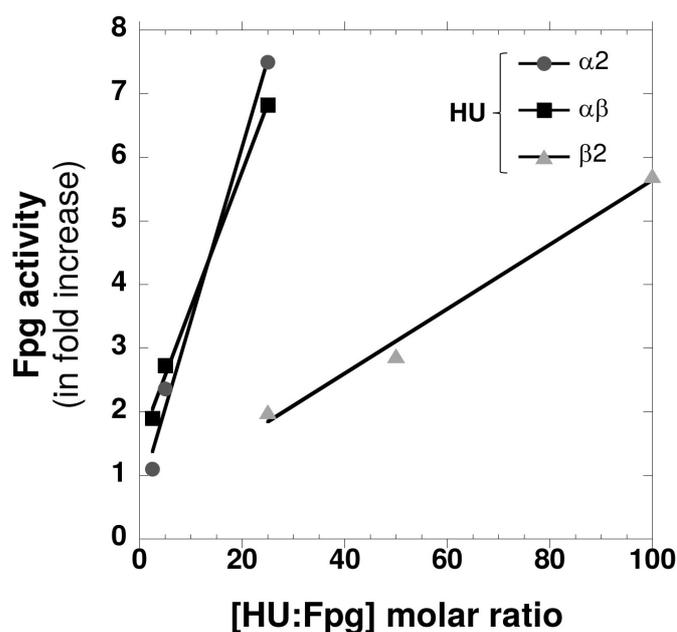


Figure S2: HU concentration-dependence of the Fpg stimulation at low HU concentrations
Experimental conditions are described in the legend of Fig.1.

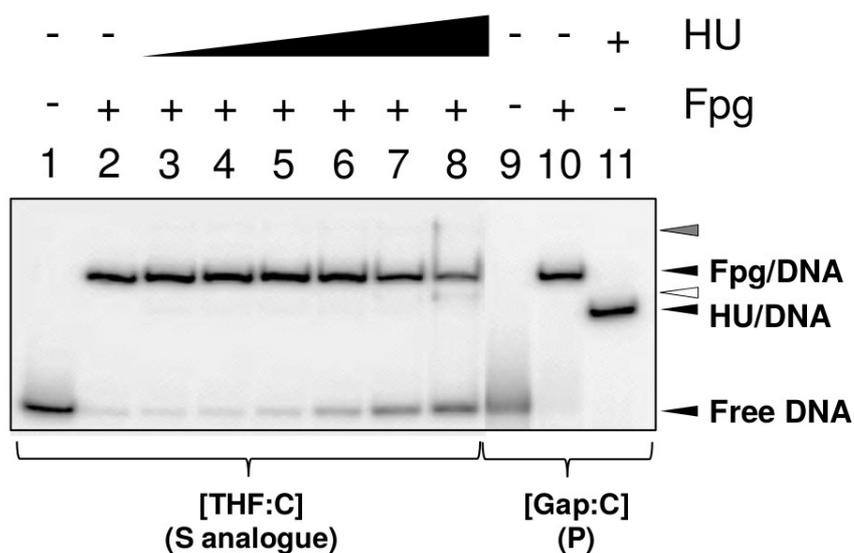


Figure S3: Effect of HU on a mimicry Fpg/DNA-substrate complex

0.1 nM of the 24-mer [THF:C] DNA duplex (lane 1, THF = tetrahydrofurane AP site analogue) was pre-incubated 30 min at 8°C with 2.2 nM of Fpg. 0, 4.5, 9, 18, 36, 72 or 144 nM of HU was added to the pre-incubation mixture (lanes 2 to 8, respectively), incubated for an additional 20 min at 8°C and analyzed by EMSA. As controls, a 24-mer [Gap:C] (lane 9) was incubated 20 min at 8°C in the presence of 2.2 nM of Fpg or HU as indicated (lanes 10 and 11, respectively). Equilibrated samples were then analyzed by EMSA. The *open* and *grey* arrows indicate the migration position of the non-specific HU/[THF:C] complex and that of a putative Fpg/[THF:C]/HU ternary complex, respectively.

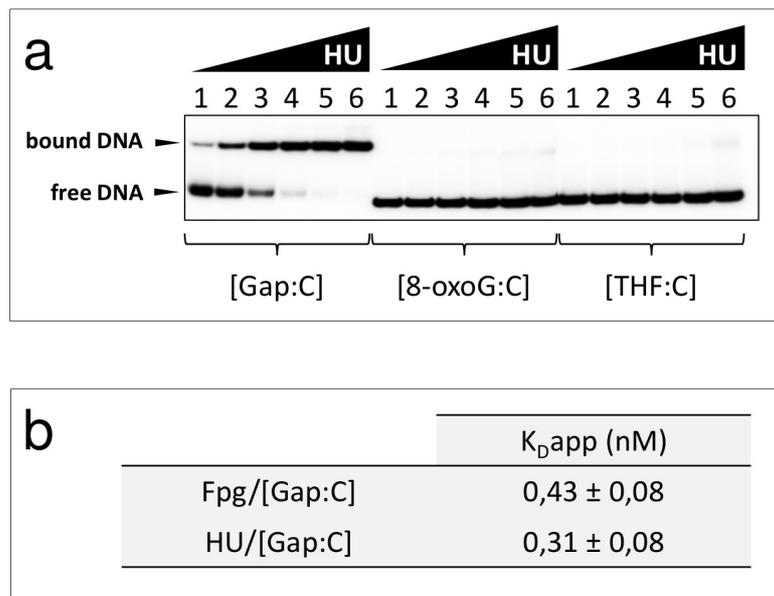


Figure S4: Preferential binding of HU to the Fpg end-DNA product

0.1 nM of indicated 24-mer DNA probe was incubated with 0.12, 0.24, 0.49, 0.98, 1.96 and 3.92 nM of HU as described in the legend of Fig.3. At equilibrium, mixtures were analyzed by EMSA. **(a)** Representative titration experiments. **(b)** Apparent dissociation constants ($K_{D,app}$). Dissociation constants were determined by EMSA titration experiments as described in the *Materials and Methods* section. Indicated values correspond to the mean values (\pm SD) obtained from three independent determinations.

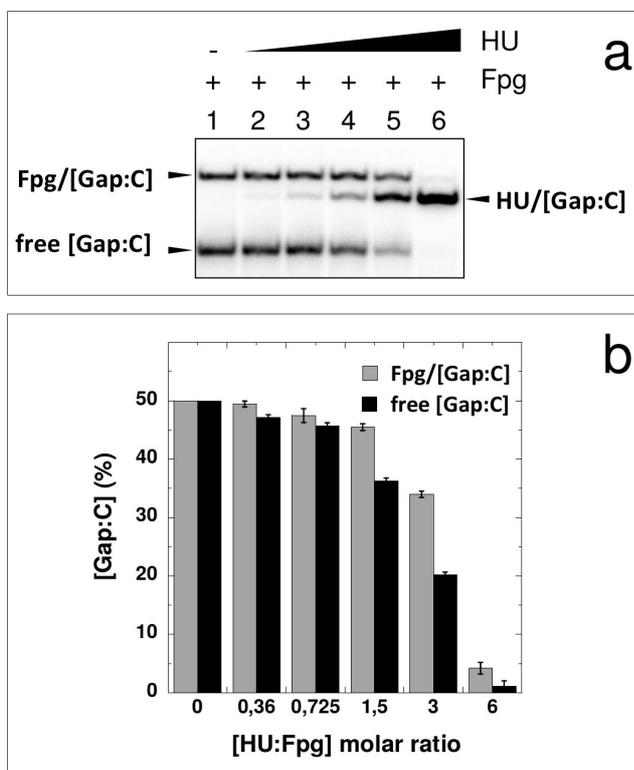


Figure S5: Comparative titration of the free and Fpg-bound DNA-product by HU

0.1 nM of 24-mer [Gap:C] (Fpg end-product) was pre-incubated for 30 min at 8°C with 0.12 nM of Fpg. This equilibrium mixture was then incubated for 10 min at 8°C with 0, 0.043, 0.087, 0.18, 0.36 and 0.72 nM of HU (lanes 1 to 6, respectively). Reaction mixtures were then analyzed by EMSA as described in Fig.3. **(a)** Representative autoradiography **(b)** Plot of the free and Fpg-bound [Gap:C] remaining as a function of the [HU:Fpg] molar ratio. Quantification was done from three independent experiments.

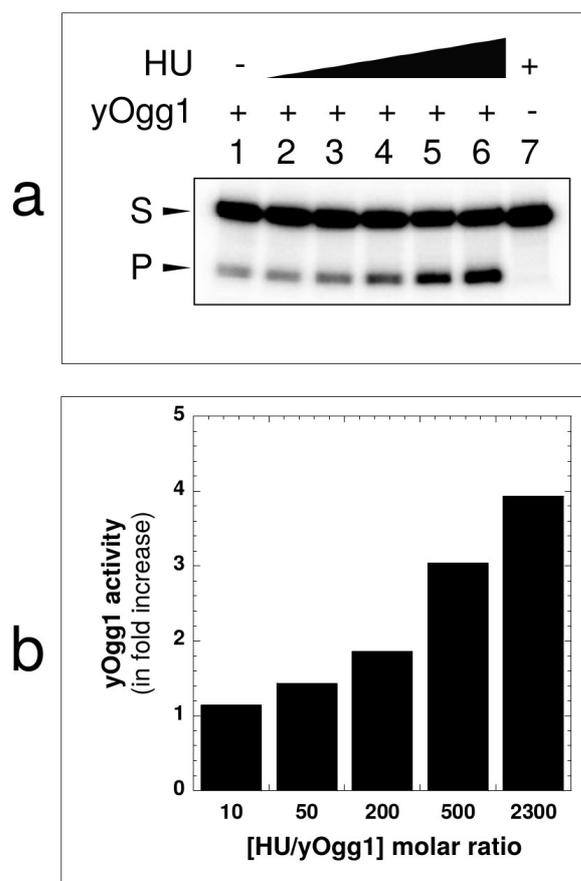


Figure S6: Modulation of the 8-oxoG DNA glycosylase activity of Ogg1 by the HU protein

25nM of a radiolabeled 24-mer-[8-oxoG:C] duplex (Fig.S1) was incubated for 10 min at 37°C with HU alone (lane 7) or with a limiting amount of yOgg1 (3nM) (lanes1) in the presence of a molar excess of [HU:yOgg1] of 10, 50, 200, 500, 2300 (lanes 2 to 6, respectively). Reactions were stopped by adding NaOH (0.2M final) followed by heating at 50° for 5 min. The products were analyzed as described in *Materials and Methods*. **(a)** Representative gel autoradiography. (S) and (P) are for Ogg1 substrate and AP-product, respectively. **(b)** Quantifications of autoradiography. The yOgg1 activity stimulation factor (fold increase) corresponds to the ratio between the relative quantity of AP site product generated by yOgg1 in the presence of HU and that without HU.

Figure S7: HU stimulates the release of Ogg1 from AP site-containing DNA

3 nM of yOgg1 was incubated at 37°C with 25 nM of a 24-mer [8-oxoG:C] DNA with or without 600 nM of HU. The removal of 8-oxoG (resulting in the formation of 24-mer [AP:C]) was measured as described in *Materials and methods*. Alkali treatment revealed AP sites which are not processed by yOgg1 under the experimental conditions used. Using the same reaction scheme of Fpg (scheme 1) raw data can be fitted using eq.1. A twofold increase in the rate constant of the [AP:C] released by yOgg1 was observed for the [HU:yOgg1] ratio of 200.

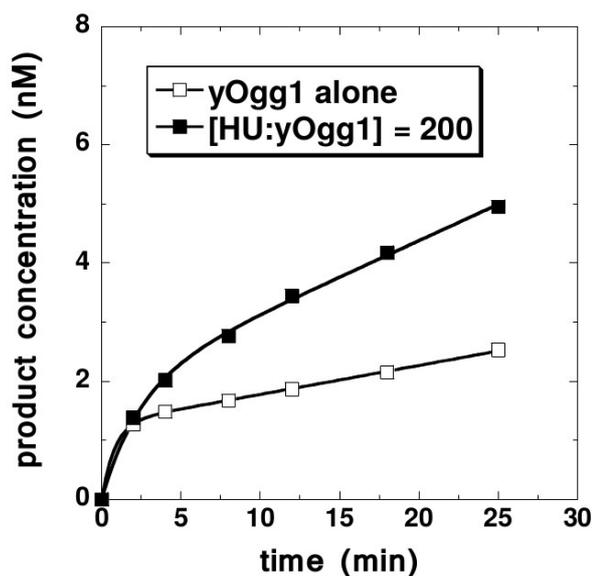
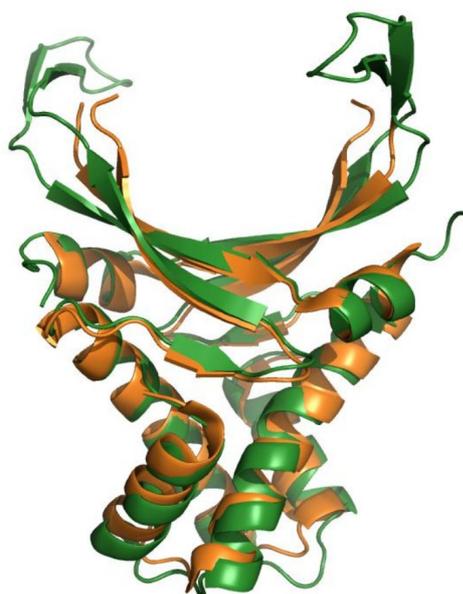


Figure S8: Superposition of the structure of HU α_2 and IHF from *E. coli*

Ribbon superposition of the peptide backbone observed in the crystal structures of HU α_2 (1MUL, in orange) and IHF (1IHF, in green) bound to DNA (14,15). For the sake of clarity, DNA in the IHF/DNA complex was omitted. The high flexible part of the arms of HU α_2 is not visible in the crystal structure. Through DNA binding, the arms of IHF adopt a fixed conformation and are entirely observed in the crystal structure. Similar observations were observed in the crystal structures of HU bound to DNA (12,13,50). The RMSD on C α between the two structures is 1.0 Å. The protein bodies of both dimers are perfectly aligned and the intrinsic dynamic of the arm is responsible for the divergence between the two protein structures.



