

UNIVERSITÉ DE MONTRÉAL

Caractérisation moléculaire du rôle de facteurs accessoires ArgR et PepA au niveau de la recombinaison spécifique sur le site *cer*.

par

Jérémy DELESQUES

DÉPARTEMENT DE MICROBIOLOGIE ET D'IMMUNOLOGIE

FACULTÉ DE MÉDECINE

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)  
en microbiologie et immunologie.

Décembre 2010

©Jérémy DELESQUES, 2010

Université de Montréal  
Faculté des études supérieures

Cette thèse intitulée :

Caractérisation moléculaire du rôle de facteurs accessoires ArgR et PepA au niveau de la recombinaison spécifique sur le site *cer*.

présentée par  
Jérémy DELESQUES

a été évaluée par un jury composé des personnes suivantes :

Président-rapporteur : Dr Patrick C. HALLENBECK  
Directeur de recherche : Dr George SZATMARI  
Membre du jury : Dr Marc DROLET  
Examineur externe : Dr Chantal AUTEXIER  
Représentant du doyen de la FES: Dr Dindial RAMOTAR

Thèse acceptée le : 20/07/2011.

À tous ceux qui m'ont soutenu et qui ont cru en moi,  
À ma famille,  
À mes amis.

## RÉSUMÉ

Mon projet de recherche avait pour but de caractériser le rôle de deux protéines, ArgR et PepA, qui agissent en tant que facteurs accessoires de la recombinaison au niveau de deux sites *cer* du plasmide ColE1 présent dans la bactérie *Escherichia coli*. Ces deux protéines, couplées aux deux recombinases à tyrosine XerC et XerD, permettent la catalyse de la recombinaison site spécifique au niveau de la séquence *cer*, convertissant les multimères instables de ColE1 en monomères stables. Cette étude a principalement porté sur la région C-terminale de la protéine ArgR. Cette région de la protéine ArgR possède une séquence en acides-aminés et une structure similaire à celle de la protéine AhrC de *Bacillus subtilis*. De plus, AhrC, le répresseur de l'arginine de cette bactérie, est capable de compléter des *Escherichia coli* mutantes déficientes en ArgR. Les régions C-terminales de ces protéines, montrent une forte similarité. De précédents travaux dans notre laboratoire ont démontré que des mutants d'ArgR comprenant des mutations dans cette région, en particulier les mutants ArgR149, une version tronquée d'ArgR de 149 acides-aminés, et ArgR5aa, une version comprenant une insertion de cinq acides-aminés dans la partie C-terminale, perdaient la capacité de permettre la recombinaison au niveau de deux sites *cer* présents dans le plasmide pCS210. Malgré cette incapacité à promouvoir la réaction de recombinaison en *cer*, ces deux mutants étaient toujours capables de se lier spécifiquement à l'ADN et de réprimer une fusion *argA :: lacZ*. Dans ce travail, les versions mutantes et sauvages d'ArgR furent surexprimées en tant que protéines de fusion 6-histidine. Des analyses *crosslinking* ont montré que la version sauvage et ArgR5aa pouvaient former des hexamères *in-vitro* de manière efficace, alors qu'ArgR149 formait des multimères de plus faible poids moléculaire. Des formes tronquées d'ArgR qui comportaient 150 acides-aminés ou plus, étaient encore capables de permettre la recombinaison en *cer*. Les mutants par substitution ArgRL149A et ArgRL151A ont tous montré que les substitutions d'un seul acide-aminé au sein de cette région avaient peu d'effets sur la recombinaison en *cer*. Les expériences de *crosslinking* protéine-à-protéine ont montré que le type sauvage et les formes mutantes d'ArgR étaient capables d'interagir avec la protéine accessoire PepA, également impliquée dans la recombinaison en *cer*. Les expériences de recombinaison *in-vitro* utilisant la forme

sauvage et les formes mutantes d'ArgR combinées avec les protéines PepA, XerC et XerD purifiées, ont montré que le mutant ArgR149 ne soutenait pas la recombinaison, mais que le mutant ArgR5aa permettait la formation d'une jonction d'Holliday. Des expériences de topologie ont montré que PepA était capable de protéger l'ADN de la topoisomérase 1, et d'empêcher ArgRWt de se lier à l'ADN. Les deux mutants ArgR149 et ArgR5aa protègent aussi l'ADN avec plus de surenroulements. Quand on ajoute PepA, les profils de migration montrent un problème de liaison des deux mutants avec PepA. D'autres expériences impliquant le triplet LEL (leucine-acide glutamique-leucine) et les acides-aminés alentour devraient être réalisés dans le but de connaître l'existence d'un site de liaison potentiel pour PepA.

**Mots clés** : ArgR, PepA, partie C-terminale d'argR, recombinaison, multimères, jonction d'Holliday, liaison protéine-à-protéine.

**ABSTRACT**

My research project involved the role of two proteins, ArgR and PepA, which act as accessory factors in the ColE1 *cer* recombination system from the gram negative bacteria *Escherichia coli*. These two proteins, in addition to the tyrosine recombinases XerC and XerD, catalyze a site-specific recombination event at the *cer* sequence which converts unstable multimeric forms of ColE1 into more stable monomers. Our study mainly focused on the C-terminal end of the ArgR. This region of the ArgR protein possesses a structural and amino acid sequence similarity with the AhrC protein from *Bacillus subtilis*. Moreover, AhrC, the Arginine repressor of this bacterium, is able to complement *Escherichia coli* mutants deficient in ArgR. The C-terminal regions of these proteins, display a very high region of similarity. Previous work from our laboratory has shown that ArgR mutants with mutations in this region, especially the mutants ArgR149, a truncated 149 amino acids form of ArgR, and ArgR5aa, a form containing a five amino acid insertion in the C-terminal part, lost the ability to perform a recombination reaction at two *cer* sites in the plasmid pCS210. Despite this defect in promoting *cer* recombination, the mutants were still able to bind specifically to DNA, and to repress an *argA :: lacZ* genetic fusion. In this work, both wild type and mutant ArgR proteins were overexpressed as 6-histidine fusion proteins. *Crosslinking* analysis showed that both wild type and ArgR5aa efficiently formed hexamers in vitro, while ArgR149 formed lower molecular weight multimers. Truncated forms of ArgR that were 150 amino acids or longer, were able to support *cer* recombination. The substitution mutants between positions 149 to 151 all showed that single amino acid substitutions at this region had little effect on *cer* recombination. Protein-protein crosslinking experiments showed that wild type and mutant forms of ArgR, were able to interact with and the other accessory protein involved in *cer* recombination, PepA. In vitro recombination experiments using wild type and mutant forms of ArgR, combined with purified PepA, XerC and XerD showed that the ArgR149 mutant did not support recombination, but the ArgR5aa mutant did promote Holliday junction formation, raising the possibility that these two mutants interact differently with the Xer recombination machinery. Topology experiments showed that after adding topoisomerase 1, PepA is able to protect DNA from

topoisomerase 1, and prevent ArgRWt binding to DNA. The two mutants ArgR149 and ArgR5aa are protecting DNA with more supercoiling. When PepA is added, migration profiles with the two mutants showed a binding problem with PepA. Other experiments involving the LEL triplet (leucine-glutamic acid-leucine) and amino-acids around it should be done in order to know the existence of a possible binding site for PepA.

**Key-words:** ArgR, PepA, ArgR C-terminus part, recombination, multimers, Holliday junction, protein to protein interaction.

**TABLE DES MATIÈRES**

Dédicace.....	III
Résumé.....	IV
Abstract.....	VI
Table des matières.....	VIII
Liste des tableaux.....	XIV
Liste des figures.....	XV
Liste des abréviations.....	XX

**CHAPITRE I – INTRODUCTION**

I.1. Quelques rappels généraux.....	2
I.2. Systèmes de ségrégation et partition du chromosome.....	4
I.2.1. Ségrégation.....	4
I.2.2. Partition.....	5
I.3. Système de ségrégation plasmidique.....	6
I.3.1. Vue générale (recombinaison site spécifique).....	6
I.3.2. Autres systèmes de ségrégation plasmidique.....	9
I.3.2.1. Généralités.....	9
I.3.2.2. Mécanismes de partition.....	11
I.3.2.3. L'accoutumance génétique.....	14
I.3.2.3.1. Le système protéique classique.....	14



I.3.2.3.2. Systèmes d'accoutumance avec ARN anti-sens.....	15
I.3.3. Le plasmide ColE1.....	17
I.3.4. La séquence <i>cer</i> .....	18
I.4. Les recombinases à tyrosine XerC et XerD.....	19
I.4.1. Généralités.....	19
I.4.2. recombinaison en <i>dif</i> .....	24
I.4.3. recombinaison en <i>psi</i> .....	28
I.4.4. recombinaison en <i>cer</i> .....	29
I.4.5. recombinaison en <i>mwr</i> et <i>nmr</i> .....	30
I.4.6. XerC.....	32
I.4.7. XerD.....	33
I.4.8. XerC et XerD et système d'intégration de <i>Vibrio cholerae</i> .....	35
I.5. La protéine FtsK.....	42
I.6. La protéine ArcA.....	47
I.7. La protéine PepA.....	50
I.7.1. Découverte.....	50
I.7.2. Caractéristiques.....	51
I.7.3. <i>CarAB</i> .....	55
I.8. La protéines ArgR.....	58
I.8.1. Découverte.....	58
I.8.2. Description de la protéine.....	59
I.8.2.1. Partie Nt.....	60
I.8.2.2. Partie Ct.....	60
I.8.3. Fonction.....	61
I.9. La protéine AhrC.....	70

I.10. La protéine ArgR de <i>Bacillus stearothermophilus</i> .....	74
I.11. La protéine ArgR de <i>Mycobacterium tuberculosis</i> .....	77
I.12. Modèles d'interaction ArgR-PepA dans le système Xer/ <i>cer</i> .....	79
I.13. Projet de thèse.....	89
CHAPITRE II - MATÉRIEL ET MÉTHODES.....	92
II.1. Le matériel.....	92
II.1.1.Plasmides.....	92
II.1.2.Souches d' <i>Escherichia coli</i> utilisées.....	94
II.1.3.Milieu de culture Luria-Bertani (LB).....	96
II.1.4.Liste des antibiotiques les plus utilisés.....	96
II.1.5. Géloses TAXI.....	96
II.1.6.Milieu MacConkey Agar.....	97
II.1.7.Fabrication de milieu auto-inductible.....	97
II.1.8.Fabrication du tampon pour cellules compétentes.....	98
II.1.9.Tampon de liaison pour retard sur gel.....	98
II.1.10.Tampons TM.....	98
II.1.11.Tampons pour la purification par dialyse de PepA employée pour les expériences de topologie.....	98
II.1.12.Tampon pour le <i>crosslinking</i> au glutaraldéhyde.....	99
II.1.13.Tampon pour les migrations SDS-PAGE.....	99
II.1.14.Tampons pour les protéines de fusion 6-HIS.....	99
II.1.15.Tampons pour la purification des protéines marquées au MBP.....	100
II.1.16.Tampons pour les expériences utilisant les billes magnétiques à amylose...	100
II.1.17.Tampons pour les expériences au DSP.....	100
II.1.18.Tampons pour les expériences de <i>pulldown</i> après <i>crosslinking</i>	

réversibles au DSP.....	99
II.1.19. Tampon de recombinaison in-vitro.....	101
II.1.20. Solutions de Southern-blot.....	102
II.1.21. Solution pour la Topoisomérase 1.....	102
II.1.22. Protéines ArgR et PepA utilisées.....	103
II.1.23. Tampon de vérification de l'activité Aminopeptidase A.....	104
II.1.24. Tampon de charge pour les gels d'ADN.....	104
II.1.25. Tampon de charge pour les gels de protéines.....	104
II.1.26. Tampon Tris-borate-EDTA.....	104
II.2. Les méthodes.....	105
II.2.1. Transformation chimique des bactéries.....	105
II.2.2. Expériences in-vivo recombinaison pCS210.....	105
II.2.3. Expériences in-vivo répression <i>argA::lacZ</i> .....	106
II.2.4. Préparation d'un gel d'agarose 1 % .....	107
II.2.5. Préparation d'un mini-gel SDS-PAGE.....	107
II.2.6. Préparation d'un gel pour un retard sur gel.....	107
II.2.7. Induction de protéines par IPTG.....	107
II.2.8. Purification partielle de la protéine ArgR.....	108
II.2.9. Purification partielle de PepA.....	109
II.2.10. Autre purification partielle pour PepA.....	110
II.2.11. Purification des protéines fusionnées à une queue N-terminale 6-HIS.....	111
II.2.12. Purification des protéines marquées au MBP.....	111
II.2.13. Protocole de récupération des protéines par billes magnétiques avec Amylose.....	112
II.2.14. Étude d'interaction protéine-ADN in vitro.....	113
II.2.15. <i>Crosslinking</i> au glutaraldéhyde.....	115
II.2.16. <i>Crosslinking</i> au DSP.....	115
II.2.17. Purification par billes magnétiques Ni-NTA.....	116
II.2.18. Construction, Surexpression et purification de PepANt et de PepACt.....	117

II.2.19. Interaction des protéines PepANt-MBP et PepACt-MBP avec les dérivés d'ArgR.....	119
II.2.20. Recombinaison in-vitro de pCS210.....	119
II.2.21. Expériences sur la topologie de l'ADN.....	121
II.2.22. Imagerie par Typhoon.....	122
II.2.23. Mutagenèse dirigée sur <i>argR</i> .....	123
II.2.24. Expérience de retard sur gel avec ArgBOX marquées par HEX.....	125
II.2.25. Expérience de retard sur gel avec ArgBOX marquées par la digoxygénine (dig).....	126
II.2.26. Expérience de retard sur gel avec les séquences <i>cer</i> marquées par la digoxygénine (dig).....	127
II.2.27. Imagerie par Imager.....	127
II.2.28. Dosage des protéines par <i>Bradford Assay</i> .....	128
II.2.29. Vérification de l'activité Aminopeptidase A.....	129
 CHAPITRE III- RÉSULTATS.....	 130
 III.1. Introduction.....	 131
 III. 2. Caractéristiques des mutants d'ArgR et PepA.....	 132
III.2.1. Recombinaison in-vivo du plasmide pCS210 dans les cellules DS956 avec différents mutants d'ArgR en C-terminal.....	132
III.2.2. Vérification de la répression de la fusion <i>argR::lacZ</i> dans les cellules EC146λ par les différentes versions d'ArgR.....	142
III. 2.3. Construction des ArgR tronquées.....	144
III. 2.4. Mutagenèse site dirigée en positions 149 et 151.....	146
 III. 3. Activité de liaison des mutants avec l'ADN <i>in vitro</i> .....	 148
III.3.1. Caractérisation des activités de liaison d'ArgRWt, ArgR149, ArgR5aa sur la séquence boîte arginine.....	148
III.3.2. Caractérisation des activités de liaison d'ArgRWt, ArgR149,	

ArgR5aa sur la séquence <i>cer</i> .....	152
III.4.Caractérisation des capacités d'interactions des protéines.....	157
III.4.1. <i>Crosslinking</i> au glutaraldéhyde.....	157
III.4.1.1. <i>Crosslinking</i> avec ArgR.....	157
III.4.1.2. <i>Crosslinking</i> avec PepA.....	162
III.4.1.3. <i>Crosslinking</i> avec les versions tronquées de PepA.....	164
III.4.2. <i>Crosslinking</i> réversibles avec DSP.....	168
III.4.3.Expériences de recombinaison <i>in vitro</i> .....	174
III.4.4.Expériences de topologie de l'ADN.....	177
CHAPITRE IV – DISCUSSION.....	182
IV.1. Travaux précédents.....	183
IV.2. Protéines ArgR générées .....	186
IV.3. Retards sur gel et crosslinking.....	188
IV.4. Possibles causes des problèmes de multimérisation.....	192
IV.5. Comparaisons expérimentales pour les retards sur gel.....	194
IV.6. Expériences avec PepA.....	196
IV.7. Expériences de recombinaison <i>in vitro</i> .....	201
IV.8. Expériences de topologie.....	204
CHAPITRE V – CONCLUSION ET PERSPECTIVES.....	213
V.1.Récapitulatif des résultats.....	215
V.2.Mutagenèse d'ArgR.....	217
V.3.Mutagenèse de PepA.....	222
V.4.Autres expériences.....	224
BIBLIOGRAPHIE.....	228

REMERCIEMENTS.....i

ANNEXE – PUBLICATION 2010

**Liste des tableaux**

<b>Tableau 1.</b> Récapitulatif de plasmides utilisant un partisome.	13
<b>Tableau 2.</b> Récapitulatif des systèmes d'accoutumance associés aux plasmides R1, F et P1.	16
<b>Tableau 3.</b> Quelques espèces de bactéries avec le nombre de gènes <i>xer</i> présents dans leur génome.	23
<b>Tableau 4.</b> Liste de gènes faisant parti du régulon ArgR.	63
<b>Tableau 5.</b> Liste des plasmides produisant les protéines sans marquage.	92
<b>Tableau 6.</b> Liste des plasmides n'étant pas dérivés de pDB169 et de pGS38 et produisant des protéines sans marquage.	93
<b>Tableau 7.</b> Liste des plasmides produisant les protéines fusionnées à une queue 6-HIS	93
<b>Tableau 8.</b> Liste des plasmides produisant les protéines fusionnées au MBP	94
<b>Tableau 9.</b> Liste des souches de cellules utilisées dans les expériences.	95
<b>Tableau 10.</b> Liste des protéines Xer utilisées.	95
<b>Tableau 11.</b> Présentation des différentes façons d'obtenir les différents mutants pour ArgR et PepA.	103
127	
<b>Tableau 12.</b> Gamme d'étalonnage pour le Bradford <i>assay</i> .	128

<b>Tableau 13.</b> Comparaison des différentes versions d'ArgR.	204
<b>Tableau 14.</b> Présentation de mutants à générer à partir d'ArgR5aa.	217