<i>Bst</i> Fpg	PELPEVETIIRRTLLPLIVGKTIEDVRIFWPNII RHPRDSEAFARMIGQT	50
<i>Ec</i> Fpg	PELPEVETSRRGIEPHLVGATILHA - VVRNGRLRWPVSEIY - -RLSDQP	47
<i>Bst</i> Fpg	VRGLERRGKFLKFLDDRDLI SHLRMEGRYAVASALEPLEPHTHVFCFT	100
<i>Ec</i> Fpg	VLSVQRRAKYLLELPEGWII IHLGMSGSLRILPEELPPEKHDHVDLVMS	97
	*	
<i>Bst</i> Fpg	DGSELRYRDVRFKFTMHVYAKEEADRRPPLAELGPEPLSPAFAVLAER	150
<i>Ec</i> Fpg	NGKVLRYTDPRRFG - AWLWTKELEGHN - VLTHLGPPEPLSDDFNGEYLHQK	145
<i>Bst</i> Fpg	AVKTKRSVKALLLDQTVVAGFGNIYVDESLFRAGILPGRPAASLSKEIE	200
<i>Ec</i> Fpg	CAKKKTAIKPWLMDNKL VVGFGNIYASESLFAAGIHPDRLASSLSLAECE	195
	*	
<i>Bst</i> Fpg	RLHEEMVATIGEAVMKGSTVTRTYVNTQGEAGTFQHHLVYVGRQGNPCKR	250
<i>Ec</i> Fpg	LLARVIKAVLLRSIEQGGTTLKDFLQSDGKPGYFAQELQVYGRKGEPCRV	245
	*	
<i>Bst</i> Fpg	CGTPIEKTVVAGRTHYCPRCQR	273
<i>Ec</i> Fpg	CGTPIVATKHAQRATFYCRQCQK	268
	*	

a

<i>Ec</i> HU α	- -MNKTQLIDVIAEK - AELSKTQAKAALESTLAAITESLKEGDAVQLVGF ^{GTF}	50
<i>Ec</i> HU β	- -MNKSQLIDKIAAG - ADISKAAAGRALDAIIASVTESLKEGDDVALVGF ^{GTF}	50
<i>Ec</i> IHF α	MALTKAEMSEYLFDK - LGLSKRDAKELVELFFEEIRRALENGEQVKLSGF ^{GNF}	50
<i>Ec</i> IHF β	- -MTKSELIERLATQQSHIPAKTVEAVKEMLEHMASTLAQGERIEIRGF ^{GSF}	50
<i>Ec</i> HU α	KVNHRAE ^{RTGRNPQ} TGKEIKIAAANVPFVSGKALKDAVK - - - - -	90
<i>Ec</i> HU β	AVKERAAR ^{TGRNPQ} TGKEITIAAAKVPSPRAGKALKDAVN - - - - -	90
<i>Ec</i> IHF α	DLRDKNQ ^{RPGRNPK} TGEDIPITARRVVTFRPGQKLKSRVENASPKDE	99
<i>Ec</i> IHF β	SLHYRAP ^{RTGRNPK} TGDKVELEGKYVPHFKPGKELDRANIYG - - - - -	94
	*	

b

Figure S9: Primary structure alignments Fpg, HU and IHF

(a) Sequence alignment of Fpg from *Escherichia coli* (*Ec*) and *Bacillus stearothermophilus* (*Bst*)

(b) Sequence alignment of α and β chains of *E. coli* HU and IHF proteins

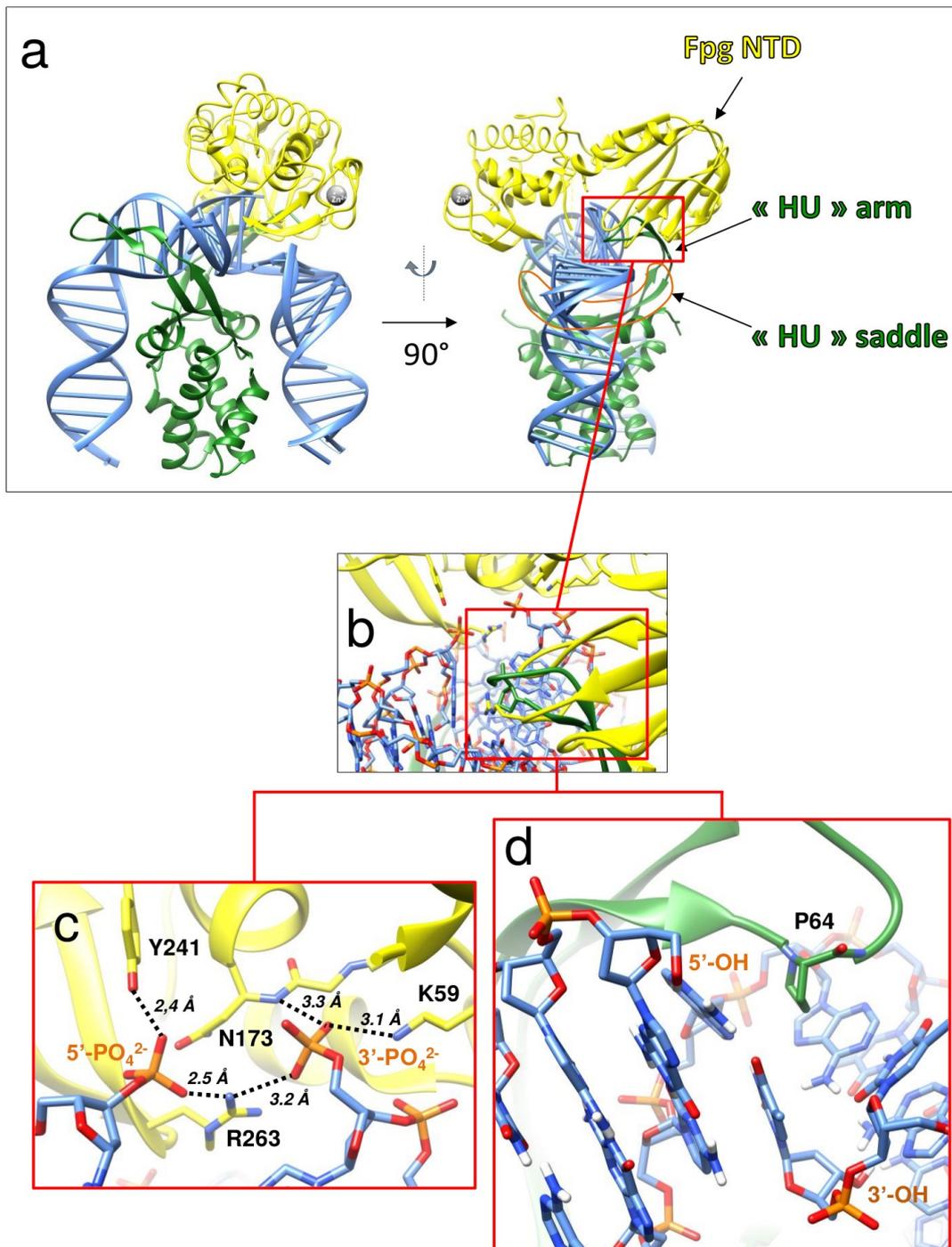


Figure S10: Proposed interaction mode of HU with the complex of Fpg bound to its end-DNA product

(a) Proposed model for the Fpg/DNA/HU ternary complex Two orthogonal views of superposition of *G. stearothermophilus* Fpg (GstFpg, yellow ribbon)/one nucleoside gap-DNA (1L2B) with *E. coli* IHF (green)/nicked DNA (1IHF) are shown. Due to the similarity of the two proteins HU and IHF (Fig.S8), IHF was used as a model for HU bound to single-strand breaks. Superposition of the structures of the two nucleoprotein complexes was made by aligning DNA molecules (blue) as well

as possible at the single-strand break level (a one nucleoside gap in 1L2B and a DNA nick in 1IHF). Steric clashes between the two structures are indicated by a red box **(b)** *Zoom of the clashes*. **(c)** and **(d)** *Separate views in the immediate vicinity of the DNA single-strand break in Fpg and IHF (HU) bound to DNA, respectively*. Residues of Fpg contacting the phosphates bordering the single-strand break are indicated. Strictly conserved P64 of the arm of the β subunit of IHF intercalated by the DNA minor groove in the vicinity of the break is indicated. The nature (-OH or -phosphate) of the DNA break ends is indicated.

Tables

Table 1: Effect of HU on the rate constant of the dissociation of Fpg from its final DNA product

The rate constants k_3 (in min^{-1}) (Scheme 1) for the concerted 8-oxoG-DNA glycosylase/AP lyase activity and for the uncoupled AP lyase activity of Fpg were extracted from the kinetic experiments presented in Fig.2a and 2b, respectively, using eq.1 as described in the *Materials and Methods* section. The reported values correspond to the mean values (\pm SD) obtained from four independent experiments. “0” is for “control experiment without HU”. For an HU/Fpg molar ratio above 20 (Fig.2), the extraction of the k_3 was not applicable because the release of the enzyme from its DNA product becomes no longer limiting.

HU/Fpg	Glycosylase/lyase		Lyase	
	k_3	Fold increase	k_3	Fold increase
“0”	0.58 ± 0.1	1	0.41 ± 0.1	1
5	1.5 ± 0.2	2.6	1.4 ± 0.3	3.4
10	2.4 ± 0.5	4.1	2.1 ± 0.4	5.1

7.5. Conclusions et perspectives

Nous avons démontré dans cette étude qu'une protéine chromatinienne de *Escherichia coli* (HU sous ses 3 isoformes dimériques) est capable de stimuler le « turnover » de la protéine Fpg. Sur la base de données cinétiques obtenues avec HU α_2 , nous proposons que le mécanisme moléculaire par lequel HU stimule Fpg passe par la formation d'un complexe ternaire instable entre l'enzyme piégée sur son produit final de réaction (une lacune de 1 nucléoside) et HU. La formation de ce complexe ternaire est immédiatement suivie par un relargage de Fpg de son produit de réaction et par la formation d'un complexe stable entre HU et le produit de réaction de Fpg (voir Figure 7, Article). Nous privilégions un mécanisme actif au cours duquel HU reconnaît l'ADN courbé dans le complexe stable Fpg/ADN. Les données structurales disponibles sur la fixation de Fpg et HU à l'ADN supporte ce modèle car Fpg interagit plutôt avec l'ADN du côté convexe de la torsion induite par l'enzyme tandis que HU interagit du côté concave. Suite à un changement structural associé à la fixation de HU à l'ADN, les contacts Fpg/ADN dans le complexe Fpg/ADN/HU sont déstabilisés puis rompus ce qui induit le relargage de Fpg et donc son recyclage (stimulation du « turnover »). Dans ce modèle, la stimulation de Fpg par HU ne peut être effective que si le complexe ternaire est transitoire (i.e., instable) et que le déplacement par HU de son produit de réaction empêche Fpg de s'y refixer. La formation d'un complexe ternaire stable aurait au mieux pas d'effet sur le « turnover » de l'enzyme ou au pire un effet négatif. Un tel mécanisme actif a déjà été proposé pour le recrutement de l'ADN polymérase β (Pol β) au niveau des lacunes dans l'ADN (lacunes de un à plusieurs nucléotides) par la protéine XRCC1.²⁶⁰ Pol β et XRCC1 sont impliquées dans le système BER chez l'homme. XRCC1 est souvent présenté comme une plate-forme de recrutement au niveau des cassures simple-brin et est connu également pour être capable de stimuler Ogg1, l'homologue fonctionnel de Fpg chez les eucaryotes supérieurs.²⁵⁶ Nous proposons en conclusion que HU pourrait-être assimilée à un homologue fonctionnel bactérien de XRCC1. Dans le cadre de cette hypothèse, le complexe HU/ADN-cassure simple-brin pourrait constituer un signal cellulaire pour des partenaires du BER ou d'autres systèmes de réparation agissant en aval des ADN glycosylases. Cette hypothèse doit être éprouvée par de nouvelles investigations. Dans le but d'élucider au niveau atomique les bases structurales du complexe ternaire Fpg/ADN/HU, nous avons voulu initier une étude structurale par cristallographie de ce complexe. Le fait que le modèle structurale historique de Fpg dans l'équipe soit l'enzyme de la bactérie *Lactococcus lactis* (LIFpg), il a été décidé de produire et purifier l'homologue de HU de cet organisme, LIHU (voir Chap.4). Nous avons également utilisé LIHU comme modèle d'homodimère unique de type mésophile pour

lequel le mécanisme d'échange de chaîne n'est pas requis. Au cours d'une étude préliminaire, nous avons observé que L₁HU, contrairement à HUβ₂ son homologue de *E. coli*, se fixe plus fortement à l'ADN normal (G:C, **Figure 114**) mais sans une préférence très marquée pour une lacune de 1 nucléoside (Gap:C, **Figure 114**).

Dans la gamme HU/Fpg 25 à 100 pour lequel nous avons observé une stimulation forte de l'enzyme dans le modèle *E. coli*, peu ou pas d'effets sont observés dans le modèle *L. lactis*. Cependant, une stimulation très forte de L₁Fpg par L₁HU est observée pour des rapports très élevés HU/Fpg (facteur de stimulation de 16 fois pour un rapport de [HU:Fpg] = 31250, **Figure 115a**). Les cinétiques enzymatiques en condition de multiple « turnover » ([Substrat] >> [Enzyme]) indiquent comme dans le modèle *E. coli* que la stimulation ne résulte pas d'un effet sur la catalyse à proprement parlé mais d'une stimulation de la vitesse de dissociation de l'enzyme de son produit final de réaction (**Figure 115 b**, voir aussi Article pour plus de détails).

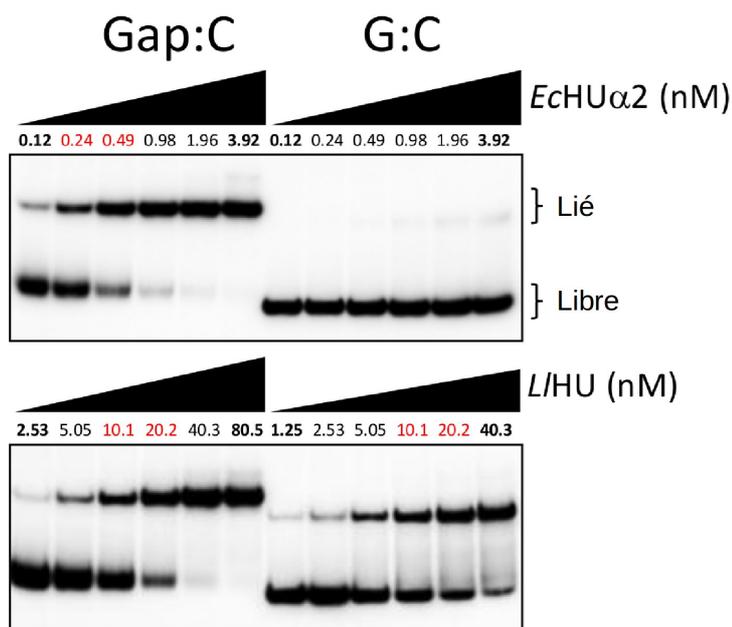


Figure 114: Analyse comparative de la fixation à l'ADN des protéines HU de *E. coli* et de *L. lactis* Les sondes nucléiques [G:C] (ADN normal) et [Gap:C] (ADN contenant une lacune de un nucléoside = produit final de réaction de Fpg) sont incubées en présence de concentrations croissantes de HU. A l'équilibre, les mélanges d'incubation sont analysés par électrophorèse en gel retard (EMSA, voir article).

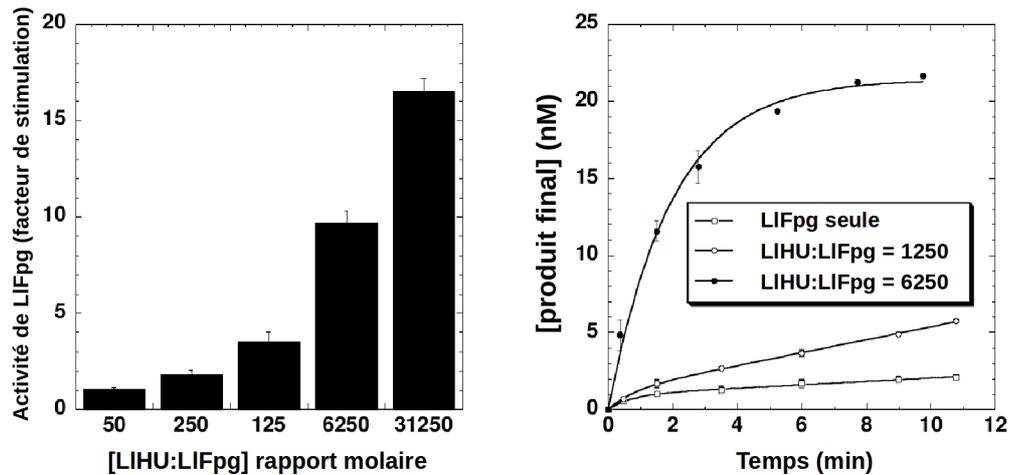


Figure 115: Stimulation de LIFpg par LIHU Les lettres « L1 » précédant le nom des protéines signifient « *L. lactis* ». Pour les détails expérimentaux se référer à l'article.

Nous ne pouvons garantir que de tels rapports entre ces 2 protéines chez *L. lactis* soient pertinents d'un point de vue physiologique. De manière non-équivalente, pour des rapports [HU:EcFpg] supérieur à 50 pour HU α_2 et HU $\alpha\beta$ et à 250 pour HU β_2 , l'activité de EcFpg est progressivement impactée jusqu'à l'observation d'une inhibition totale pour des rapports supérieurs à 500 (Fig.1, Article). Ce n'est pas du tout ce que l'on observe avec les protéines de *L. lactis*. L'ensemble de ces observations suggèrent que les modes de fixation à l'ADN des isoformes de HU de *E. coli* et LIHU sont très différents en particulier en fonction du rapport HU/ADN. Cette étude fonctionnelle comparative entre protéines HU suggère qu'outre son rôle général dans structuration du chromosome bactérien (qui semble vrai pour tous les systèmes bactériens étudiés), le système HU peut selon les phylums bactériens ne pas jouer complètement les mêmes rôles et être plus au moins complexe (3 protéines chez les entérobactéries et 1 protéine chez les autres bactéries). Ce type de situation a déjà été observé pour la stimulation de l'activité de relaxation de l'ADN de la topoisomérase I par HU chez *Mycobacterium tuberculosis*. En effet, une telle stimulation de la Topo I par HU n'est pas observée chez *E. coli*.²⁶¹ En conséquence, nous ne pouvons pas affirmer que HU soit capable de stimuler les enzymes du BER chez toutes les bactéries. Cette stimulation est probablement très dépendante d'une forte affinité de la protéine pour les cassures simple-brin. Dans un système hétérologue, nous avons montré aussi que HU α_2 pouvait stimuler l'activité ADN

glycosylase Ogg1, l'homologue fonctionnel eucaryotes de Fpg. Cette dernière observation confirme qu'une interaction physique directe entre HU et l'ADN glycosylase n'est pas absolument nécessaire pour stimuler son activité et passe plutôt par une reconnaissance de l'ADN ayant subi une forte courbure imposée par la fixation à l'ADN de l'ADN glycosylase. Dans le cas de Ogg1, la stimulation n'est pas dépendante de la formation d'une cassure simple-brin, l'excision de la base est suffisante (formation d'un site abasique). Nous ne pouvons pas complètement exclure une interaction directe entre les protéines bactériennes, mais nous n'avons pas pour l'instant été capables de la mettre en évidence expérimentalement. L'approche structurale du complexe ternaire par cristallographie ou en solution (SAXS) doit donc se faire avec les protéines d'*E. coli*. Cette étude est en cours et nécessite au préalable que l'on soit capable de stabiliser artificiellement ce complexe.

Bibliographie

1. Pellicer, J., Fay, M. F. & Leitch, I. J. The largest eukaryotic genome of them all? *Bot. J. Linn. Soc.* **164**, 10–15 (2010).
2. Wu, H. Y., Shyy, S. H., Wang, J. C. & Liu, L. F. Transcription generates positively and negatively supercoiled domains in the template. *Cell* **53**, 433–40 (1988).
3. Wang, J. C. Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine. *Q. Rev. Biophys.* **31**, 107–44 (1998).
4. Stewart, L., Ireton, G. C. & Champoux, J. J. The domain organization of human topoisomerase I. *J. Biol. Chem.* **271**, 7602–8 (1996).
5. Bergerat, A. *et al.* An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* **386**, 414–7 (1997).
6. Fussner, E., Ching, R. W. & Bazett-Jones, D. P. Living without 30nm chromatin fibers. *Trends Biochem. Sci.* **36**, 1–6 (2011).
7. Daban, J. R. & Bermúdez, A. Interdigitated solenoid model for compact chromatin fibers. *Biochemistry* **37**, 4299–304 (1998).
8. Robinson, P. J. J., Fairall, L., Huynh, V. A. T. & Rhodes, D. EM measurements define the dimensions of the “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6506–11 (2006).
9. Robinson, P. J. J. & Rhodes, D. Structure of the “30 nm” chromatin fibre: a key role for the linker histone. *Curr. Opin. Struct. Biol.* **16**, 336–43 (2006).
10. Dorigo, B. *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**, 1571–3 (2004).
11. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436**, 138–41 (2005).

12. Li, F. *et al.* The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutS α . *Cell* **153**, 590–600 (2013).
13. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–80 (2001).
14. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421**, 448–53 (2003).
15. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–60 (1997).
16. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319**, 1097–113 (2002).
17. White, C. L., Suto, R. K. & Luger, K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207–18 (2001).
18. Sperling, R. & Bustin, M. Histone dimers: a fundamental unit in histone assembly. *Nucleic Acids Res.* **3**, 1263–75 (1976).
19. Bustin, M. Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem. Sci.* **26**, 152–3 (2001).
20. Reeves, R. & Nissen, M. S. The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J. Biol. Chem.* **265**, 8573–82 (1990).
21. Hill, D. A., Pedulla, M. L. & Reeves, R. Directional binding of HMG-I(Y) on four-way junction DNA and the molecular basis for competitive binding with HMG-1 and histone H1. *Nucleic Acids Res.* **27**, 2135–44 (1999).
22. Nissen, M. S. & Reeves, R. Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y. *J. Biol. Chem.* **270**, 4355–60 (1995).
23. Sgarra, R. *et al.* HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* **1799**, 37–47 (2010).

24. Reeves, R. & Nissen, M. S. Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles. *J. Biol. Chem.* **268**, 21137–46 (1993).
25. Knapp, S. *et al.* The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. *Biochemistry* **43**, 11992–7 (2004).
26. Pil, P. M. & Lippard, S. J. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science* **256**, 234–7 (1992).
27. Travers, A. A. Priming the nucleosome: a role for HMGB proteins? *EMBO Rep.* **4**, 131–6 (2003).
28. Bonaldi, T., Längst, G., Strohner, R., Becker, P. B. & Bianchi, M. E. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. *EMBO J.* **21**, 6865–73 (2002).
29. Bustin, M. Chromatin unfolding and activation by HMGN(*) chromosomal proteins. *Trends Biochem. Sci.* **26**, 431–7 (2001).
30. Körner, U., Bustin, M., Scheer, U. & Hock, R. Developmental role of HMGN proteins in *Xenopus laevis*. *Mech. Dev.* **120**, 1177–92 (2003).
31. Furusawa, T. *et al.* Down-regulation of nucleosomal binding protein HMGN1 expression during embryogenesis modulates Sox9 expression in chondrocytes. *Mol. Cell. Biol.* **26**, 592–604 (2006).
32. Lehtonen, S. & Lehtonen, E. HMG-17 is an early marker of inductive interactions in the developing mouse kidney. *Differentiation.* **67**, 154–63 (2001).
33. Catez, F., Brown, D. T., Misteli, T. & Bustin, M. Competition between histone H1 and HMGN proteins for chromatin binding sites. *EMBO Rep.* **3**, 760–6 (2002).
34. Li, W. T., Sandman, K., Pereira, S. L. & Reeve, J. N. MJ1647, an open reading frame in the genome of the hyperthermophile *Methanococcus jannaschii*, encodes a very thermostable archaeal histone with a C-terminal extension. *Extremophiles* **4**, 43–51 (2000).
35. Pereira, S. L., Grayling, R. A., Lurz, R. & Reeve, J. N. Archaeal nucleosomes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12633–7 (1997).

36. Bailey, K. A., Chow, C. S. & Reeve, J. N. Histone stoichiometry and DNA circularization in archaeal nucleosomes. *Nucleic Acids Res.* **27**, 532–6 (1999).
37. Sandman, K., Grayling, R. A., Dobrinski, B., Lurz, R. & Reeve, J. N. Growth-phase-dependent synthesis of histones in the archaeon *Methanothermobacter thermautotrophicus*. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12624–8 (1994).
38. Xue, H., Guo, R., Wen, Y., Liu, D. & Huang, L. An abundant DNA binding protein from the hyperthermophilic archaeon *Sulfolobus shibatae* affects DNA supercoiling in a temperature-dependent fashion. *J. Bacteriol.* **182**, 3929–33 (2000).
39. Bell, S. D., Botting, C. H., Wardleworth, B. N., Jackson, S. P. & White, M. F. The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science* **296**, 148–51 (2002).
40. Guo, R., Xue, H. & Huang, L. Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. *Mol. Microbiol.* **50**, 1605–15 (2003).
41. Bohrmann, B. *et al.* Localization of Histone-like Proteins in Thermophilic Archaea by Immunogold Electron Microscopy. *J. Struct. Biol.* **112**, 70–78 (1994).
42. Jelinska, C. *et al.* Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure* **13**, 963–71 (2005).
43. Biyani, K. *et al.* Solution structure, stability, and nucleic acid binding of the hyperthermophile protein Sso10b2. *Biochemistry* **44**, 14217–30 (2005).
44. Wu, X. *et al.* Thermal unfolding of the archaeal DNA and RNA binding protein Ssh10. *Biochem. Biophys. Res. Commun.* **373**, 482–7 (2008).
45. Guo, L. *et al.* Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res.* **36**, 1129–37 (2008).
46. Driessen, R. P. C. *et al.* Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. *Nucleic Acids Res.* **41**, 196–205 (2013).
47. Niu, Y. *et al.* A prototypic lysine methyltransferase 4 from archaea with degenerate sequence specificity methylates chromatin proteins Sul7d and Cren7

- in different patterns. *J. Biol. Chem.* **288**, 13728–40 (2013).
48. Edmondson, S. P. & Shriver, J. W. DNA binding proteins Sac7d and Sso7d from *Sulfolobus*. *Methods Enzymol.* **334**, 129–45 (2001).
 49. Luo, X. *et al.* CC1, a novel crenarchaeal DNA binding protein. *J. Bacteriol.* **189**, 403–9 (2007).
 50. Murzin, A. G. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861–7 (1993).
 51. Culard, F., Laine, B., Sautière, P. & Maurizot, J. C. Stoichiometry of the binding of chromosomal protein MC1 from the archaeobacterium, *Methanosarcina* spp. CHTI55, to DNA. *FEBS Lett.* **315**, 335–9 (1993).
 52. Teyssier, C. *et al.* Preferential binding of the archaeobacterial histone-like MC1 protein to negatively supercoiled DNA minicircles. *Biochemistry* **35**, 7954–8 (1996).
 53. Paradinas, C., Gervais, A., Maurizot, J. C. & Culard, F. Structure-specific binding recognition of a methanogen chromosomal protein. *Eur. J. Biochem.* **257**, 372–9 (1998).
 54. Paquet, F. *et al.* Model of a DNA-protein complex of the architectural monomeric protein MC1 from Euryarchaea. *PLoS One* **9**, e88809 (2014).
 55. Cam, E. L., Culard, F., Larquet, E., Delain, E. & Cognet, J. A. DNA bending induced by the archaeobacterial histone-like protein MC1. *J. Mol. Biol.* **285**, 1011–21 (1999).
 56. DeLange, R. J., Green, G. R. & Searcy, D. G. A histone-like protein (HTa) from *Thermoplasma acidophilum*. I. Purification and properties. *J. Biol. Chem.* **256**, 900–4 (1981).
 57. DeLange, R. J., Williams, L. C. & Searcy, D. G. A histone-like protein (HTa) from *Thermoplasma acidophilum*. II. Complete amino acid sequence. *J. Biol. Chem.* **256**, 905–11 (1981).
 58. Worcel, A. & Burgi, E. On the structure of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* **71**, 127–147 (1972).

59. Postow, L., Hardy, C. D., Arsuaga, J. & Cozzarelli, N. R. Topological domain structure of the Escherichia coli chromosome. *Genes Dev.* **18**, 1766–79 (2004).
60. Teleman, A. A., Graumann, P. L., Lin, D. C., Grossman, A. D. & Losick, R. Chromosome arrangement within a bacterium. *Curr. Biol.* **8**, 1102–9 (1998).
61. Niki, H., Yamaichi, Y. & Hiraga, S. Dynamic organization of chromosomal DNA in Escherichia coli. *Genes Dev.* **14**, 212–23 (2000).
62. Valens, M., Penaud, S., Rossignol, M., Cornet, F. & Boccard, F. Macrodomain organization of the Escherichia coli chromosome. *EMBO J.* **23**, 4330–41 (2004).
63. Dillon, S. C. & Dorman, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**, 185–95 (2010).
64. Rimsky, S. & Travers, A. Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. *Curr. Opin. Microbiol.* **14**, 136–41 (2011).
65. Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D. & Possoz, C. MukB colocalizes with the oriC region and is required for organization of the two Escherichia coli chromosome arms into separate cell halves. *Mol. Microbiol.* **65**, 1485–92 (2007).
66. Graumann, P. L. SMC proteins in bacteria: condensation motors for chromosome segregation? *Biochimie* **83**, 53–9 (2001).
67. Li, Y., Schoeffler, A. J., Berger, J. M. & Oakley, M. G. The crystal structure of the hinge domain of the Escherichia coli structural maintenance of chromosomes protein MukB. *J. Mol. Biol.* **395**, 11–9 (2010).
68. Cui, Y., Petrushenko, Z. M. & Rybenkov, V. V. MukB acts as a macromolecular clamp in DNA condensation. *Nat. Struct. Mol. Biol.* **15**, 411–8 (2008).
69. Petrushenko, Z. M., Lai, C.-H., Rai, R. & Rybenkov, V. V. DNA reshaping by MukB. Right-handed knotting, left-handed supercoiling. *J. Biol. Chem.* **281**, 4606–15 (2006).
70. Petrushenko, Z. M., Cui, Y., She, W. & Rybenkov, V. V. Mechanics of DNA bridging by bacterial condensin MukBEF in vitro and in singulo. *EMBO J.* **29**, 1126–35 (2010).