**Liste des figures**

<b>Figure 1.</b> Effet de la recombinaison sur le plasmide ColE1 dimèrisé.	8
<b>Figure 2.</b> Partition du plasmide R1.	12
<b>Figure 3.</b> Représentation de la séquence <i>cer</i> sur le plasmide ColE1	17
<b>Figure 4.</b> Schéma de la séquence <i>cer</i> .	18
<b>Figure 5.</b> Alignement des parties Ct entre les protéines XerC et XerD chez <i>Escherichia coli</i> .	21
<b>Figure 6.</b> Modèle d'interaction entre le site <i>dif</i> et les protéines XerC et XerD.	25
<b>Figure 7.</b> Recombinaison médiée par XerC et XerD sur le site <i>dif</i> .	26
<b>Figure 8.</b> Recombinaisons FtsK dépendantes et FtsK indépendante au niveau du site <i>dif</i> .	27
<b>Figure 9.</b> Schéma de la structure des sites <i>cer</i> et <i>mwr</i> .	31
<b>Figure 10.</b> Comparaison entre les séquences d'ADN de l'extrémité 3' de 150 pbs des séquences <i>cer</i> et <i>nmr</i> .	31
<b>Figure 11.</b> Représentation schématique du polycistron contenant le gène <i>xerC</i> .	33
<b>Figure 12.</b> Représentation schématique du polycistron contenant <i>xerD</i> .	34
<b>Figure 13.</b> Structure du domaine Ct de XerD.	35

<b>Figure 14.</b> Schéma du phage CTX $\phi$ .	36
<b>Figure 15.</b> Repliement de l'ADN monobrin (soit ssDNA sur la figure) du phage CTX $\phi$ , au niveau du site <i>attP</i> .	38
<b>Figure 16.</b> Modèle d'intégration du phage dans le chromosome 1.	39
<b>Figure 17.</b> Appariement le plus probable des brins du variant El Tor de CTX $\phi$ .	40
<b>Figure 18.</b> Schéma des interactions de type Watson et Crick au niveau des paires de bases entre les sites <i>dif1</i> du chromosome 1 et <i>dif2</i> du chromosome 2.	41
<b>Figure 19.</b> Alignement des séquences de fixation de XerC et de XerD.	42
<b>Figure 20.</b> Structure 3D de la sous-unité FtsK chez <i>E.coli</i> et <i>P.aeruginosa</i> .	43
<b>Figure 21.</b> Représentation d'un dimère de deux sous-unités de la protéine FtsK de <i>P.aeruginosa</i> .	44
<b>Figure 22.</b> Représentation simplifiée de l'hexamère formé par FtsK au moment de la ségrégation.	46
<b>Figure 23.</b> Schéma d'un dimère de protéine ArcA.	48
<b>Figure 24.</b> Représentation de la région de liaison entre les deux monomères d'ArcA.	49
<b>Figure 25.</b> Représentation de la sous-unité de PepA et de son hexamère.	53
<b>Figure 26.</b> Séquence et alignement des séquences de différents homologues de la protéine PepA issues de différents organismes, <i>E.coli</i> (ECOLI), <i>S.typhimurium</i> (SALTY),	

<i>V.cholerae</i> (VIBCH), <i>H.influenzae</i> (HAEIN), <i>P.aeruginosa</i> (PSEAE), <i>R.prowazekii</i> (RICPR), <i>Bos taurus</i> (BOVIN).	54
<b>Figure 27.</b> Schéma de la région de contrôle de l'opéron <i>carAB</i> comportant les promoteurs <i>carP1</i> et <i>carP2</i> , sur lesquels se lie la protéine PepA au niveau des séquences PEPA1 et PEPA2.	55
<b>Figure 28.</b> Schéma interaction de PepA avec <i>carAB</i>	57
<b>Figure 29.</b> Séquence de la protéine ArgR d' <i>E.coli</i> .	60
<b>Figure 30.</b> Parties Ct de ArgR	64
<b>Figure 31.</b> Schéma agrandi de la zone d'interaction entre les molécules de L-arginine et d'ArgR.	65
<b>Figure 32.</b> Représentation des opérons <i>argT-hisJQMP</i> et <i>artPIQM-artJ</i> permettant respectivement la fabrication des systèmes de transport pour l'arginine et le système de transport pour l'arginine, de la lysine, de l'ornithine et de l'histidine.	66
<b>Figure 33.</b> Schémas représentant les systèmes de transports générés respectivement par les opérons <i>artPIQM-artJ</i> et <i>argT-hisJQMP</i> .	66
<b>Figure 34.</b> Séquence nucléotidique du gène <i>argR</i> avec les promoteurs P1 et P2.	67
<b>Figure 35.</b> La boîte arginine.	67
<b>Figure 36.</b> Promoteur de l'opéron <i>astCDABE</i> .	68
<b>Figure 37.</b> La protéine AhrC, sous-unité et hexamère.	71

- Figure 38.** Différents modèles pour la liaison de l'hexamère d'AhrC sur la boîte arginine. 73
- Figure 39.** Alignement de séquences entre les protéines ArgR d'*Escherichia coli* et AhrC de *Bacillus subtilis*. 74
- Figure 40.** Représentation tridimensionnelle d'un monomère d'ArgRBst. 75
- Figure 41.** Représentation de l'hexamère d'ArgRBst. 76
- Figure 42.** Représentation tridimensionnelle de l'interface entre la partie Ct et Nt d'ArgRBst. 76
- Figure 43.** Représentation tridimensionnelle de l'hexamère d'ArgR de *Mycobacterium tuberculosis*. 77
- Figure 44.** Modèle de liaison de l'hexamère d'ArgR de *Mycobacterium tuberculosis* à sa boîte arginine. 78
- Figure 45.** Représentations des complexes ADN-MtbArgR possibles. 79
- Figure 46.** Complexe ADN/protéines ArgR-PepA. 82
- Figure 47.** Nouvelle modélisation du complexe ADN-PepA-ArgR utilisant la méthode de calcul dite Tangle. 84
- Figure 48.** Seconde modélisation du complexe ADN-PepA-ArgR utilisant la méthode de calcul dite Tangle désormais appelée modèle de Sträter modifié. 85

<b>Figure 49.</b> Modélisation du complexe formé par les protéines accessoires ArgR/PepA.	86
<b>Figure 50.</b> Modèle modifié de Sträter <i>et al</i> proposant deux sites <i>cer</i> (vert et magenta) assemblés par 2 hexamères de PepA. L'ADN entre en contact avec la boucle Ct (en jaune) et les parties Nt adjacentes. L'hexamère ArgR est représenté en mauve.	88
<b>Figure 51.</b> Représentation de la protéine PepA d' <i>E.coli</i> .	118
<b>Figure 52.</b> Recombinaison pCS210 en pCS211.	134
<b>Figure 53.</b> Image en négatif d'une migration électrophorétique sur gel d'agarose 0,8% des plasmides pCS210, pCS211, pGS38 et pHSArgR5aa.	135
<b>Figure 54.</b> Migration sur gel d'agarose 0,8% des plasmides pCS210, pCS211, pGS38 et pHSArgR149.	136
<b>Figure 55.</b> Représentation des plasmides pDB169 et pGS38 utilisés pour générer les plasmides pBDArgR149 et pBDArgR5aa.	137
<b>Figure 56.</b> Migration sur gel d'agarose 0,8 % des plasmides pBDArgR149, pBDArgR5aa et pDB169 et leur influence sur la recombinaison du plasmide pCS210.	138
<b>Figure 57.</b> Purification d'ArgRWt native.	139
<b>Figure 58.</b> Purification d'ArgR5aa native.	140
<b>Figure 59.</b> Purification de la protéine ArgR149.	141
<b>Figure 60.</b> Purification de la protéine ArgRWt fusionnée avec 6-HIS.	142

<b>Figure 61.</b> Purification de la protéine ArgR5aa fusionnée avec 6-HIS.	142
<b>Figure 62.</b> Purification des protéines ArgR149 fusionnées avec 6-HIS.	143
<b>Figure 63.</b> migration sur gel d'agarose 0,8 % des plasmides pHSArgR150, pHSArgR151, pGS38, pHSArgR5aa, pHSArgRL149A, pHSArgRL151A et leur influence sur la recombinaison du plasmide pCS210.	146
<b>Figure 64.</b> Migration sur gel d'agarose 0,8% de nombreux clones contenant les plasmides pCS210, pCS211, pGS38 et pHSArgRE150A.	147
<b>Figure 65 :</b> Migration sur gel d'agarose 0,8 % des plasmides pQE31ArgRWt, pQE31ArgR149, pQE31ArgR5aa, pQE31ArgRL149A, pQE31ArgRL151A et leur influence sur la recombinaison du plasmide PCS210.	149
<b>Figure 66.</b> Retard sur gel d'ArgRWt avec la boîte arginine.	150
<b>Figure 67.</b> Retard sur gel d'ArgR5aa avec la boîte arginine.	151
<b>Figure 68.</b> Retard sur gel d'ArgR149 avec la boîte arginine.	152
<b>Figure 69.</b> Retard sur gel d'ArgRWt avec deux séquences <i>cer</i> .	154
<b>Figure 70.</b> Retard sur gel d'ArgR5aa avec deux séquences <i>cer</i> .	155
<b>Figure 71.</b> Retard sur gel d'ArgR149 avec deux séquences <i>cer</i> .	156
<b>Figure 72.</b> Retard sur gel de complexes PepA-ArgRWt avec deux séquences <i>cer</i> .	157
<b>Figure 73.</b> <i>Crosslinking</i> au glutaraldéhyde de la protéine ArgRWt.	159

<b>Figure 74.</b> <i>Crosslinking</i> au glutaraldéhyde d'ArgR5aa.	160
<b>Figure 75.</b> <i>Crosslinking</i> au glutaraldéhyde d'ArgR149.	161
<b>Figure 76.</b> <i>Crosslinking</i> au glutaraldéhyde sur un gel SDS-PAGE à gradient des protéines ArgRWt, ArgR5aa et ArgR149.	162
<b>Figure 77.</b> Purification partielle de PepA.	163
<b>Figure 78.</b> <i>Crosslinking</i> au glutaraldéhyde de la protéine PepA	164
<b>Figure 79.</b> Purification de la protéine PepANtMBP.	165
<b>Figure 80.</b> Purification de la protéine PepACtMBP.	165
<b>Figure 81.</b> <i>Crosslinking</i> au glutaraldéhyde de PepANtMBP.	166
<b>Figure 82.</b> <i>Crosslinking</i> au glutaraldéhyde de PepACtMBP.	167
<b>Figure 83.</b> Purification de la protéine ArgRL149A.	169
<b>Figure 84.</b> Purification de la protéine ArgRL151A.	169
<b>Figure 85.</b> Migration sur gel SDS-PAGE 15 % des différentes formes d'ArgR avec L-arginine.	170
<b>Figure 86.</b> Migration sur gel SDS-PAGE 15 % des différentes formes d'ArgR sans L-arginine.	171
<b>Figure 87.</b> Migration des protéines PepA tronquées lors des interactions protéine à protéine entre ArgR et les formes tronquées de PepA.	173

<b>Figure 88.</b> Purification de XerC.	174
<b>Figure 89.</b> Purification de XerD.	175
<b>Figure 90.</b> Migration <i>Southern-blot</i> des différentes recombinaisons <i>in vitro</i> .	176
<b>Figure 91.</b> Carte du plasmide pNCK1.	178
<b>Figure 92.</b> Purification de PepA après chauffage à 70°C.	179
<b>Figure 93.</b> Photo des différents profils des plasmides pNCK1 sur un gel d'agarose 0,8 % avec 0,5 µg/ml de chloroquine.	180
<b>Figure 94.</b> Détails des différences entre les séquences en Ct d'ArgR et d'AhrC.	184
<b>Figure 95.</b> Répression de l'activité β-galactosidase dans les cellules EC146λ (λAZ-7).	185
<b>Figure 96.</b> Représentation d'un hexamère des parties Ct d'ArgR d' <i>E.coli</i> .	193
<b>Figure 97.</b> Comparaison de la fin de la partie Ct des protéines ArgRWt et ArgR180.	209
<b>Figure 98.</b> Analyse Jpred3 de la séquence Ct de la protéine ArgR180.	209
<b>Figure 99.</b> Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgR5aa et ArgR149.	210



**Figure 100.** Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgRE150A, ArgRL149A, ArgRL151A et ArgRL149AL151A. 211

**Figure 101.** Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgR149, ArgR150 et ArgR151. 212

**Figure 102.** Alignement de séquence Nt entre les protéines ArgR d'*E.coli* (ArgRNt) de *B. stearothermophilus* (ArgRBstNt) et de *B.subtilis* (AhrCNt). 219

**Figure 103.** Alignement des séquences des hélices  $\alpha_6$  des protéines ArgR d'*E.coli*, d'AhrC de *B.subtilis* et ArgRBst de *B.stearothermophilus*. 221

**Liste des abréviations****Acides-aminés :**

A : Alanine  
C : Cystéine  
D : Acide aspartique  
E : Acide glutamique  
F : Phénylalanine  
G : Glycine  
H: Histidine  
I: Isoleucine  
K: Lysine  
L: Leucine  
M: Méthionine  
N: Asparagine  
P: Proline  
Q : Glutamine  
R : Arginine  
S : Sérine  
T : Thréonine  
V : Valine  
W : Tryptophane  
Y : Tyrosine

**Acides nucléiques :**

A : Adénine  
C : Cytosine  
G : Guanine

N : peut-être n'importe quel nucléotide

R : Purine

S : Cytosine ou Guanine

T : Thymine

U : Uracile

*Abréviations des expériences et produits:*

aa : acide-aminé

AAA : ATPases Associated with various cellular Activities

ADN : Acide Désoxyribonucléique

Ap<sup>R</sup> : résistance à l'ampicilline

ARN : acide ribonucléique

ATP : Adénosine triphosphate

$\alpha$  : Alpha

$\beta$  : Béta

BFP : Blue Fluorescent Protein

BSA : Bovine serum albumin

Dig : Digoxygénine

DMSO : Diméthylsulfoxyde

dNTP : Désoxyribonucléotide triphosphate

DNaseI : Désoxyribonucléase I

DSP : Dithiobis succinimidyl propionate

DST : Disuccinimidyl tartarate

EDTA : Acide 1-(4-Aminbenzyl)éthylènediamide-N,N,N',N'-tétraacétique

FRET : Fluorescence resonance energy transfert

GFP : Green Fluorescent Protein

IPTG : isopropyl $\beta$ -D-thiogalactopyranoside

LB : Luria Bertani

M : Molaire

MBP : Maltose binding protein (protéine se liant au Maltose)  
Ni-NTA : Nickel-nitrilotriacétique  
Nt : amino-terminal  
N-Terminal : amino-terminal  
DO : Densité Optique  
pb : paire de bases  
PCR : Polymerase chain reaction  
PM : Poids Moléculaire  
rpm : rotation par minute  
SDS : Sodium Dodécyl Sulfate  
SDS-PAGE : gel d'électrophorèse de polyacrylamide contenant du SDS  
SPR : surface plasmon resonance  
Tet<sup>R</sup> : résistance à la tétracycline  
<sup>32</sup>P : phosphore 32  
UV : ultra-violet  
X-gal : 5-bromo-4-chloro-3-indolyl-β D-galactopyranoside

*Abréviations de mesures :*

nM : nanomolaire  
Da : Dalton  
g : gramme  
h : heure  
μg : microgramme  
μg/mL : microgramme par millilitre  
μL : microlitre  
μM : micromolaire  
Kb : kilobases  
kDa : kilodalton  
min : minute

mM : millimolaire

ng : Nanogramme

nM : Nanomolaire

pmol : picomole

s : seconde

V : volt

° : degré

°C : degré Celsius

**CHAPITRE I**  
**INTRODUCTION**

## CHAPITRE I. INTRODUCTION

### I.1. Quelques rappels généraux

La plupart des organismes unicellulaires eucaryotes et procaryotes, sinon tous, se multiplient grâce au phénomène dit de division cellulaire où la cellule, après avoir dupliqué son information génétique, c'est-à-dire son ADN, va séparer les deux copies de cet ADN et distribuer les deux molécules filles de manière équitable dans chaque cellule fille. Ce qui signifie que chaque exemplaire produit par la réplication de l'ADN se retrouvera dans une seule cellule fille.

Les eucaryotes et les procaryotes présentent de nombreuses différences au niveau de leur organisation génétique et de leur division cellulaire:

- Présence de plusieurs chromosomes non circulaires chez les eucaryotes, généralement en forme de bâtonnets, où la molécule d'ADN linéarisée est liée à des protéines appelées histones.
- Un grand chromosome circulaire chez les bactéries, portant toute l'information génétique principale. Cependant, certaines bactéries ont plusieurs chromosomes (2 à 3), et d'autres ont des chromosomes linéaires. De plus, il existe des plasmides, petites molécules circulaires ou linéaires d'ADN existant chez toutes les bactéries, mais portant par exemple des gènes de résistance pour certains antibiotiques. Ces plasmides varient en nature et en nombre selon les souches bactériennes au sein d'une même espèce.
- Si les chromosomes bactériens ne peuvent que rarement s'échanger entre deux bactéries (la souche Hfr d'*E.coli* étant une de ces exceptions), les plasmides eux le peuvent, même parfois entre deux bactéries d'espèces différentes. C'est le phénomène de conjugaison.
- La division d'une cellule eucaryote se fera par étranglement de la membrane cellulaire grâce au cytosquelette, suite aux phases suivantes : prophase, métaphase, anaphase et télophase.

- La division cellulaire d'une cellule procaryote se fait par formation progressive d'un septum bactérien à partir de la membrane interne de la bactérie. De plus, contrairement aux eucaryotes, les deux copies du chromosome bactérien vont se retrouver au niveau du septum bactérien avant séparation et déplacement vers le pôle opposé des deux chromosomes « fils » et fermeture du septum, qui sera suivi de séparation des deux cellules filles.

Malgré toutes ces différences certains mécanismes de séparation des chromosomes sont conservés entre les eucaryotes et les procaryotes : nous nous attarderons sur les systèmes procaryotes.

Lors de la réplication du génome bactérien, comprenant à la fois le chromosome et les plasmides, le matériel génétique est répliqué de manière bidirectionnelle, bien que certains plasmides soient répliqués unidirectionnellement. Cette réplication a lieu sur le chromosome au niveau d'une séquence nommée *oriC* qui correspond à l'origine de réplication. La réplication va se poursuivre jusqu'à une région diamétralement opposée nommée *ter* (22 pb) pour terminateur, qui sera le lieu de terminaison de la réplication. Cette région est composée de deux séquences, *ter1* et *ter2*, qui servent de sites de liaison pour la protéine Tus. Cette liaison va permettre le blocage du système de réplication composé à la fois de l'hélicase et de la polymérase bactérienne. Après ce blocage, les deux copies du chromosome seront séparées grâce aux tyrosines recombinases XerC et XerD si il y a eu recombinaison entre ces deux chromosomes. C'est toujours la topoisomérase IV qui agit pour séparer les chromosomes liés sous forme de caténanes.