71. Hu, K. H. *et al.* Overproduction of three genes leads to camphor resistance and chromosome condensation in *Escherichia coli*. *Genetics* **143**, 1521–32 (1996).
72. Chen, N. *et al.* ATP-induced shrinkage of DNA with MukB protein and the MukBEF complex of *Escherichia coli*. *J. Bacteriol.* **190**, 3731–7 (2008).
73. Thanbichler, M. & Shapiro, L. Chromosome organization and segregation in bacteria. *J. Struct. Biol.* **156**, 292–303 (2006).
74. Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* **181**, 6361–70 (1999).
75. Johnson, R. C., Bruist, M. F. & Simon, M. I. Host protein requirements for in vitro site-specific DNA inversion. *Cell* **46**, 531–9 (1986).
76. Koch, C. & Kahmann, R. Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. *J. Biol. Chem.* **261**, 15673–8 (1986).
77. Thompson, J. F., Moitoso de Vargas, L., Koch, C., Kahmann, R. & Landy, A. Cellular factors couple recombination with growth phase: characterization of a new component in the lambda site-specific recombination pathway. *Cell* **50**, 901–8 (1987).
78. Pan, C. Q. *et al.* Variable structures of Fis-DNA complexes determined by flanking DNA-protein contacts. *J. Mol. Biol.* **264**, 675–95 (1996).
79. Kostrewa, D. *et al.* Crystal structure of the factor for inversion stimulation FIS at 2.0 Å resolution. *J. Mol. Biol.* **226**, 209–26 (1992).
80. Hancock, S. P. *et al.* Control of DNA minor groove width and Fis protein binding by the purine 2-amino group. *Nucleic Acids Res.* **41**, 6750–60 (2013).
81. Schneider, R. *et al.* An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Res.* **29**, 5107–14 (2001).
82. Skoko, D., Yan, J., Johnson, R. C. & Marko, J. F. Low-force DNA condensation and discontinuous high-force decondensation reveal a loop-stabilizing function of the protein Fis. *Phys. Rev. Lett.* **95**, 208101 (2005).
83. Schneider, R., Travers, A., Kutateladze, T. & Muskhelishvili, G. A DNA

- architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. *Mol. Microbiol.* **34**, 953–64 (1999).
84. Schneider, R., Travers, A. & Muskhelishvili, G. FIS modulates growth phase-dependent topological transitions of DNA in *Escherichia coli*. *Mol. Microbiol.* **26**, 519–30 (1997).
 85. Xu, J. & Johnson, R. C. *aldB*, an RpoS-dependent gene in *Escherichia coli* encoding an aldehyde dehydrogenase that is repressed by Fis and activated by Crp. *J. Bacteriol.* **177**, 3166–75 (1995).
 86. González-Gil, G., Bringmann, P. & Kahmann, R. FIS is a regulator of metabolism in *Escherichia coli*. *Mol. Microbiol.* **22**, 21–9 (1996).
 87. Travers, A., Schneider, R. & Muskhelishvili, G. DNA supercoiling and transcription in *Escherichia coli*: The FIS connection. *Biochimie* **83**, 213–7 (2001).
 88. Topping, T. B., Hoch, D. A. & Gloss, L. M. Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation. *J. Mol. Biol.* **335**, 1065–81 (2004).
 89. Meinhold, D., Boswell, S. & Colón, W. P61A mutation in the factor for inversion stimulation results in a thermostable dimeric intermediate. *Biochemistry* **44**, 14715–24 (2005).
 90. Meinhold, D., Beach, M., Shao, Y., Osuna, R. & Colón, W. The location of an engineered inter-subunit disulfide bond in factor for inversion stimulation (FIS) affects the denaturation pathway and cooperativity. *Biochemistry* **45**, 9767–77 (2006).
 91. Merickel, S. K., Sanders, E. R., Vázquez-Ibar, J. L. & Johnson, R. C. Subunit exchange and the role of dimer flexibility in DNA binding by the Fis protein. *Biochemistry* **41**, 5788–98 (2002).
 92. Shindo, H. *et al.* Solution structure of the DNA binding domain of a nucleoid-associated protein, H-NS, from *Escherichia coli*. *FEBS Lett.* **360**, 125–31 (1995).
 93. Bloch, V. *et al.* The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat. Struct. Biol.* **10**, 212–8 (2003).

94. Falconi, M., Gualtieri, M. T., La Teana, A., Losso, M. A. & Pon, C. L. Proteins from the prokaryotic nucleoid: primary and quaternary structure of the 15-kD *Escherichia coli* DNA binding protein H-NS. *Mol. Microbiol.* **2**, 323–9 (1988).
95. Smyth, C. P. *et al.* Oligomerization of the chromatin-structuring protein H-NS. *Mol. Microbiol.* **36**, 962–72 (2000).
96. Suzuki, C. *et al.* Oligomerization Mechanisms of an H-NS Family Protein, Pmr, Encoded on the Plasmid pCAR1 Provide a Molecular Basis for Functions of H-NS Family Members. *PLoS One* **9**, e105656 (2014).
97. Arold, S. T., Leonard, P. G., Parkinson, G. N. & Ladbury, J. E. H-NS forms a superhelical protein scaffold for DNA condensation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15728–32 (2010).
98. Dame, R. T. *et al.* DNA bridging: a property shared among H-NS-like proteins. *J. Bacteriol.* **187**, 1845–8 (2005).
99. Dame, R. T., Wyman, C. & Goosen, N. H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28**, 3504–10 (2000).
100. Luijsterburg, M. S., Noom, M. C., Wuite, G. J. L. & Dame, R. T. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* **156**, 262–72 (2006).
101. Lim, C. J., Kenney, L. J. & Yan, J. Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. *Nucleic Acids Res.* **42**, 8369–78 (2014).
102. McGovern, V., Higgins, N. P., Chiz, R. S. & Jaworski, A. H-NS over-expression induces an artificial stationary phase by silencing global transcription. *Biochimie* **76**, 1019–29 (1994).
103. Spurio, R. *et al.* Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. *Mol. Gen. Genet.* **231**, 201–11 (1992).
104. Ali, S. S. *et al.* Structural insights into the regulation of foreign genes in *Salmonella* by the Hha/H-NS complex. *J. Biol. Chem.* **288**, 13356–69 (2013).
105. Salomon, D., Klimko, J. A. & Orth, K. H-NS regulates the *Vibrio parahaemolyticus* type VI secretion system 1. *Microbiology* **160**, 1867–73

(2014).

106. Brambilla, L., Morán-Barrio, J. & Viale, A. M. Expression of the *Escherichia coli* ompW colicin S4 receptor gene is regulated by temperature and modulated by the H-NS and StpA nucleoid-associated proteins. *FEMS Microbiol. Lett.* **352**, 238–44 (2014).
107. Singh, S. S. *et al.* Widespread suppression of intragenic transcription initiation by H-NS. *Genes Dev.* **28**, 214–9 (2014).
108. Sharadamma, N., Harshavardhana, Y., Singh, P. & Muniyappa, K. *Mycobacterium tuberculosis* nucleoid-associated DNA-binding protein H-NS binds with high-affinity to the Holliday junction and inhibits strand exchange promoted by RecA protein. *Nucleic Acids Res.* **38**, 3555–69 (2010).
109. Shiraishi, K., Ogata, Y., Hanada, K., Kano, Y. & Ikeda, H. Roles of the DNA binding proteins H-NS and StpA in homologous recombination and repair of bleomycin-induced damage in *Escherichia coli*. *Genes Genet. Syst.* **82**, 433–9 (2007).
110. Friedman, D. I. Integration host factor: a protein for all reasons. *Cell* **55**, 545–54 (1988).
111. Wang, S., Cosstick, R., Gardner, J. F. & Gumport, R. I. The specific binding of *Escherichia coli* integration host factor involves both major and minor grooves of DNA. *Biochemistry* **34**, 13082–90 (1995).
112. Murtin, C., Engelhorn, M., Geiselman, J. & Boccard, F. A quantitative UV laser footprinting analysis of the interaction of IHF with specific binding sites: re-evaluation of the effective concentration of IHF in the cell. *J. Mol. Biol.* **284**, 949–61 (1998).
113. Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**, 1295–306 (1996).
114. Dame, R. T. *et al.* Analysis of scanning force microscopy images of protein-induced DNA bending using simulations. *Nucleic Acids Res.* **33**, e68 (2005).
115. Goosen, N. & van de Putte, P. The regulation of transcription initiation by integration host factor. *Mol. Microbiol.* **16**, 1–7 (1995).
116. Dos Santos, M. T. & Rodrigues, P. S. A genomic-scale search for regulatory

- binding sites in the integration host factor regulon of *Escherichia coli* K12. *Genet. Mol. Res.* **4**, 783–9 (2005).
117. Kolodrubetz, D., Phillips, L. & Burgum, A. Repression of aerobic leukotoxin transcription by integration host factor in *Aggregatibacter actinomycetemcomitans*. *Res. Microbiol.* **161**, 541–8 (2010).
 118. Pitre, C. A. J., Tanner, J. R., Patel, P. & Brassinga, A. K. C. Regulatory control of temporally expressed integration host factor (IHF) in *Legionella pneumophila*. *Microbiology* **159**, 475–92 (2013).
 119. Freundlich, M., Ramani, N., Mathew, E., Sirko, A. & Tsui, P. The role of integration host factor in gene expression in *Escherichia coli*. *Mol. Microbiol.* **6**, 2557–63 (1992).
 120. Fiévet, A., Cascales, E., Valette, O., Dolla, A. & Aubert, C. IHF is required for the transcriptional regulation of the *Desulfovibrio vulgaris* Hildenborough *orp* operons. *PLoS One* **9**, e86507 (2014).
 121. Silva-Rocha, R., Chavarría, M., Kleijn, R. J., Sauer, U. & de Lorenzo, V. The IHF regulon of exponentially growing *Pseudomonas putida* cells. *Environ. Microbiol.* **15**, 49–63 (2013).
 122. Torres-Escobar, A., Juárez-Rodríguez, M. D. & Demuth, D. R. Integration host factor is required for replication of pYGK-derived plasmids in *Aggregatibacter actinomycetemcomitans*. *FEMS Microbiol. Lett.* **357**, 184–94 (2014).
 123. Ryan, V. T., Grimwade, J. E., Nievera, C. J. & Leonard, A. C. IHF and HU stimulate assembly of pre-replication complexes at *Escherichia coli* *oriC* by two different mechanisms. *Mol. Microbiol.* **46**, 113–24 (2002).
 124. Nozaki, S., Yamada, Y. & Ogawa, T. Initiator titration complex formed at *datA* with the aid of IHF regulates replication timing in *Escherichia coli*. *Genes Cells* **14**, 329–41 (2009).
 125. Ali, B. M. *et al.* Compaction of single DNA molecules induced by binding of integration host factor (IHF). *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10658–63 (2001).
 126. Krylov, A. S., Zasedateleva, O. A., Prokopenko, D. V., Rouviere-Yaniv, J. & Mirzabekov, A. D. Massive parallel analysis of the binding specificity of

- histone-like protein HU to single- and double-stranded DNA with generic oligodeoxyribonucleotide microchips. *Nucleic Acids Res.* **29**, 2654–60 (2001).
127. Prieto, A. I. *et al.* Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Res.* (2011). doi:10.1093/nar/gkr1236
 128. Castaing, B., Zelwer, C., Laval, J. & Boiteux, S. HU protein of *Escherichia coli* binds specifically to DNA that contains single-strand breaks or gaps. *J. Biol. Chem.* **270**, 10291–6 (1995).
 129. Kamashev, D., Balandina, A. & Rouviere-Yaniv, J. The binding motif recognized by HU on both nicked and cruciform DNA. *EMBO J.* **18**, 5434–44 (1999).
 130. Pinson, V., Takahashi, M. & Rouviere-Yaniv, J. Differential binding of the *Escherichia coli* HU, homodimeric forms and heterodimeric form to linear, gapped and cruciform DNA. *J. Mol. Biol.* **287**, 485–97 (1999).
 131. Kamashev, D. & Rouviere-Yaniv, J. The histone-like protein HU binds specifically to DNA recombination and repair intermediates. *EMBO J.* **19**, 6527–35 (2000).
 132. Balandina, A., Claret, L., Hengge-Aronis, R. & Rouviere-Yaniv, J. The *Escherichia coli* histone-like protein HU regulates *rpoS* translation. *Mol. Microbiol.* **39**, 1069–79 (2001).
 133. Wang, H.-C., Wu, M.-L., Ko, T.-P. & Wang, A. H.-J. *Neisseria* conserved hypothetical protein DMP12 is a DNA mimic that binds to histone-like HU protein. *Nucleic Acids Res.* **41**, 5127–38 (2013).
 134. Parsons, L. M., Liu, F. & Orban, J. HU- α binds to the putative double-stranded DNA mimic HI1450 from *Haemophilus influenzae*. *Protein Sci.* **14**, 1684–7 (2005).
 135. Dame, R. T. & Goosen, N. HU: promoting or counteracting DNA compaction? *FEBS Lett.* **529**, 151–156 (2002).
 136. Koh, J., Saecker, R. M. & Record, M. T. DNA binding mode transitions of *Escherichia coli* HU(α): evidence for formation of a bent DNA--protein complex on intact, linear duplex DNA. *J. Mol. Biol.* **383**, 324–46 (2008).

137. Endo, T., Sasaki, N., Tanaka, I. & Nakata, M. Compact form of DNA induced by DNA-binding protein HU. *Biochem. Biophys. Res. Commun.* **290**, 546–51 (2002).
138. Esser, D. *et al.* A hyperthermostable bacterial histone-like protein as an efficient mediator for transfection of eukaryotic cells. *Nat. Biotechnol.* **18**, 1211–3 (2000).
139. Noort, J. Van, Verbrugge, S., Goosen, N., Dekker, C. & Dame, R. T. Dual architectural roles of HU: Formation of flexible hinges and rigid filaments. *PNAS* **101**, 6969–6974 (2004).
140. Xiao, B., Johnson, R. C. & Marko, J. F. Modulation of HU-DNA interactions by salt concentration and applied force. *Nucleic Acids Res.* **38**, 6176–85 (2010).
141. Wada, M., Kano, Y., Ogawa, T., Okazaki, T. & Imamoto, F. Construction and characterization of the deletion mutant of hupA and hupB genes in *Escherichia coli*. *J. Mol. Biol.* **204**, 581–91 (1988).
142. Kar, S., Edgar, R. & Adhya, S. Nucleoid remodeling by an altered HU protein: reorganization of the transcription program. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16397 (2005).
143. Oberto, J., Nabti, S., Jooste, V., Mignot, H. & Rouviere-Yaniv, J. The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS One* **4**, e4367 (2009).
144. Mangan, M. W. *et al.* Nucleoid-associated protein HU controls three regulons that coordinate virulence, response to stress and general physiology in *Salmonella enterica* serovar Typhimurium. *Microbiology* **157**, 1075–87 (2011).
145. Aki, T., Choy, H. E. & Adhya, S. Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor. *Genes Cells* **1**, 179–88 (1996).
146. Aki, T. & Adhya, S. Repressor induced site-specific binding of HU for transcriptional regulation. *EMBO J.* **16**, 3666–74 (1997).
147. Lyubchenko, Y. L., Shlyakhtenko, L. S., Aki, T. & Adhya, S. Atomic force microscopic demonstration of DNA looping by GalR and HU. *Nucleic Acids Res.* **25**, 873–6 (1997).

148. Geanacopoulos, M., Vasmatzis, G., Zhurkin, V. B. & Adhya, S. Gal repressosome contains an antiparallel DNA loop. *Nat. Struct. Biol.* **8**, 432–6 (2001).
149. Koli, P., Sudan, S., Fitzgerald, D., Adhya, S. & Kar, S. Conversion of Commensal Escherichia coli K-12 to an Invasive Form via Expression of a Mutant Histone-Like Protein. *MBio* **2**, (2011).
150. Chodavarapu, S., Felczak, M. M., Yaniv, J. R. & Kaguni, J. M. Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin. *Mol. Microbiol.* **67**, 781–92 (2008).
151. Oberto, J. & Rouviere-Yaniv, J. Does the parallel evolution pattern between the replication-segregation proteins and HU have a biological significance? *Biochimie* **83**, 61–6 (2001).
152. Li, S. & Waters, R. Escherichia coli strains lacking protein HU are UV sensitive due to a role for HU in homologous recombination. *J. Bacteriol.* **180**, 3750–6 (1998).
153. Boubrik, F. & Rouviere-Yaniv, J. Increased sensitivity to gamma irradiation in bacteria lacking protein HU. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3958–62 (1995).
154. Kamashev, D. *et al.* Mycoplasma gallisepticum produces a histone-like protein that recognizes base mismatches in DNA. *Biochemistry* **50**, 8692–702 (2011).
155. Mukherjee, A., Bhattacharyya, G. & Grove, A. The C-terminal domain of HU-related histone-like protein Hlp from Mycobacterium smegmatis mediates DNA end-joining. *Biochemistry* **47**, 8744–53 (2008).
156. Hashimoto, M., Imhoff, B., Ali, M. M. & Kow, Y. W. HU protein of Escherichia coli has a role in the repair of closely opposed lesions in DNA. *J. Biol. Chem.* **278**, 28501–7 (2003).
157. Rouvière-Yaniv, J. & Kjeldgaard, N. O. Native Escherichia coli HU protein is a heterotypic dimer. *FEBS Lett.* **106**, 297–300 (1979).
158. Oberto, J. & Rouviere-Yaniv, J. Serratia marcescens contains a heterodimeric HU protein like Escherichia coli and Salmonella typhimurium. *J. Bacteriol.* **178**, 293–7 (1996).
159. Cases, I. & de Lorenzo, V. The genomes of Pseudomonas encode a third HU

- protein. *Microbiology* **148**, 1243–5 (2002).
160. Bartels, F., Fernández, S., Holtel, A., Timmis, K. N. & de Lorenzo, V. The essential HupB and HupN proteins of *Pseudomonas putida* provide redundant and nonspecific DNA-bending functions. *J. Biol. Chem.* **276**, 16641–8 (2001).
 161. Kano, Y. *et al.* Molecular cloning and nucleotide sequence of the HU-1 gene of *Escherichia coli*. *Mol. Gen. Genet.* **201**, 360–2 (1985).
 162. Kano, Y., Osato, K., Wada, M. & Imamoto, F. Cloning and sequencing of the HU-2 gene of *Escherichia coli*. *Mol. Gen. Genet.* **209**, 408–10 (1987).
 163. Claret, L. & Rouviere-Yaniv, J. Regulation of HU alpha and HU beta by CRP and FIS in *Escherichia coli*. *J. Mol. Biol.* **263**, 126–39 (1996).
 164. Claret, L. & Rouviere-Yaniv, J. Variation in HU composition during growth of *Escherichia coli*: the heterodimer is required for long term survival. *J. Mol. Biol.* **273**, 93–104 (1997).
 165. Jones, P. G., Krah, R., Tafuri, S. R. & Wolffe, A. P. DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. *J. Bacteriol.* **174**, 5798–802 (1992).
 166. Ramstein, J. *et al.* Evidence of a Thermal Unfolding Dimeric Intermediate for the *Escherichia coli* Histone-like HU Proteins: Thermodynamics and Structure. *J. Mol. Biol.* **331**, 101–121 (2003).
 167. Garnier, N. *et al.* An alternative flexible conformation of the *E. coli* HU β_2 protein: structural, dynamics, and functional aspects. *Eur. Biophys. J.* **40**, 117–29 (2011).
 168. Coste, F., Hervouet, N., Oberto, J., Zelwer, C. & Castaing, B. Crystallization and preliminary X-ray diffraction analysis of the homodimeric form α_2 of the HU protein from *Escherichia coli*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 1952–1954 (1999).
 169. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–32 (2010).
 170. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–42 (2011).
 171. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D. Biol. Crystallogr.* **63**, 32–41 (2007).

172. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–21 (2010).
173. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126–32 (2004).
174. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
175. Ikura, M., Kay, L. E. & Bax, A. A novel approach for sequential assignment of ¹H, ¹³C, and ¹⁵N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **29**, 4659–67 (1990).
176. Schanda, P., Van Melckebeke, H. & Brutscher, B. Speeding up three-dimensional protein NMR experiments to a few minutes. *J. Am. Chem. Soc.* **128**, 9042–3 (2006).
177. Lescop, E., Schanda, P. & Brutscher, B. A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–9 (2007).
178. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–93 (1995).
179. Kazimierczuk, K. & Orekhov, V. Y. Accelerated NMR spectroscopy by using compressed sensing. *Angew. Chem. Int. Ed. Engl.* **50**, 5556–9 (2011).
180. Orekhov, V. Y. & Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 271–92 (2011).
181. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–96 (2005).
182. Rieping, W. *et al.* ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–2 (2007).
183. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–33 (2007).

184. Cheung, M.-S., Maguire, M. L., Stevens, T. J. & Broadhurst, R. W. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J. Magn. Reson.* **202**, 223–33 (2010).
185. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of ¹⁵N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–34 (1994).
186. Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123**, 11341–52 (2001).
187. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.4.1.
188. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).
189. Liang, J., Edelsbrunner, H., Fu, P., Sudhakar, P. V & Subramaniam, S. Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins. *Proteins* **33**, 18–29 (1998).
190. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
191. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–91 (2009).
192. Pellegrini, O., Oberto, J., Pinson, V., Wery, M. & Rouviere-Yaniv, J. Overproduction and improved strategies to purify the threonine forms of nuclease-free HU protein. *Biochimie* **82**, 693–704 (2000).
193. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–54 (2013).
194. Monticelli, L. *et al.* The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **4**, 819–834 (2008).
195. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–24 (2007).

196. Vis, H., Dobson, C. M. & Robinson, C. V. Selective association of protein molecules followed by mass spectrometry. *Protein Sci.* **8**, 1368–70 (1999).
197. Gabant, G. & Cadene, M. Mass spectrometry of full-length integral membrane proteins to define functionally relevant structural features. *Methods* **46**, 54–61 (2008).
198. Strohal, M., Hassman, M., Kosata, B. & Kodíček, M. mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* **22**, 905–8 (2008).
199. Strohal, M., Kavan, D., Novák, P., Volný, M. & Havlíček, V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal. Chem.* **82**, 4648–51 (2010).
200. Huisman, O. *et al.* Multiple defects in *Escherichia coli* mutants lacking HU protein. *J. Bacteriol.* **171**, 3704–12 (1989).
201. Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J. & Tan, S. Comparison of affinity tags for protein purification. *Protein Expr. Purif.* **41**, 98–105 (2005).
202. Gasparian, M. E., Bychkov, M. L., Dolgikh, D. A. & Kirpichnikov, M. P. Strategy for improvement of enteropeptidase efficiency in tag removal processes. *Protein Expr. Purif.* **79**, 191–6 (2011).
203. Chial, H. J., Thompson, H. B. & Splittgerber, A. G. A spectral study of the charge forms of Coomassie blue G. *Anal. Biochem.* **209**, 258–66 (1993).
204. Le Bihan, Y.-V. *et al.* 5-Hydroxy-5-methylhydantoin DNA lesion, a molecular trap for DNA glycosylases. *Nucleic Acids Res.* (2011). doi:10.1093/nar/gkr215
205. Marchuk, D., Drumm, M., Saulino, A. & Collins, F. S. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**, 1154 (1991).
206. Walhout, A. J. *et al.* GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* **328**, 575–92 (2000).
207. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA cloning using in vitro site-specific recombination. *Genome Res.* **10**, 1788–95 (2000).

208. Muyrers, J. P., Zhang, Y. & Stewart, A. F. Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA. *Trends Biochem. Sci.* **26**, 325–31 (2001).
209. Danthinne, X. & Imperiale, M. J. Production of first generation adenovirus vectors: a review. *Gene Ther.* **7**, 1707–14 (2000).
210. Bubeck, P., Winkler, M. & Bautsch, W. Rapid cloning by homologous recombination in vivo. *Nucleic Acids Res.* **21**, 3601–2 (1993).
211. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **4**, 251–6 (2007).
212. Van den Ent, F. & Löwe, J. RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74 (2006).
213. Andera, L., Schneider, G. J. & Geiduschek, E. P. An analysis of subunit exchange in the dimeric DNA-binding and DNA-bending protein, TF1. *Biochimie* **76**, 933–40 (1994).
214. Durney, M. a *et al.* An alternate conformation of the hyperthermostable HU protein from *Thermotoga maritima* has unexpectedly high flexibility. *FEBS Lett.* **563**, 49–54 (2004).
215. Kawamura, S., Tanaka, I., Yamasaki, N. & Kimura, M. Contribution of a salt bridge to the thermostability of DNA binding protein HU from *Bacillus stearothermophilus* determined by site-directed mutagenesis. *J. Biochem.* **121**, 448–55 (1997).
216. Kawamura, S. *et al.* Investigation of the structural basis for thermostability of DNA-binding protein HU from *Bacillus stearothermophilus*. *J. Biol. Chem.* **273**, 19982–7 (1998).
217. Guo, F. & Adhya, S. Spiral structure of *Escherichia coli* HU α provides foundation for DNA supercoiling. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 4309–14 (2007).
218. Bhowmick, T. *et al.* Targeting *Mycobacterium tuberculosis* nucleoid-associated protein HU with structure-based inhibitors. *Nat. Commun.* **5**, 4124 (2014).
219. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–9 (2010).

220. Tilly, K., Fuhrman, J., Campbell, J. & Samuels, D. S. Isolation of *Borrelia burgdorferi* genes encoding homologues of DNA-binding protein HU and ribosomal protein S20. *Microbiology* **142** (Pt 9, 2471–9 (1996).
221. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–97 (2007).
222. Bueno, M., Cremades, N., Neira, J. L. & Sancho, J. Filling small, empty protein cavities: structural and energetic consequences. *J. Mol. Biol.* **358**, 701–12 (2006).
223. Chakrabarty, A., Kortemme, T. & Baldwin, R. L. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**, 843–52 (1994).
224. Cierpicki, T. & Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21**, 249–61 (2001).
225. Baxter, N. J. & Williamson, M. P. Temperature dependence of ¹H chemical shifts in proteins. *J. Biomol. NMR* **9**, 359–69 (1997).
226. Baxter, N. J., Hosszu, L. L., Waltho, J. P. & Williamson, M. P. Characterisation of low free-energy excited states of folded proteins. *J. Mol. Biol.* **284**, 1625–39 (1998).
227. Korzhnev, D. M., Religa, T. L. & Kay, L. E. Transiently populated intermediate functions as a branching point of the FF domain folding pathway. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17777–82 (2012).
228. Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647–51 (1992).
229. Saitô, H. Conformation-dependent ¹³C chemical shifts: A new means of conformational characterization as obtained by high-resolution solid-state ¹³C NMR. *Magn. Reson. Chem.* **24**, 835–852 (1986).
230. Wishart, D. S., Sykes, B. D. & Richards, F. M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–33 (1991).
231. Spera, S. & Bax, A. Empirical correlation between protein backbone

- conformation and C.alpha. and C.beta. ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492 (1991).
232. Marrink, S. J., de Vries, A. H. & Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **108**, 750–760 (2004).
233. Kannan, A., Camilloni, C., Sahakyan, A. B., Cavalli, A. & Vendruscolo, M. A conformational ensemble derived using NMR methyl chemical shifts reveals a mechanical clamping transition that gates the binding of the HU protein to DNA. *J. Am. Chem. Soc.* **136**, 2204–7 (2014).
234. Neudecker, P. *et al.* Structure of an intermediate state in protein folding and aggregation. *Science* **336**, 362–6 (2012).
235. Hilton, G. R. & Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *J. R. Soc. Interface* **9**, 801–16 (2012).
236. Uetrecht, C., Rose, R. J., van Duijn, E., Lorenzen, K. & Heck, A. J. R. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem. Soc. Rev.* **39**, 1633–55 (2010).
237. Bach, R. D., Dmitrenko, O. & Thorpe, C. Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J. Org. Chem.* **73**, 12–21 (2008).
238. Haynes, W. M. *CRC Handbook of Chemistry and Physics*. (CRC Press, 2011).
239. Pegoraro, S., Fiori, S., Cramer, J., Rudolph-Böhner, S. & Moroder, L. The disulfide-coupled folding pathway of apamin as derived from diselenide-quenched analogs and intermediates. *Protein Sci.* **8**, 1605–13 (1999).
240. Cadene, M. & Chait, B. T. A robust, detergent-friendly method for mass spectrometric analysis of integral membrane proteins. *Anal. Chem.* **72**, 5655–8 (2000).
241. Amici, A., Levine, R. L., Tsai, L. & Stadtman, E. R. Conversion of amino acid residues in proteins and amino acid homopolymers to carbonyl derivatives by metal-catalyzed oxidation reactions. *J. Biol. Chem.* **264**, 3341–6 (1989).
242. GARRISON, W. M., JAYKO, M. E. & BENNETT, W. Radiation-induced oxidation of protein in aqueous solution. *Radiat. Res.* **16**, 483–502 (1962).