Il faut savoir que le système d'arrêt de la réplication employant la protéine Tus peut occasionner des bris d'ADN, mais le système employé par la bactérie pour réparer les cassures, le système RecBCD, peut entraîner la formation de caténanes et de dimères chromosomiques (Horiuchi *et al*, 1994), d'où l'importance des systèmes de résolution Xer dépendants.



Durant ma thèse, j'ai étudié un de ces systèmes de réparation de dimères : celui impliquant la séquence *cer* présente sur le plasmide ColE1.

Mais avant de parler plus en détail de la résolution des dimères, il faut présenter les systèmes de ségrégation et de partition du chromosome, voire des plasmides, qui ne peuvent donc pas avoir lieu si les dimères ne sont pas résolus, compromettant la survie de la cellule et la transmission de son patrimoine génétique d'une génération à une autre.

## **I.2. Systèmes de ségrégation et de partition.**

### **I.2.1.Ségrégation**

Lors de cet événement, il y a d'abord décaténation ou résolution des dimères puis migration de chaque chromosome vers les pôles de chaque cellule fille.

On suggère que c'est la topoisomérase IV qui catalyserait l'échange des brins d'ADN permettant aux deux chromosomes fils enchevêtrés de se séparer lors d'une décaténation (Zechiedrich and Cozzarelli, 1995).

Le système de résolution des dimères chromosomiques utilise les protéines XerC et XerD et a lieu au niveau du site *dif*, ce système sera décrit plus en aval de cette thèse.

Une propagation stable d'un plasmide bactérien peut se faire par une ségrégation au hasard du plasmide ou via une partition active. Si le nombre de copies est assez élevé, il n'y a pas de barrière à une diffusion libre et la ségrégation au hasard sera suffisante (il faut tout de même un plasmide avec un fort nombre de copies.) et la probabilité qu'une cellule fille n'ait aucun plasmide reste faible.

De surcroît, les mécanismes de contrôle opérant au niveau de la réplication peuvent corriger toute déviation du nombre normal de copies due à une ségrégation inégale.

Par contre, des plasmides à faible nombre de copies doivent utiliser des systèmes de partition actifs pour éviter tout problème de ségrégation. Certains plasmides possèdent même un système d'accoutumance génétique pour éliminer la compétition avec les cellules sans plasmides (Gosh *et al*, 2006). Dans ces systèmes, il y a fabrication d'une toxine et d'un antidote à durée de vie plus courte.

### **I.2.2.Partition**

Le modèle le plus récent met en avant le déplacement actif des deux chromosomes fils séparés vers les pôles de la bactérie (telle *E.coli*). Ce serait la protéine MukB (un analogue à la dynamine eucaryote) qui se lierait, sous forme de dimères, au chromosome et serait responsable de ce déplacement (Hiraga *et al*, 1989 ; Hiraga 1993 ; Niki *et al*, 1992). Chez *E.coli*, chaque copie du chromosome possède une origine de réplication qui va être orientée vers une des extrémités de la cellule avant sa partition. La partie terminale de ces mêmes copies, quant à elle, va être localisée au centre de la cellule au niveau du septum. Cette migration est très rapide (10 fois plus que l'élongation) et pourrait être permise par des séquences comprises dans la région en boucle contenant la séquence *oriC*. Et ce n'est pas tout, les régions citées ci-dessus pourraient être ancrées dans la membrane cellulaire, ce qui faciliterait la migration vers les pôles (Gordon *et al*, 1997 ; Webb *et al*, 1998). Les équipes d'Hiraga et Imai ont également avancé que tout le système de réplication, appelé réplisome, resterait accroché à la membrane cellulaire au niveau du septum cellulaire. L'ADN nouvellement recopié glisserait donc dans ce système puis serait rejeté dans des directions opposées alors que la région possédant le site *ter* resterait au niveau du septum (remarquons que cette région chez *E.coli*, comprend aussi le site *dif*.) Cependant, en 2008, Reyes-Lamothe *et al* ont observé que le réplisome ne serait fixé à

aucune structure cellulaire particulière. Les deux fourchettes de réplication sœurs se comportent indépendamment durant la phase S. L'ADN ne fait que s'ancrer dans le réplisome.

Après phénomène de condensation de l'ADN, il y a ségrégation de l'ADN bactérien. Chez *Escherichia coli*, ce sont les protéines MukB qui la permettent en se couplant avec les topoisomérases IV, de plus, MukB s'assemble avec MukE et MukF pour permettre cette condensation de l'ADN (Sawitzke and Austin, 2000 ; Holmes and Cozzarelli, 2000 ; Li *et al*, 2010) servant ainsi de condensine. Si ces protéines semblent prépondérantes pour la condensation du chromosome d'*Escherichia coli*, d'autres ont été mises en évidence via les gènes suivants : *crcA*, *crcB*, *cspE*. En effet, ces gènes coderaient pour de petites protéines qui seraient également partie prenante dans la condensation du chromosome (Hu *et al*, 1996).

A noter que chez d'autres bactéries, ce sont des protéines nommées SMC (*structural maintenance of chromosome*) catégorisées comme des « *histone-like proteins* » qui permettent cette condensation, avec la participation de la gyrase (Drlica, 1990 ; Gellert *et al*, 1976).

### **I.3. Système de ségrégation plasmidique**

#### **I.3.1. Vue générale (recombinaison site spécifique)**

Tout comme les chromosomes, les plasmides doivent également passer par les systèmes de ségrégation et de partition pour être transmis d'une génération à une autre.

Pour le plasmide ColE1, les mécanismes mis en place passent par le système de recombinases à tyrosines XerC et XerD, et ressemblent beaucoup à celui utilisé pour le chromosome bactérien. La principale différence étant que l'on ne fait pas migrer les

plasmides vers les pôles bactériens. Il existe de nombreux systèmes de ségrégation plasmidique, dont plusieurs ont été décrits pour des plasmides à faible nombre de copies, contrairement au plasmide ColE1. Il s'agit d'ailleurs de recombinaisons à spécificité de site. Ces recombinaisons ont la particularité de nécessiter quatre brins d'ADN pour pouvoir se faire. Ces quatre brins seront ensuite clivés grâce à l'existence de deux sites de recombinaison, qui seront échangés et reliés de manière définie (Hallet and Sherratt, 1997). Les mécanismes de recombinaison sont intramoléculaires, et donnent deux molécules d'ADN si les sites de recombinaison sont directement répétés. Différents types d'enzymes peuvent permettre ce type de recombinaison, les invertases, les résolvasés et les intégrases qui sont des recombinases.

Dans le cas du site *cer* du plasmide ColE1, il y a intervention de recombinases de la famille des recombinases à tyrosine (recombinases du phage  $\lambda$ ). Les deux molécules obtenues après la recombinaison sont des monomères de plasmides dimérisés.

Il est à noter que si les séquences ne sont pas directement répétées sur la molécule d'ADN, il n'y aura pas, par exemple, de séparation d'un dimère en deux monomères, mais juste une inversion de la séquence d'ADN sur la même molécule (voir figure 1).

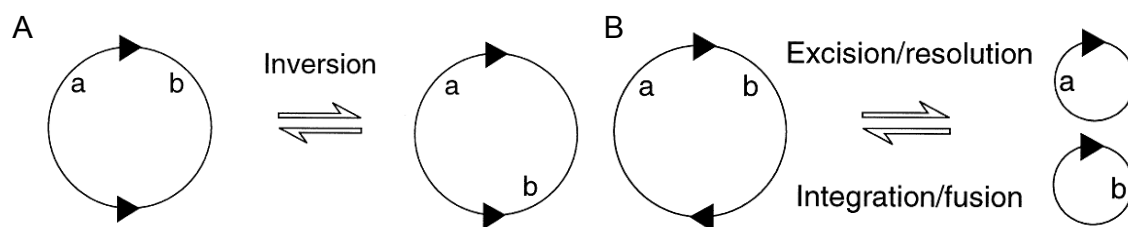
Les systèmes de recombinaison sites spécifiques diffèrent des transpositions dans le fait que pour les transpositions, on a affaire à un élément d'ADN mobile, généralement un petit fragment d'ADN appelé transposon.

Ce dernier (il existe de très nombreux types de transposons) peut s'insérer dans plusieurs endroits d'un génome bactérien grâce à des enzymes appelées transposases. Contrairement à la recombinaison site spécifique, la transposition répllicative est généralement accompagnée de la synthèse du transposon (donc synthèse d'ADN, qui est ensuite inséré dans le génome, ce qui n'a pas lieu dans la transposition

conservative). De plus la recombinaison site spécifique ne s'accompagne pas de dégradation de l'ADN, et généralement la consommation d'ATP n'a pas lieu.

Les protéines impliquées dans le système de résolution site spécifique des multimères plasmidiques étudié lors de cette thèse proviennent de la bactérie *E.coli*. On a étudié particulièrement le site *cer* du plasmide ColE1 contenant une séquence de 280 paires de bases. Sur cette séquence, on retrouve deux sites de liaison pour les recombinases à tyrosine XerC et XerD, mais aussi une séquence de 18 paires de bases correspondant à une demie-boîte arginine. Lorsque le plasmide est au moins dimérisé, la séquence *cer* est également dupliquée, ce qui reconstitue la boîte arginine en entier et génère quatre sites de liaison pour les recombinases.

Ce phénomène permet donc la résolution des dimères de ColE1 en monomères. Grâce à une séquence *cer* qui est directement répétée sur la molécule dimérisée.



**Figure 1.** Effet de la recombinaison sur un plasmide dimérisé.

**A)** La partie de cette figure montre ce qui se passe lorsque les sites de recombinaison sont inversement répétés, il n'y a pas de résolution des dimères, mais il y a juste une inversion d'un des segments d'ADN (représenté ici par a et b).

**B)** La partie B montre ce qui se passe lorsque les sites de recombinaison sont directement répétés, comme pour le site *cer*. Si la partie B montre une recombinaison bidirectionnelle, ce n'est pas le cas lorsque la recombinaison est médiée par les protéines du complexe XerC, XerD, ArgR et PepA pour le plasmide ColE1. Ces protéines favorisent le sens excision/résolution et non l'autre sens, c'est-à-dire, l'intégration/fusion, qui est plutôt favorisée par la protéine RecA. Il s'agit bien d'une recombinaison intramoléculaire où les sites de recombinaison sont directement répétés (avec permission de reproduction, Hallet *et al*, 1997).

Cette résolution des dimères en monomères permet de garder les plasmides de façon stable au cours des générations bactériennes (figure 1). En effet, la multimérisation d'un plasmide peut ralentir la croissance de la bactérie, mais surtout, le

plasmide multimérisé peut-être perdu au fil des générations, avec les gènes qu'il contient s'il n'y a pas de résolution des multimères en monomères.

Pour la réparation des multimères en monomères plasmidiques, des protéines accessoires sont nécessaires. Ce sont les protéines ArgR (*Arginine Repressor*) et PepA (Aminopeptidase A) qui servent de protéines accessoires dans le système impliquant la séquence *cer*, et qui seront étudiées au cours de cette thèse.

### **I.3.2. Autres systèmes de ségrégation plasmidique**

#### **I.3.2.1. Généralités**

Si le plasmide ColE1 possède un système de ségrégation pour un plasmide à grand nombre de copies, il existe bien sûr d'autres systèmes de ségrégation et de partition plasmidique, notamment pour des plasmides à faible nombre de copies. Tous ces systèmes visent à résoudre des dimères en monomères et à garantir leur pérennité et stabilité lors des ségrégations au fur et à mesure des divisions cellulaires. Citons par exemple les systèmes de partitions employés par les plasmides suivants :

- Plasmide F, c'est le plasmide qui porte les facteurs de conjugaison chez *E. coli* (Mori *et al*, 1989 ; Ogura and Hiraga, 1983).
- Plasmide R1 (et R100) (Hiraga, 2000 ; Jensen and Gerdes, 1997 ; Jensen *et al*, 1998).
- Plasmide P1 (Abeles *et al*, 1985 ; Bouet *et al*, 2000 ; Funnel, 1991 ; Funnel and Gagnier, 1993 ; Bouet and Funnel, 1999).

Tous ces systèmes ont un point en commun, ils voient leur partition médiée par un partisome, c'est-à-dire plusieurs protéines capables de se lier à l'ADN, dont une possède une fonction ATPase, et au moins une autre se liant en un site *cis* à l'ADN. Ce

sont souvent des systèmes composés au moins de deux protéines *trans* codées par un opéron et qui ne font pas intervenir de recombinaisons.

Les plasmides F, R1 et P1 possèdent tous une séquence agissant comme des centromères. Pour le plasmide F, il s'agit de la séquence *sopC*, pour le plasmide R1, de la séquence *parC* et pour le plasmide P1, de la séquence *parS* (ces séquences sont considérées comme des loci de même type appelé *par*). Pour chacun de ces systèmes de partition cités plus haut, il faut une ATPase qui s'associe à une autre protéine dite partenaire qui elle-même se lie à la séquence centromérique. Ce complexe protéine de liaison-centromère est appelé système de pré-partition, et va recruter l'ATPase liée à de l'ATP pour déclencher la partition du plasmide (Gosh *et al*, 2006). La localisation des centromères est variable et les loci du type *par* se caractérisent par de multiples séquences répétées directes en amont ou en aval des gènes de type *par*.

Actuellement, les ATPases citées précédemment sont divisées en deux groupes :

- Les ATPases de types Walker, SopA du plasmide F et ParA du plasmide P1.
- Les ATPases de type actine, dont fait partie ParM du plasmide R1.

Les ATPases de type Walker possèdent un domaine de liaison aux nucléotides qui est ion dépendant, dans ce cas  $Mg^{2+}$ . Les ATPases de type actine contiennent un motif de liaison à l'ATP qui leur est propre. En général, les loci des centromères sont sous autorégulation. Par exemple, les protéines ParA et SopA s'associent aux séquences des opérateurs dans la région promotrice grâce à leur domaine Nt. ParB et SopB agissent comme des corépresseurs.

Pour le plasmide F, une répression maximale requiert le centromère, mais probablement avec des boucles d'ADN.

Pour le plasmide P1, ParA interagit avec le complexe centromère-ParB lorsqu'il est lié à l'ATP (rappelons que le centromère est la séquence *parS*). ParA fonctionnerait comme un agent de partition seulement si lié à l'ATP, et comme répresseur de la

transcription si lié à l'ADP. Les phénomènes sont les mêmes pour le plasmide F et ses protéines SopA et SopB.

Pour le plasmide R1, la protéine ParM (ATPase) ne régule pas l'opéron, c'est la protéine ParR qui le régule complètement.

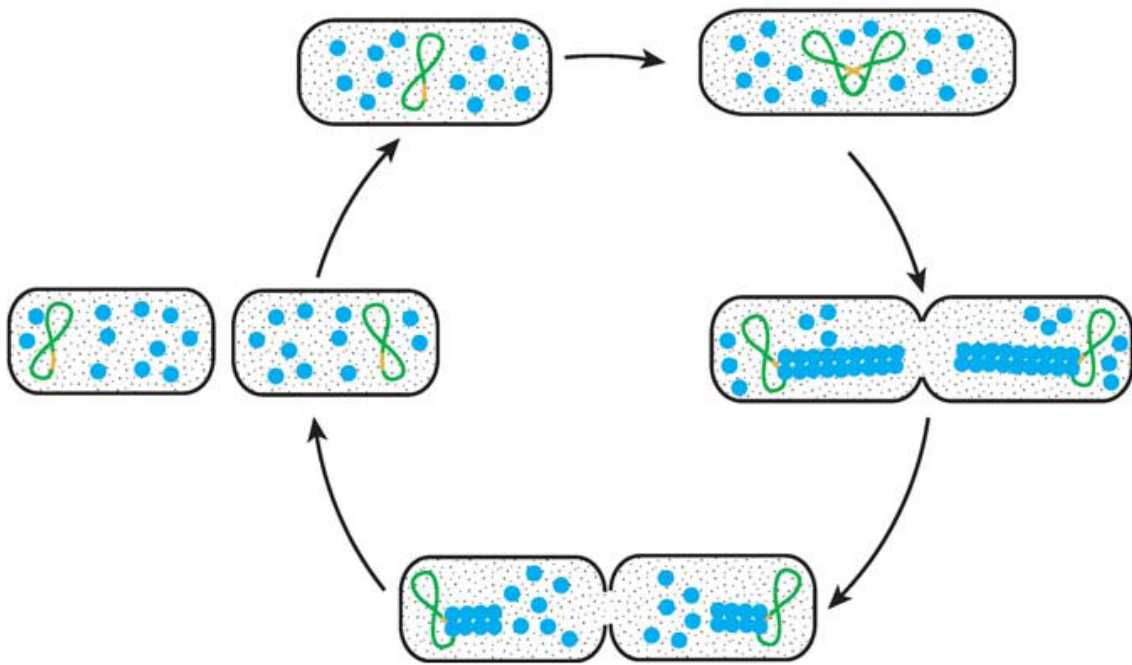
### **I.3.2.2.Mécanismes de partition**

Avec les systèmes de type Par, il y a association entre les protéines Par ou Sop avec la séquence faisant office de centromère. Ce qui entraîne une localisation précise des plasmides dans la bactérie. On sait par exemple, que l'IHF (*Integration Host Factor*) facilite l'assemblage du système de partition de P1 en augmentant l'affinité de ParB (sous forme de dimère) pour son centromère (Bouet *et al*, 2000), mais ce n'est pas essentiel, car ParB peut très bien se lier à parS sans intervention d'IHF (Bouet *et al*, 2000), mais son affinité est grandement réduite (Funnell, 1991). Dans ce type de mécanisme, les plasmides P1 et F sont présents au centre de la cellule fille et y restent jusqu'à la réplication. Puis après leur réplication, chaque plasmide va respectivement au quart et à la position trois-quarts de la cellule, avant la formation du septum (même si certaines données semblent indiquer que ce n'est pas juste avant la formation de ce septum (Gosh *et al*, 2006). Il ne semble pas qu'il y ait un seul centre de partition pour chaque plasmide cité ci-dessus. De plus, les protéines ParB et SopB peuvent être responsables de l'appariement de par leur association, à moins que ParA et SopA ne soient responsables du désappariement des plasmides. Pour le plasmide P1, ParB est requis pour « attacher » le plasmide au milieu de la cellule et ParA va « éjecter » le plasmide de cet endroit après la réplication (Gosh *et al*, 2006).