243. Requena, J. R., Chao, C. C., Levine, R. L. & Stadtman, E. R. Glutamic and aminoadipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 69–74 (2001).
244. Stadtman, E. R. Oxidation of free amino acids and amino acid residues in proteins by radiolysis and by metal-catalyzed reactions. *Annu. Rev. Biochem.* **62**, 797–821 (1993).
245. Uchida, K., Mitsui, M. & Kawakishi, S. Monooxygenation of N-acetylhistamine mediated by L-ascorbate. *Biochim. Biophys. Acta* **991**, 377–9 (1989).
246. Uchida, K. & Kawakishi, S. Site-specific oxidation of angiotensin I by copper(II) and L-ascorbate: conversion of histidine residues to 2-imidazolones. *Arch. Biochem. Biophys.* **283**, 20–6 (1990).
247. Liu, Y. & Eisenberg, D. 3D domain swapping: as domains continue to swap. *Protein Sci.* **11**, 1285–99 (2002).
248. Bennett, M. J., Schlunegger, M. P. & Eisenberg, D. 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4**, 2455–68 (1995).
249. Placek, B. J. & Gloss, L. M. Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: the N-terminal tails affect the transition state between a dimeric intermediate and the native dimer. *J. Mol. Biol.* **345**, 827–36 (2005).
250. Hobart, S. A., Meinhold, D. W., Osuna, R. & Colón, W. From two-state to three-state: the effect of the P61A mutation on the dynamics and stability of the factor for inversion stimulation results in an altered equilibrium denaturation mechanism. *Biochemistry* **41**, 13744–54 (2002).
251. Liu, Z. *et al.* Temperature-induced partially unfolded state of hUBF HMG Box-5: conformational and dynamic investigations of the Box-5 thermal intermediate ensemble. *Proteins* **77**, 432–47 (2009).
252. Lunsford, R. D., Nguyen, N. & London, J. DNA-binding activities in *Streptococcus gordonii*: identification of a receptor-nickase and a histonelike protein. *Curr. Microbiol.* **32**, 95–100 (1996).
253. Goodman, S. D. *et al.* Biofilms can be dispersed by focusing the immune system on a common family of bacterial nucleoid-associated proteins. *Mucosal*

Immunol. **4**, 625–37 (2011).

254. Brandstetter, K. A., Jurcisek, J. A., Goodman, S. D., Bakaletz, L. O. & Das, S. Antibodies directed against integration host factor mediate biofilm clearance from Nasopore. *Laryngoscope* **123**, 2626–32 (2013).
255. Dalhus, B., Laerdahl, J. K., Backe, P. H. & Bjørås, M. DNA base repair-- recognition and initiation of catalysis. *FEMS Microbiol. Rev.* **33**, 1044–78 (2009).
256. Campalans, A. *et al.* XRCC1 interactions with multiple DNA glycosylases: a model for its recruitment to base excision repair. *DNA Repair (Amst)*. **4**, 826–35 (2005).
257. Hill, J. W., Hazra, T. K., Izumi, T. & Mitra, S. Stimulation of human 8-oxoguanine-DNA glycosylase by AP-endonuclease: potential coordination of the initial steps in base excision repair. *Nucleic Acids Res.* **29**, 430–8 (2001).
258. Vidal, A. E., Boiteux, S., Hickson, I. D. & Radicella, J. P. XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. *EMBO J.* **20**, 6530–9 (2001).
259. Vidal, A. E., Hickson, I. D., Boiteux, S. & Radicella, J. P. Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step. *Nucleic Acids Res.* **29**, 1285–92 (2001).
260. Marintchev, A. *et al.* Solution structure of the single-strand break repair protein XRCC1 N-terminal domain. *Nat. Struct. Biol.* **6**, 884–93 (1999).
261. Ghosh, S., Mallick, B. & Nagaraja, V. Direct regulation of topoisomerase activity by a nucleoid-associated protein. *Nucleic Acids Res.* (2014). doi:10.1093/nar/gku804
262. Bates, A. D. & Maxwell, A. *DNA Topology*. 220 (OUP Oxford, 2005). at <<http://books.google.com/books?id=D64BXcRvMscC&pgis=1>>
263. Wang, X., Montero Llopis, P. & Rudner, D. Z. Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.* **14**, 191–203 (2013).
264. Liao, J.-H. *et al.* Binding and cleavage of E. coli HUbeta by the E. coli Lon protease. *Biophys. J.* **98**, 129–37 (2010).

265. Giangrossi, M., Giuliadori, A. M., Gualerzi, C. O. & Pon, C. L. Selective expression of the beta-subunit of nucleoid-associated protein HU during cold shock in *Escherichia coli*. *Mol. Microbiol.* **44**, 205–16 (2002).
266. Driessen, R. P. C. & Dame, R. T. Nucleoid-associated proteins in Crenarchaea. *Biochem. Soc. Trans.* **39**, 116–21 (2011).
267. J, R. *et al.* Archaeal histones: structures, stability and DNA binding. (2004). at <<http://www.biochemsoctrans.org/bst/032/0227/bst0320227.htm>>

Annexes

Tableau 21: Compositon des 48 solutions de la boîte 1 du kit JCSG-plus™ de chez Molecular

Dimensions

Tube No.	Salt	Buffer	pH	Precipitant
1.1	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	50 % v/v PEG 400
1.2	None	0.1 M sodium citrate	5.5	20 % w/v PEG 3000
1.3	0.2 M di-ammonium hydrogen citrate	None	-	20 % w/v PEG 3350
1.4	0.02 M calcium chloride	0.1 M sodium acetate	4.6	30 % v/v MPD
1.5	0.2 M magnesium formate	None	-	20 % w/v PEG 3350
1.6	0.2 M lithium sulfate	0.1 M phosphate/citrate	4.2	20 % w/v PEG 1000
1.7	None	0.1 M CHES	9.5	20 % w/v PEG 8000
1.8	0.2 M ammonium formate	None	-	20 % w/v PEG 3350
1.9	0.2 M ammonium chloride	None	-	20 % w/v PEG 3350
1.10	0.2 M potassium formate	None	-	20 % w/v PEG 3350
1.11	0.2 M ammonium dihydrogen phosphate	0.1 M Tris	8.5	50 % v/v MPD
1.12	0.2 M potassium nitrate	None	-	20 % w/v PEG 3350
1.13	None	0.1 M citrate	4.0	0.8 M ammonium sulfate
1.14	0.2 M sodium thiocyanate	None	-	20 % w/v PEG 3350
1.15	None	0.1 M Bicine	9.0	20 % w/v PEG 6000
1.16	None	0.1 M HEPES	7.5	10 % w/v PEG 8000/ 8 % v/v Ethylene glycol
1.17	None	0.1 M sodium cacodylate	6.5	40 % v/v MPD/ 5 % w/v PEG 8000
1.18	None	0.1 M phosphate/citrate	4.2	40 % v/v Ethanol/ 5 % w/v PEG 1000
1.19	None	0.1 M sodium acetate	4.6	8 % w/v PEG 4000
1.20	0.2 M magnesium chloride	0.1 M Tris	7.0	10 % w/v PEG 8000
1.21	None	0.1 M citrate	5.0	20 % w/v PEG 6000
1.22	0.2 M magnesium chloride	0.1 M sodium cacodylate	6.5	50 % v/v PEG 200
1.23	None	None	6.5	1.6 M tri-sodium citrate
1.24	0.2 M tri-potassium citrate	None	-	20 % w/v PEG 3350
1.25	0.2 M sodium chloride	0.1 M phosphate/citrate	4.2	20 % w/v PEG 8000
1.26	1.0 M lithium chloride	0.1 M Na citrate	4.0	20 % w/v PEG 6000
1.27	0.2 M ammonium nitrate	None	-	20 % w/v PEG 3350
1.28	None	0.1 M Na HEPES	7.0	10 % w/v PEG 6000
1.29	None	0.1 M Na HEPES	7.5	0.8 M sodium dihydrogen phosphate 0.8 M potassium dihydrogen phosphate
1.30	None	0.1 M phosphate/citrate	4.2	40 % v/v PEG 300
1.31	0.2 M zinc acetate	0.1 M sodium acetate	4.5	10 % w/v PEG 3000
1.32	None	0.1 M Tris	8.5	20 % v/v Ethanol
1.33	None	0.1 M Na/K phosphate	6.2	25 % v/v 1,2-propanediol 10 % v/v Glycerol
1.34	None	0.1 M Bicine	9.0	10 % w/v PEG 20,000/ 2% v/v Dioxane
1.35	None	0.1 M sodium acetate	4.6	2.0 M ammonium sulfate
1.36	None	None	-	10 % w/v PEG 1000/ 10 % w/v PEG 8000
1.37	None	None	-	24 % w/v PEG 1500/ 20 % v/v Glycerol
1.38	0.2 M magnesium chloride	0.1 M Na HEPES	7.5	30 % v/v PEG 400
1.39	0.2 M sodium chloride	0.1 M Na/K phosphate	6.2	50 % v/v PEG 200
1.40	0.2 M lithium sulfate	0.1 M sodium acetate	4.5	30 % w/v PEG 8000
1.41	None	0.1 M HEPES	7.5	70 % v/v MPD
1.42	0.2 M magnesium chloride	0.1 M Tris	8.5	20 % w/v PEG 8000
1.43	0.2 M lithium sulfate	0.1 M Tris	8.5	40 % v/v PEG 400
1.44	None	0.1 M Tris	8.0	40 % v/v MPD
1.45	0.17 M ammonium sulfate	None	-	25.5 % w/v PEG 4000/ 15 % v/v Glycerol
1.46	0.2 M calcium acetate	0.1 M sodium cacodylate	6.5	40 % v/v PEG 300
1.47	0.14 M calcium chloride	0.07 M sodium acetate	4.6	14 % v/v 2-propanol/ 30 % v/v Glycerol
1.48	0.04 M potassium dihydrogen phosphate	None	-	16 % w/v PEG 8000/ 20 % v/v Glycerol

Tableau 22: Compositon des 48 solutions de la boîte 2 du kit JCSG-plus™ de chez Molecular**Dimensions**

Tube	Salt	Buffer	pH	Precipitant
2.1	None	0.1 M sodium cacodylate	6.5	1.0 M tri-sodium citrate
2.2	0.2 M sodium chloride	0.1 M sodium cacodylate	6.5	2.0 M ammonium sulfate
2.3	0.2 M sodium chloride	0.1 M HEPES	7.5	10 % v/v 2-propanol
2.4	0.2 M lithium sulfate	0.1 M Tris	8.5	1.26 M ammonium sulfate
2.5	None	0.1 M CAPS	10.5	40 % v/v MPD
2.6	0.2 M zinc acetate	0.1 M imidazole	8.0	20 % w/v PEG 3000
2.7	0.2 M zinc acetate	0.1 M sodium cacodylate	6.5	10 % v/v 2-propanol
2.8	None	0.1 M sodium acetate	4.5	1.0 M di-ammonium hydrogen phosphate
2.9	None	0.1 M MES	6.5	1.6 M magnesium sulfate
2.10	None	0.1 M Bicine	9.0	10 % w/v PEG 6000
2.11	0.16 M calcium acetate	0.08 M sodium cacodylate	6.5	14.4 % w/v PEG 8000/ 20 % v/v glycerol
2.12	None	0.1 M imidazole	8.0	10 % w/v PEG 8000
2.13	0.05 M caesium chloride	0.1 M MES	6.5	30 % v/v Jeffamine M-600
2.14	None	0.1 M Na Citrate	5.0	3.2 M ammonium sulfate
2.15	None	0.1 M Tris	8.0	20 % v/v MPD
2.16	None	0.1 M HEPES	7.5	20 % v/v Jeffamine M-600
2.17	0.2 M magnesium chloride	0.1 M Tris	8.5	50 % v/v ethylene glycol
2.18	None	0.1 M Bicine	9.0	10 % v/v MPD
2.19	None	None	7.0	0.8 M succinic acid
2.20	None	None	7.0	2.1 M DL-malic acid
2.21	None	None	7.0	2.4 M sodium malonate
2.22	1.1 M sodium malonate	0.1 M HEPES	7.0	0.5 % v/v Jeffamine ED-2001
2.23	1.0 M succinic acid	0.1 M HEPES	7.0	1 % w/v PEG 2000 MME
2.24	None	0.1 M HEPES	7.0	30 % v/v Jeffamine M-600
2.25	None	0.1 M HEPES	7.0	30 % v/v Jeffamine ED-2001
2.26	0.02 M magnesium chloride	0.1 M HEPES	7.5	22 % w/v polyacrylic acid 5100 sodium salt
2.27	0.01 M cobalt chloride	0.1 M Tris	8.5	20 % w/v polyvinylpyrrolidone K15
2.28	0.2 M tri-methylamine N-oxide	0.1 M Tris	8.5	20 % w/v PEG 2000 MME
2.29	0.005 M cobalt chloride	0.1 M HEPES	7.5	12 % w/v PEG 3350
	0.005 M cadmium chloride			
	0.005 M magnesium chloride			
	0.005 M nickel chloride			
2.30	0.2 M sodium malonate	None	7.0	20 % w/v PEG 3350
2.31	0.1 M succinic acid	None	7.0	15 % w/v PEG 3350
2.32	0.15 M DL - malic acid	None	7.0	20 % w/v PEG 3350
2.33	0.1 M potassium thiocyanate	None	-	30 % w/v PEG 2000 MME
2.34	0.15 M potassium bromide	None	-	30 % w/v PEG 2000 MME
2.35	None	0.1 M Bis Tris	5.5	2.0 M ammonium sulfate
2.36	None	0.1 M Bis Tris	5.5	3.0 M sodium chloride
2.37	None	0.1 M Bis Tris	5.5	0.3 M magnesium formate
2.38	1.0 M ammonium sulfate	0.1 M Bis Tris	5.5	1 % w/v PEG 3350
2.39	None	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.40	0.2 M calcium chloride	0.1 M Bis Tris	5.5	45 % v/v MPD
2.41	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	45 % v/v MPD
2.42	0.1 M ammonium acetate	0.1 M Bis Tris	5.5	17 % w/v PEG 10000
2.43	0.2 M ammonium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.44	0.2 M sodium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.45	0.2 M lithium sulfate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.46	0.2 M ammonium acetate	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.47	0.2 M magnesium chloride	0.1 M Bis Tris	5.5	25 % w/v PEG 3350
2.48	0.2 M ammonium acetate	0.1 M HEPES	7.5	45 % v/v MPD