Le système pour le plasmide R1 est plus connu. Le plasmide R1 va d'abord se positionner près d'un des pôles de la cellule fille. Il y a ensuite déplacement du plasmide vers le centre de la cellule pour l'amener vers le réplisome pendant la croissance cellulaire. Après septation, le plasmide retourne au pôle.



L'ATPase ParM est de type actine et s'oligomérise pour former des filaments dépendants de l'ATP et de  $Mg^{2+}$ . Dans ce système, ParR se lie au centromère *parC* et entraîne la médiation des deux plasmides nouvellement répliqués. La formation du complexe ParR-*parC* serait donc essentielle à l'oligomérisation de ParM. Dans ce système, les plasmides R1 sont situés sur les pointes distales des filaments de ParM oligomérisés en fin de réplication. Après la duplication du plasmide, les deux répliquats sont maintenus ensemble par ParR en *parC*. Ce complexe étant le site de polymérisation de ParM+ATP, et plus les filaments grandissent, plus les plasmides sont séparés, jusqu'au pôle. La dépolymérisation des filaments à partir de leur queue va libérer ParM+ADP sous forme de monomères (Gosh *et al*, 2006) (figure 2).



**Figure 2.** Partition du plasmide R1. Le plasmide R1 (en vert), se trouve près des pôles de chaque cellule fille. Le plasmide va alors se déplacer au centre de la cellule au moment de la réplication. La protéine ParM-ATP (en bleu) va se polymériser en filament à partir du site *parC* (en jaune), poussant les nouveaux plasmides vers les pôles de la cellule avant septation. Il y a ensuite dépolymérisation de ParM-ADP (avec permission de reproduction, Gosh *et al*, 2006).

Les protéines et sites de liaison des trois systèmes précédents ont été récapitulés dans le tableau suivant.

Plasmide	ATPase	Facteur de liaison	Site d'ADN
Plasmide F	SopA (s'associe à la boîte <i>sopA</i> , qui fait partie de l'opéron <i>sopAB</i> ). Elle est membranaire	SopB (se lie à une région de 7 pb dans le motif de <i>sopC</i> , stimule SopA.)	<i>sopC</i> (12 répétitions d'un motif de 43 pb)
Plasmide R1 (et R100)	ParM, son activité est stimulée par ParR lié à l'ADN	ParR, se liant à <i>parC</i> , activé par l'hydrolyse de l'ATP.	<i>parC</i> , situé dans la région de l'opéron <i>parM-parR</i>
Plasmide P1 (prophage)	ParA. Il faut la forme ParA-ATP pour former le complexe. ParA-ADP réprime l'opéron.	ParB, se liant à <i>parS</i> , il en faut une forte concentration pour recruter ParA.	<i>parS</i> situé dans l'opéron <i>parAB</i>

**Tableau 1.** Récapitulatif de plasmides utilisant un partisome.

Ces systèmes entraînent des changements dans la topologie des plasmides dont ils permettent la ségrégation/partition. Ces changements permettent donc la mise en marche du partisome, et donc la recombinaison.

Pour ces plasmides cités plus haut, il existe d'autres systèmes pour garantir leur stabilité au sein de la cellule, par exemple :

- Plasmide F, utilisation de la protéine D, une résolvasse.
- Plasmide R1, utilisation de XerC et XerD sur un site *dif* propre au plasmide.
- Plasmide P1, utilisation du système Cre-Lox.

### **I.3.2.3.L'accoutumance génétique**

De plus, on s'est aperçu que les plasmides F, R1 et P1 possédaient un système dit d'accoutumance génétique qui assure sa pérennité au sein des cellules d'*Escherichia coli*.

Ce système a été découvert par introduction d'un plasmide F endommagé par les UV dans une bactérie, ce plasmide endommagé a engendré un prophage et une mise en marche du système SOS. Or, cette réponse SOS est programmée par le locus *ccd* sur le plasmide F (Borek *et al*, 1958). On en a donc une maintenance médiée pour le plasmide F portant *ccd* qui entraîne la mort de la cellule si le plasmide n'est plus présent après ségrégation en partition de la cellule (Jaffe *et al*, 1985). Pour le plasmide R1, un processus similaire a été décrit pour le locus *hok/sok*, on a donc mis-à-jour une notion de « mort post-ségrégationnelle » (Gerdes *et al*, 1986). Cette mort post-ségrégationnelle implique la perte d'éléments génétiques internes à la cellule. Si un antibiotique empêche l'expression de certains gènes, il y a aussi mort cellulaire. De plus on s'est aperçu que les gènes impliqués dans cette mort post-ségrégationnelle, appelée aussi accoutumance génétique, peuvent être aussi bien plasmidiques que chromosomiques (Kobayashi, 2004).

#### **I.3.2.3.1.Le système protéique classique**

Il existe plusieurs types de mécanismes, mais l'un de ceux rencontrés dans F, P1 et R1 est le système dit protéique classique. Dans ce système dit classique, on a un opéron situé sur un plasmide, qui produit à la fois une toxine et une antitoxine, or, dans tous les systèmes décrits, l'antitoxine se dégrade plus vite que la toxine, ce qui au final peut entraîner la mort de la cellule.

On va tenter de décrire les systèmes présents dans chaque plasmide ci-dessus.

Pour les plasmides R100 et R1, il existe deux modules d'accoutumance, respectivement *pem* et *parD*, qui codent la toxine PemK (alias Kid) et l'antitoxine Pem1 (alias Kis). Il existe deux homologues chromosomiques connus pour ce système, *mazEF* et *chpB* qui possèdent les mêmes effets (Masuda *et al*, 1993). Tous ces modules ont une faible similarité avec *ccd*, le système propre au plasmide F.

Ce plasmide F code la toxine CcdB et l'antitoxine CcdA. On remarquera que les protéines Kid et MazF (Kis) sont similaires à la toxine CcdB. Le plasmide F possède donc le système Ccd, où la toxine CcdB se lie à la DNA gyrase et entraîne une cassure double brin. L'antitoxine CcdA inhibe la toxine en se liant à elle, permettant à la cellule de survivre (Hiraga *et al*, 1986 ; Jaffe *et al*, 1985).

Pour le plasmide P1, on a le système Phd-Doc où Phd est l'antitoxine et Doc la toxine. De plus, P1 serait un prophage qui se réplique comme un plasmide. Dans le système, la toxine inhibe la synthèse des protéines, il y a une réduction de la taille de la cellule et pas d'induction du système SOS. Plus encore, le système MazEF serait mis en œuvre (Hazan *et al*, 2001 ; Kobayashi, 2004).

#### **I.3.2.3.2. Systèmes d'accoutumance avec ARN anti-sens**

Mis-à-part les systèmes classiques, il existe des systèmes d'accoutumance avec des ARN anti-sens, que l'on retrouve avec les plasmides R1 et F. Pour le plasmide R1, on a le système *hok/sok*, où Hok (*hok* pour *host killing gene*) est la toxine, et ARN-Sok l'antitoxine (*sok* pour *suppression of killing*). Dans ces systèmes, la toxine est une protéine de 52 acides-aminés qui attaque la membrane. Les antitoxines sont des ARN anti-sens de 64 nucléotides complémentaires à l'ARNm de *hok* (Gerdes *et al*, 1997). La protéine Hok s'associerait à la membrane de par sa partie Nt (qui est transmembranaire), tandis que la partie Ct reste dans le périplasme. Il en résulte que la toxine entraîne une perte du potentiel membranaire, voire une perforation de la cellule. Les ARN anti-sens empêchent la synthèse des toxines. Le plasmide F, quant à lui,

présente un système à ARN anti-sens similaire à celui du plasmide R1, où c'est SrnB' qui sert de toxine et l'ARN-SrnC d'antitoxine. Ce système est appelé *srnB* (Nielsen *et al*, 1991). Il existe une troisième catégorie de systèmes d'accoutumance, celui de la restriction-modification, qui n'est pas utilisé par les plasmides R1, F et P1. Mais ce système est fréquent et met en jeu une protéine de restriction (toxine), qui cible l'ADN, le meilleur exemple est l'enzyme EcoRI, codé par le plasmide RTF-1. L'antitoxine, qui est une méthylase, protège l'ADN de la toxine par méthylation de ce dernier (Kobayashi, 2004).