Tableau 23: Composition des 50 solution du kit basique de cristallisation de chez Sigma-Aldrich

1. Ca-chloride 0.02M, Na-acetate (pH 4.6) 0.1M, 2-Methyl-2,4-pentanediol 30%
2. K-,Na-tartrate 0.4M
3. NH₄-dihydrogenphosphate 0.4M
4. TRIS-HCl (pH 8.5) 0.1M, NH₄-sulfate 2.0M
5. Na-citrate 0.2M, HEPES Na-salt (pH 7.5) 0.1M, 2-Methyl-2,4-pentanediol 30%
6. Mg-chloride 0.2M, TRIS-HCl (pH 8.5) 0.1M, PEG 4000 30%
7. Na-cacodylate (pH 6.5) 0.1M, Na-acetate 1.4M
8. Na-citrate 0.2M, Na-cacodylate (pH 6.5) 0.1M, 2-Propanol 30%
9. NH₄-acetate 0.2M, Na-citrate (pH 5.6) 0.1M, PEG 4000 30%
10. NH₄-acetate 0.2M, Na-acetate (pH 4.6) 0.1M, PEG 4000 30%
11. Na-citrate (pH 5.6) 0.1M, NH₄-dihydrogenphosphate 1.0M
12. Mg-chloride 0.2M, HEPES Na-salt (pH 7.5) 0.1M, 2-Propanol 30%
13. Na-citrate 0.2M, TRIS-HCl (pH 8.5) 0.1M, PEG 400 30%
14. Ca-chloride 0.2M, HEPES Na-salt (pH 7.5) 0.1M, PEG 400 28%
15. NH₄-sulfate 0.2M, Na-cacodylate (pH 6.5) 0.1M, PEG 8000 30%
16. HEPES Na-salt (pH 7.5) 0.1M, Li-sulfate 1.5M
17. Li-sulfate 0.2M, TRIS-HCl (pH 8.5) 0.1M, PEG 4000 30%
18. Mg-acetate 0.2M, Na-cacodylate pH (6.5) 0.1M, PEG 8000 20%
19. NH₄-acetate 0.2M, TRIS-HCl (pH 8.5) 0.1M, 2-Propanol 30%
20. NH₄-sulfate 0.2M, Na-acetate (pH 4.6) 0.1M, PEG 4000 25%
21. Mg-acetate 0.2M, Na-cacodylate (pH 6.5) 0.1M, 2-Methyl-2,4-pentanediol 30%
22. Na-acetate 0.2M, TRIS-HCl (pH 8.5) 0.1M, PEG 4000 30%
23. Mg-chloride 0.2M, HEPES Na-salt (pH 7.5) 0.1M, PEG 400 30%
24. Ca-chloride 0.2M, Na-acetate (pH 4.6) 0.1M, 2-Propanol 20%
25. Imidazole (pH 6.5) 0.1M, Na-acetate 1M
26. NH₄-acetate 0.2M, Na-citrate (pH 5.6) 0.1M, 2-Methyl-2,4-pentanediol 30%
27. Na-citrate 0.2M, HEPES Na-salt (pH 7.5) 0.1M, 2-Propanol 20%
28. Na-acetate 0.2M, Na-cacodylate pH (6.5) 0.1M, PEG 8000 30%
29. HEPES Na-salt (pH 7.5) 0.1M, K-,Na-tartrate 0.8M
30. NH₄-sulfate 0.2M, PEG 8000 30%
31. NH₄-sulfate 0.2M, PEG 4000 30%
32. NH₄-sulfate 2M
33. Na-formiate 4M
34. Na-acetate (pH 4.6) 0.1M,Na-formiate 2M
35. HEPES Na-salt (pH 7.5) 0.1M, K-dihydrogenphosphate 0.8M, Na-dihydrogenphosphate 0.8M
36. TRIS-HCl (pH 8.5) 0.1M, PEG 8000 8%
37. Na-acetate (pH 4.6) 0.1M, PEG 4000 8%
38. HEPES Na-salt (pH 7.5) 0.1M, Na-citrate 1.4M
39. HEPES Na-salt (pH 7.5) 0.1M, PEG 400 2%, NH₄-sulfate 2.0M
40. Na-citrate (pH 5.6) 0.1M, 2-Propanol 20%, PEG 4000 20%
41. HEPES Na-salt (pH 7.5) 0.1M, 2-Propanol 10%, PEG 4000 20%
42. K-dihydrogenphosphate 0.05M, PEG 8000 20%
43. PEG 1500 30%
44. Mg-formiate 0.2M
45. Zn-acetate 0.2M, Na-cacodylate (pH 6.5) 0.1M, PEG 8000 18%
46. Ca-acetate 0.2M, Na-cacodylate (pH 6.5) 0.1M, PEG 8000 18%
47. Na-acetate (pH 4.6) 0.1M, NH₄-sulfate 2.0M
48. TRIS-HCl (pH 8.5) 0.1M, NH₄-dihydrogenphosphate 2.0M
49. Li-sulfate 1.0M, PEG 8000 2%
50. Li-sulfate 0.5M, PEG 8000 15%

Tableau 24: Oligonucléotides

N°	Dénomination	Séquence (5' → 3')	Taille(pb)
1	P28-8hisE-Nfor	GCCATCATCATCATCATCACCACCACTCTTCATATGG CCTGGTGCCGCGC	50
2	P28-8hisE-Nrev	GCGCGGCACCAGGCCATATGAAGAGTGGTGGTGGTGA TGATGATGATGGC	50
3	hupA-nde1-E_for	ACTCTTCATATGGCGATGATGATGATCGCATGAACAA GACTCAACTG	47
4	hupA-bamh1-2	GAATTCGGATCCGTTACTTAACTGCGTCTTTCAG	34
5	hupB-nde1-E_for	ACTCTTCATATGGCGATGATGATGATCGCATGAATAA ATCTCAATTG	47
6	hupB-bamh1-2	GAATTCGGATCCGTTAGTTTACCGCGTCTTTCAG	37
7	pjes-A-nco1-for	GGAGATATACCCATGGACAAGACTC	25
8	pjes-A-nco1-rev	GAGTCTTGTCCATGGGTATATCTCC	25
9	pjes-B-nco1-for	GGAGATATACCCATGGATAAATCTC	25
10	pjes-B-nco1-rev	GAGATTTATCCATGGGTATATCTCC	25
11	Pmal-c5e-hupa-for	CCTCGGGGATGACGATGACAAGATGAACAAGACTCAA CTGATTGATG	47
12	Pmal-c5e-hupa-rev	ACCTGCAGGGAATTCGGATCCTTACTTAACTGCGTCT TTCAGTGCC	46
13	Pmal-c5e-hupb-for	CCTCGGGGATGACGATGACAAGATGAATAAATCTCAA TTGATCGACAAG	50
14	Pmal-c5e-hupb-rev	ACCTGCAGGGAATTCGGATCCTTAGTTTACCGCGTCT TTCAGTGC	45
15	pmalc5e-hsla-F	CCTCGGGGATGACGATGACAAGATGGCTAACAAACAA GATC	41
16	pmalc5e-hsla-R	ACCTGCAGGGAATTCGGATCCTTATTTAACTGCGTCT TTC	40

Tableau 25: Code génétique et fréquence d'utilisation des codons chez *E. coli* Pour chaque codons est indiqué l'acide aminé correspondant et la fréquence d'utilisation de ce codon pour l'acide aminé correspondant (Données de la « codon usage database » sur la souche E.coli K12 <http://www.kazusa.or.jp/codon/>)

		2 ^{ème}													
		U			C			A			G				
1 ^{er}	U	UUU	Phe	0.57	UCU	Ser	0.11	UAU	Tyr	0.53	UGU	Cys	0.42	U	3 ^{ème}
		UUC	Phe	0.43	UCC	Ser	0.11	UAC	Tyr	0.47	UGC	Cys	0.58	C	
		UUA	Leu	0.15	UCA	Ser	0.15	UAA	stop	0.64	UGA	stop	0.36	A	
		UUG	Leu	0.12	UCG	Ser	0.16	UAG	stop	0.00	UGG	Trp	1.00	G	
	C	CUU	Leu	0.12	CCU	Pro	0.17	CAU	His	0.55	CGU	Arg	0.36	U	
		CUC	Leu	0.10	CCC	Pro	0.13	CAC	His	0.45	CGC	Arg	0.44	C	
		CUA	Leu	0.05	CCA	Pro	0.14	CAA	Gln	0.30	CGA	Arg	0.07	A	
		CUG	Leu	0.46	CCG	Pro	0.55	CAG	Gln	0.70	CGG	Arg	0.07	G	
	A	AUU	Ile	0.58	ACU	Thr	0.16	AAU	Asn	0.47	AGU	Ser	0.14	U	
		AUC	Ile	0.35	ACC	Thr	0.47	AAC	Asn	0.53	AGC	Ser	0.33	C	
		AUA	Ile	0.07	ACA	Thr	0.13	AAA	Lys	0.73	AGA	Arg	0.02	A	
		AUG	Met	1.00	ACG	Thr	0.24	AAG	Lys	0.27	AGG	Arg	0.03	G	
	G	GUU	Val	0.25	GCU	Ala	0.11	GAU	Asp	0.65	GGU	Gly	0.29	U	
		GUC	Val	0.18	GCC	Ala	0.31	GAC	Asp	0.35	GGC	Gly	0.46	C	
		GUA	Val	0.17	GCA	Ala	0.21	GAA	Glu	0.70	GGA	Gly	0.13	A	
		GUG	Val	0.40	GCG	Ala	0.38	GAG	Glu	0.30	GGG	Gly	0.12	G	

Tableau 26: Ajustement linéaire des coefficients de températures des protons H_N Équation de fit : $y = A.x + B$. A correspondant au coefficient de température (ppb/K).

Résidu	A	B	R ²	Résidu	A	B	R ²
K3	-2,31	9,13	0,9928	Q43	-1,13	9,09	0,9748
T4	-3,26	8,84	0,9967	L44	-5,48	8,86	0,4271
Q5	-7,33	1,05	0,9984	V45	-7,76	10,6	0,9975
L6	-4,25	9,63	0,9973	F47	0,17	8,29	0,3400
I7	-4,97	9,64	0,9943	G48	-0,96	8,80	0,9581
D8	-0,61	7,52	0,9874	T49	-1,26	8,23	0,9612
V9	1,69	7,23	0,9970	F50	-2,61	10,1	0,9958
I10	-1,78	9,15	0,8020	K51	-3,68	10,5	0,9951
	0,31 / -2,98	8,52 / 9,53	0,56 / 0,95	V52	-4,08	9,72	0,9981
A11	-2,92	9,35	0,9951	N53	-2,65	9,78	0,9793
Q12	1,61	6,97	0,9785	H54	-2,50	9,70	0,6678
K13	-4,91	1,03	0,9969	R55	-7,19	11,0	0,9841
A14	-3,40	9,80	0,9960	A56	-5,16	9,83	0,9770
Q15	0,83	7,13	0,9108	E57	-8,79	11,3	0,9976
L16	-2,96	9,25	0,9967	R58	-4,41	9,90	0,9810
S17	-4,68	10,15	0,9988	T59	-8,22	10,9	0,9982
K18	-4,42	10,37	0,9950	G60	-6,58	10,7	0,9922
T19	-4,21	9,27	0,9943	R61	-7,61	10,8	0,9969
Q20	-4,23	9,07	0,9977	N62	-7,23	11,0	0,9751
A21	-2,49	9,52	0,9868	Q64	-0,53	8,43	0,9149
K22	-1,05	8,52	0,9590	T65	0,66	7,47	0,7305
A23	-0,35	7,92	0,5938	G66	-2,43	9,24	0,9313
	0,057 / -1,08	7,80 / 8,15	0,11/0,84	K67	1,91	7,06	0,9352
A24	-2,80	9,74	0,9962	E68	-6,17	10,5	0,9977
L25	-2,44	9,31	0,9941	I69	-4,67	9,77	0,9681
E26	-1,53	9,44	0,9553	K70	-7,10	1,06	0,9991
S27	-1,56	8,86	0,9682	I71	-10,86	12,1	0,9894
T28	-1,27	8,31	0,7770	A72	-5,35	9,98	0,9963
	-1,93 / 1,26	8,51 / 7,51	0,97 / 0,74	A73	-10,4	11,7	0,9943
L29	-0,74	7,99	0,8900	A74	-3,70	9,48	0,9836
	-0,91 / 0,95	8,04 / 7,46	0,96 / 0,95	N75	-8,08	11,1	0,9956
A30	-1,04	8,51	0,9755	V76	-6,42	10,7	0,9678
A31	-0,30	8,22	0,7362	A78	-2,61	10,1	0,9959
I32	-1,18	8,30	0,9869	F79	-5,05	9,96	0,9807
T33	0,12	7,8	0,2832	V80	-3,34	9,48	0,9956
E34	0,57	8,29	0,8459	G82	-2,20	8,99	0,9813
S35	-1,26	8,23	0,9612	A84	-1,8	8,93	0,9069
L36	-1,53	8,53	0,9865	L85	-4,01	8,46	0,9862
K37	-4,19	9,75	0,9959	K86	-7,65	7,99	0,9590
E38	1,68	6,94	0,9778	D87	-5,01	10,0	0,9950
G39	-0,65	8,22	0,9237	A88	-2,82	8,56	0,9942
D40	-1,04	8,51	0,9755	V89	0,78	6,88	0,9047
A41	-4,99	9,64	0,9968	K90	1,53	6,65	0,9538
V42	-2,29	9,90	0,9800				

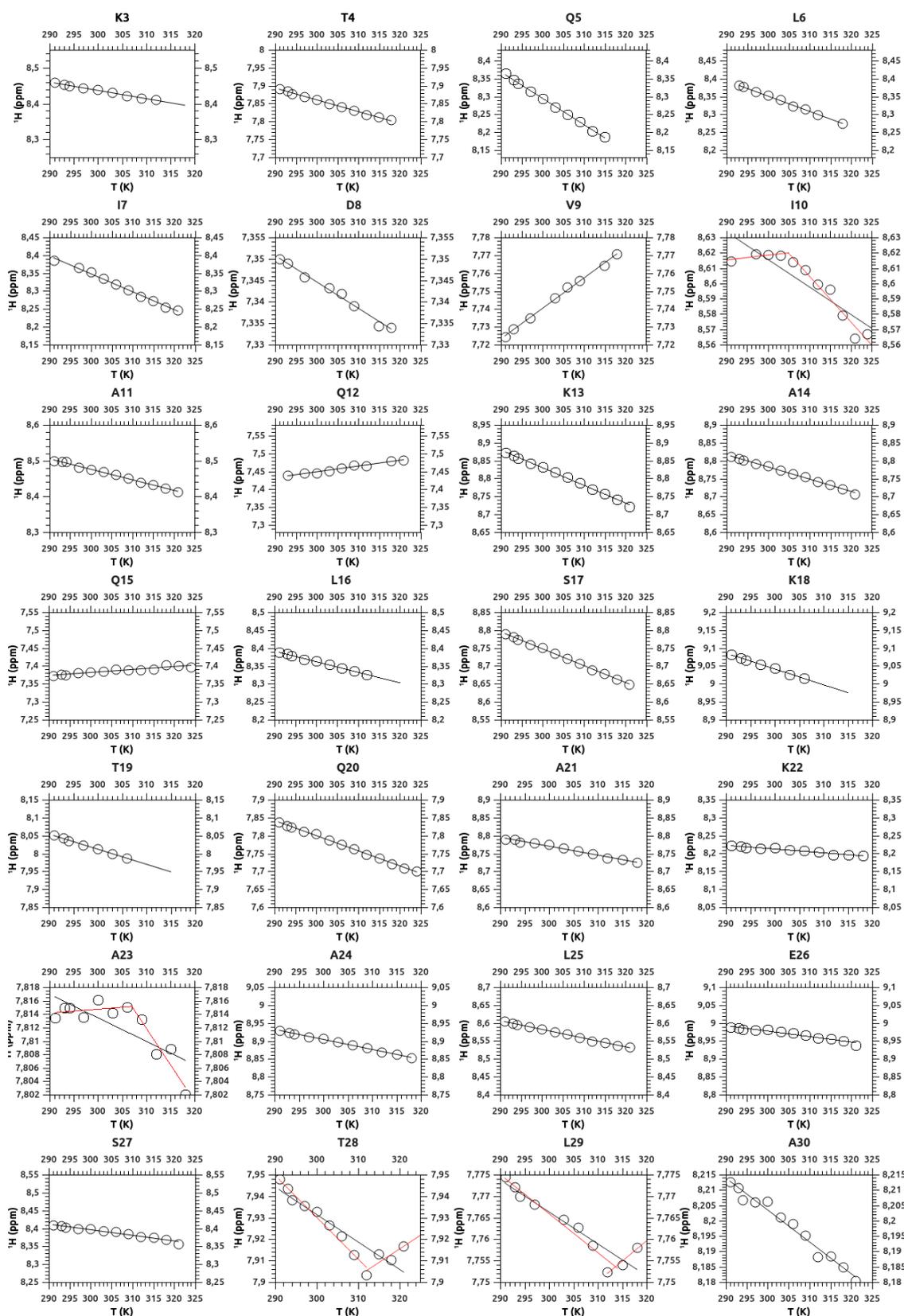


Figure 114: Représentations graphiques des déplacements chimiques H_N en fonction de la température (1/3)

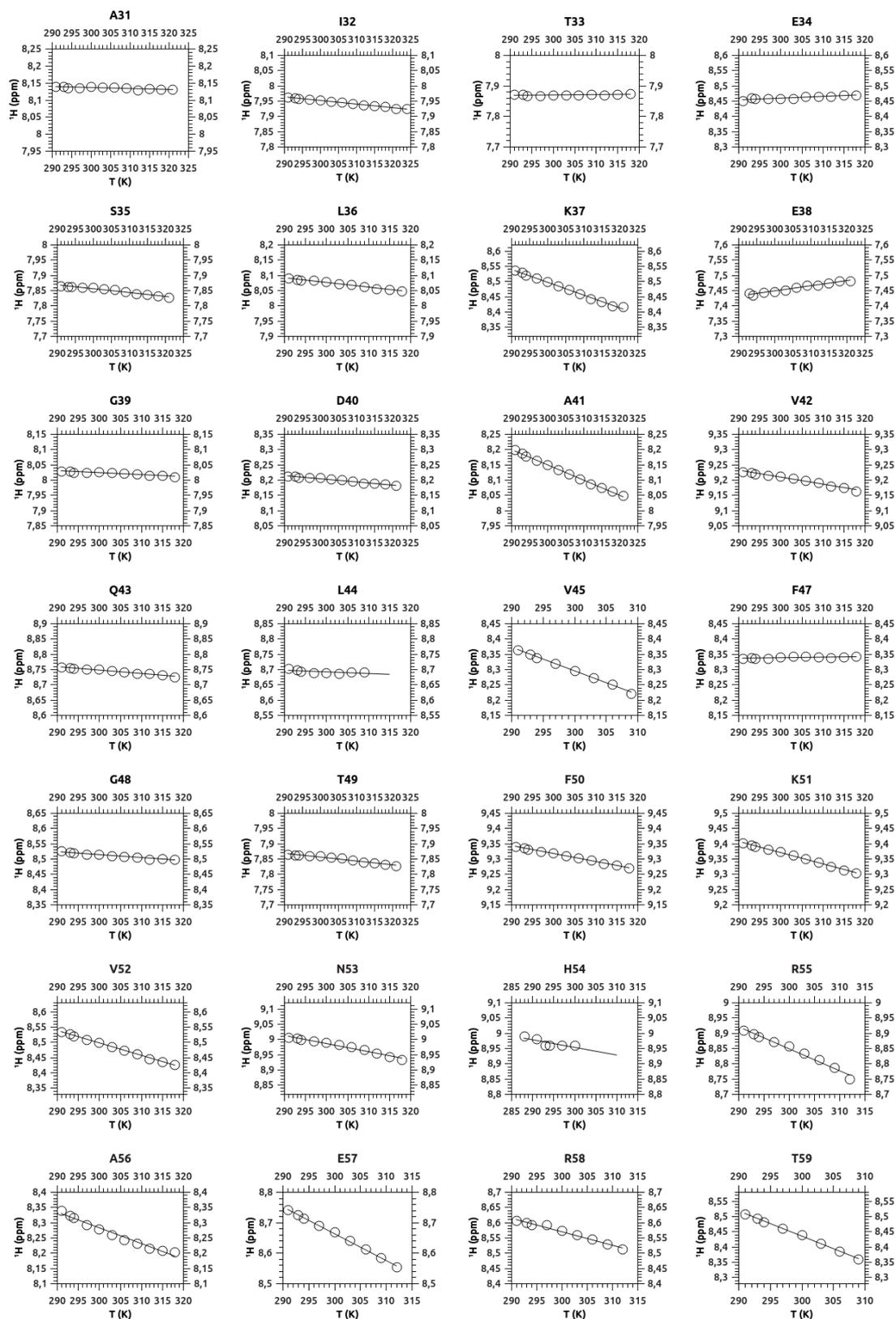


Figure 115: Représentations graphiques des déplacements chimiques HN en fonction de la température (2/3)

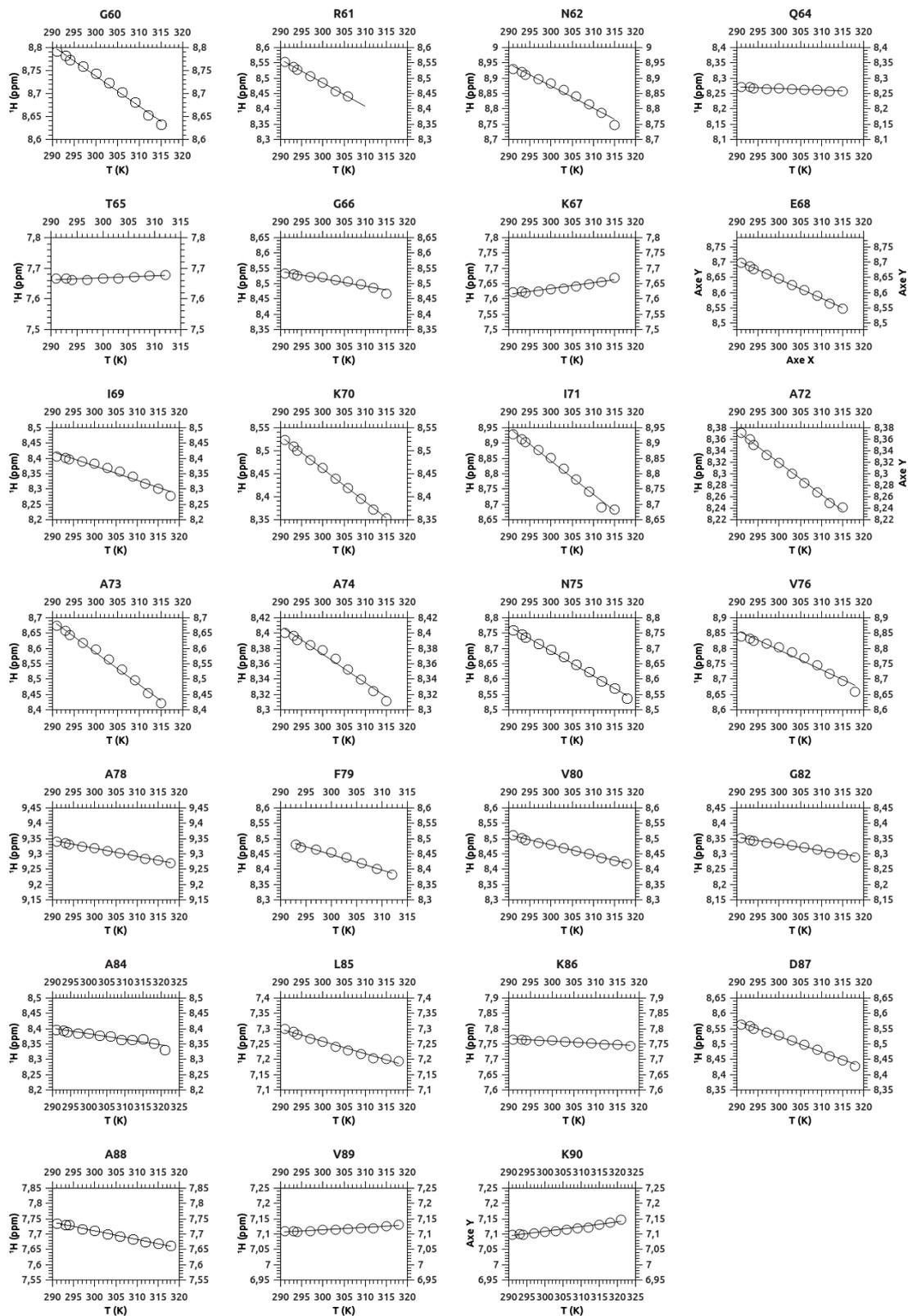


Figure 116: Représentations graphiques des déplacements chimiques HN en fonction de la température (3/3)

Rémy LE MEUR

Étude structurale du mécanisme d'échange de chaînes des dimères de la protéine HU d'*E. coli*

HU est une protéine bactérienne de type histone impliquée dans de nombreuses fonctions biologiques telles que la compaction, la transcription, la réplication et la réparation de l'ADN. Chez *E. coli*, il existe trois espèces dimériques de HU ($HU\alpha_2$, $HU\beta_2$ et $HU\alpha\beta$) ayant des rôles biologiques distincts. La formation de l'hétérodimère repose sur un échange de chaînes peptidiques entre les homodimères. Un mécanisme modèle a été proposé par Ramstein et collaborateurs (*J.M.B.* **331**, 101-121 2003) et a servi de point de départ à ce travail. Dans ce modèle, les homodimères transitent d'une conformation native (N_2) vers une conformation intermédiaire (I_2). Les homodimères sous forme I_2 s'associent ensuite dans un hétérotétramère transitoire qui se redissocie en formant des hétérodimères. L'objectif principal de ce travail de thèse a été de caractériser chaque étape du mécanisme d'échange du point de vue structural et cinétique.

Parmi les principaux résultats de ce travail, deux structures originales, $HU\beta_2$ de *E. coli* et HU de *L. lactis*, ont été obtenues par diffraction des rayons X et complètent la caractérisation structurale des protéines HU sous leur forme N_2 . Un modèle de la conformation I_2 , partiellement désordonnée, a été élaboré à partir des résultats obtenus en RMN et en simulation de dynamique moléculaire. De plus, l'existence de la conformation tétramérique a été mise en évidence en faible concentration par spectrométrie de masse en conditions natives. Des protocoles de production/purification/oxydation ont été mis au point pour l'introduction de ponts disulfure afin de stabiliser la conformation tétramérique en vue de sa caractérisation structurale.

L'ensemble des données acquises par ces différentes méthodes biophysiques affine la compréhension du mécanisme d'échange de chaînes d'un point de vue structural et cinétique et met en lumière le rôle clef de la conformation I_2 dans le contrôle de la composition des dimères de HU.

Mot-Clefs : protéine de type histone, HU, conformation intermédiaire, RMN, cristallographie des rayons X, dynamique moléculaire

Structural study of the chain exchange mechanism of *E. coli* HU dimers

HU is a histone like protein of bacteria involved in numerous biological functions such as DNA compaction, transcription, replication and repair. In *E. coli*, three HU dimers types are present ($HU\alpha_2$, $HU\beta_2$ et $HU\alpha\beta$) and show distinct biological roles. The heterodimer is formed from homodimers through a peptidic chain exchange. A model of this mechanism has been proposed by Ramstein and coworkers (*J.M.B.* **331**, 101-121 2003) and was used as a starting point for this study. In this model, homodimers undergo a conformational change from a native state (N_2) toward an intermediate state (I_2). Then, I_2 homodimers associate to form a transient heterotetramer which then dissociate into heterodimers. The main aim of this work was to characterize the structure and kinetic of each step of this mechanism.

Major results of this work include the elucidation of two original crystal structures : $HU\beta_2$ from *E. coli* and HU from *L. lactis* in N_2 states. A model of the partially disordered I_2 state has also been proposed for $HU\beta_2$ and $HU\alpha_2$, and is consistent with results obtained from both NMR and molecular dynamics experiments. In addition, the existence of a low concentration tetrameric conformation has been evidenced by native mass spectrometry experiments. A protocol of production/purification/oxydation as been developed for the introduction of disulfide bridges in order to stabilize this conformation and characterize its structure.

Together, results obtained from these different biophysical means refine our understanding of the chain exchange mechanism at the molecular level and highlight the role of the I_2 conformation in controlling HU dimers composition.

Keywords : histone-like protein, HU, intermediate state, NMR, X ray crystallography, molecular dynamics

Centre de Biophysique Moléculaire (CNRS)
Rue Charles Sadron
CS 80054
45071 ORLEANS CEDEX 2