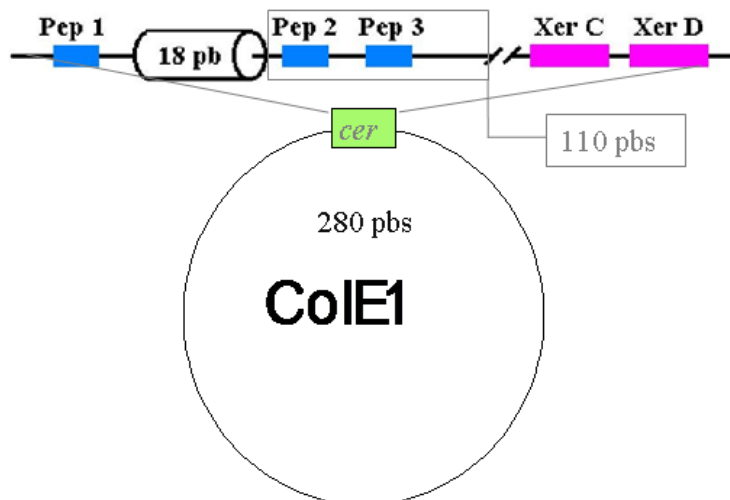
type	Locus (plasmid)	Toxin	Target	Antitoxin	Protease/RNase-destroying antitoxin
classical proteic systems	<i>ccd</i> (F)	CcdB	DNA gyrase	CcdA (=LetA)	Lon
classical proteic systems	<i>pem/parD</i> (R100/R1)	PemK/Kid	DnaB dependent DNA replication	PemI/Kis	Lon
classical proteic systems	<i>phd/doc</i> (P1)	Doc	MazEF addiction system ?	Phd	ClpXP
Antisense-RNA-regulated systems	<i>srnB</i> (F)	SrnB'	Membrane	ARN-SrnC	RNaseIII
Antisense-RNA-regulated systems	<i>hok/sok</i> (R1)	Hok	Membrane	ARN-Sok	RNaseIII

**Tableau 2.** Récapitulatif des systèmes d'accoutumance associés aux plasmides R1, F et P1 (adapté de Kobayashi, 2004, autorisation de reproduction en anglais seulement).

### I.3.3. Le plasmide ColE1

C'est un plasmide présent dans la bactérie *E. coli*, qui ne possède pas de mécanisme de partition actif, même s'il est habituellement présent en grand nombre de copies. Malgré ce manque de mécanisme actif pour sa partition, il est présent de façon stable à chaque génération (Summers and Sherratt, 1984).

On a constaté que c'était le site *cer* (pour ColE1 resolution) qui permettait cette stabilité au fur et à mesure des divisions cellulaires. En effet, la séquence *cer* permet la résolution des dimères du plasmide en monomères (figure 3). Comme il a été dit plus haut, il faut deux séquences *cer*, de 280 pb chacune, directement répétées pour générer la résolution du dimère en deux monomères.

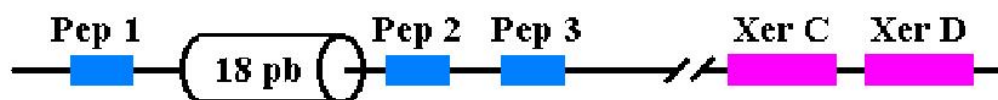


**Figure 3.** Représentation de la séquence *cer* sur le plasmide ColE1 (maîtrise Hélène Sénéchal).

Mais pour ce faire il faut appairer les deux séquences directement répétées. C'est ce que permettent les protéines ArgR, PepA, protéines accessoires à la recombinaison du plasmide médiée par les recombinases à tyrosine XerC et XerD. Il a été proposé que le surenroulement (favorisé dans ce cas par les protéines accessoires) de l'ADN favorise la formation d'un complexe de surenroulement actif, tant que les

séquences *cer* sont toujours orientées dans le sens direct. Sinon le rapprochement des séquences ne peut se faire (Craigie and Mizuuchi, 1986), le coût en énergie serait en effet trop important.

#### I.3.4. La séquence *cer*



**Figure 4.** Schéma de la séquence *cer*, avec les séquences où PepA est supposée interagir avec l'ADN (les séquences nommées Pep1, Pep2 et Pep3), les séquences spécifiques à XerC et XerD. Enfin, nous avons les 18 pb correspondant à la demie-boîte arginine (maîtrise Hélène Sénéchal).

Comme indiqué précédemment, la séquence *cer* a été découverte dans le plasmide ColE1 et la dimérisation du plasmide entraîne un dédoublement de la séquence *cer*. A ce moment précis une séquence particulière se retrouve dédoublée, c'est une séquence de 18 paires de bases qui est identique à une demie-boîte arginine, donc on restitue la boîte arginine en entier (c'est-à-dire deux fois les 18 pb). Cette boîte arginine est la séquence reconnue par la protéine ArgR, comme par exemple, lors de la répression des gènes de synthèse de l'arginine (figure 4). Cette dimérisation du plasmide, donc de *cer*, permet le recrutement et la fixation de l'hexamère d'ArgR sur lequel va s'enrouler l'ADN dimérisé.

À l'extrémité terminale en 5' de *cer*, il y a une séquence de 30 pb permettant la fixation des recombinases XerC et XerD (Blakely *et al*, 1993), les deux régions de fixation de XerC et XerD sont séparées de 8 pb. En plus de la demie-boîte arginine présente dans la séquence, il y a des séquences sur lesquelles la protéine PepA pourrait se fixer, les séquences Pep1, Pep2 et Pep3. De plus, la demie-boîte arginine se trouve environ à 100 paires de bases du site de recombinaison (Van Duyne *et al*, 1996). Il existe un promoteur *cer* (*Pcer*) lui-même inclus dans la séquence *cer*. Il permettrait un

héritage stable de ColE1 au fur et à mesure des divisions cellulaires, mais son inactivation n'empêche pas la résolution des dimères de ColE1 en monomère (Patient and Summers, 1993 ; Summers and Sherratt, 1988).

*P<sub>cer</sub>* entraîne la synthèse d'un ARN de 75 nucléotides nommé Rcd ressemblant au niveau de sa structure secondaire à un ARN antisens. Il a été décrit que Rcd bloque le cycle cellulaire tant que sa concentration est élevée dans la cellule, ce qui arrive lorsqu'il y a beaucoup de multimères, permettant par ce blocage du cycle cellulaire de résoudre complètement le problème des multimères (et donc d'obtenir des monomères) (Patient and Summers, 1993). On suppose, que Rcd serait activé lors de la multimérisation de ColE1 par alignement des séquences situées en -35 et -10 de *P<sub>cer</sub>*, dû à la distorsion de l'ADN entre les deux *cer*, qui à ce moment précis porte le complexe synaptique. Tous ces événements permettraient à l'ARN polymérase de se fixer sur le promoteur, et de synthétiser Rcd.

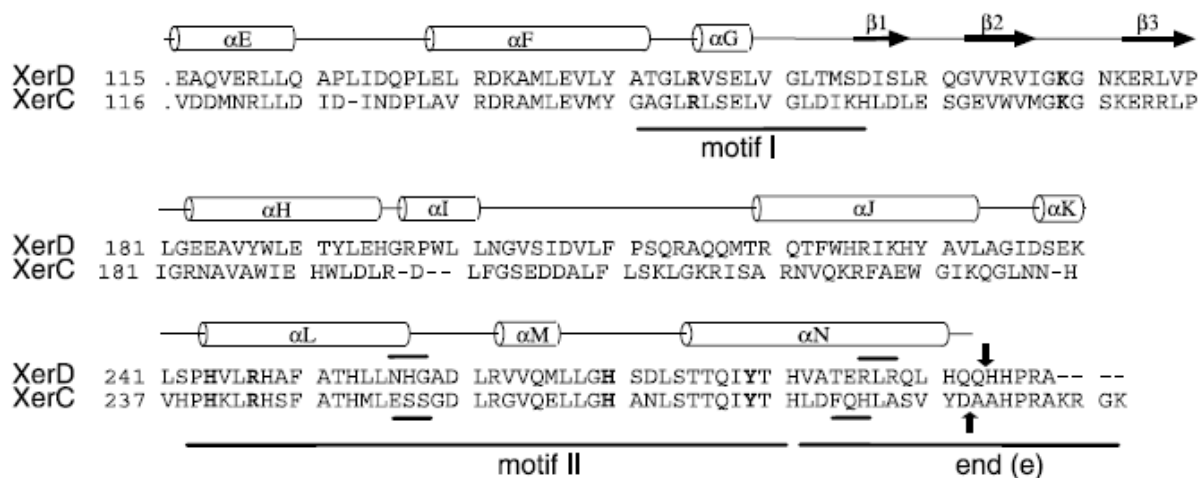
En 1999, on se demandait si Rcd s'appariait avec une protéine, mais là, étant donné que les multimères de ColE1 se retrouvent au centre de la cellule lors de la réplication (comme tous les plasmides), on se demande si une protéine du septum, ou une protéine responsable de la division cellulaire pourrait le faire (Chatwin and Summers, 2001). En 2009, Blaby *et al* ont proposé que la transcription de la protéine Rcd (dont le promoteur *P<sub>cer</sub>* est placé sur le plasmide ColE1) était déclenchée lorsque les dimères de ColE1 se forment et retirent les protéines XerC et XerD d'un protomère, ainsi que la protéine FIS.



## I.4. Les recombinases à tyrosine XerC et XerD.

### I.4.1. Généralités

Ces deux recombinases d'origine chromosomique font partie de la vaste famille des tyrosines recombinases caractérisées par un motif d'acides-aminés conservé au sein de leur site actif, le motif RKHRHY. On retrouve ce motif également chez les intégrases du phage  $\lambda$ , une famille de plus de 130 protéines (dont nos deux recombinases font partie), où le résidu tyrosine (Y) est le résidu catalytique. Les autres résidus sont responsables de la stabilité du complexe, voire de la recombinaison elle-même. Lorsque ces deux protéines effectuent une recombinaison sur des sites spécifiques (tels *dif* et *cer*), elles s'apparient pour former deux hétérodimères XerC/XerD qui réalisent une jonction d'Holliday, avant de résoudre la recombinaison et séparer les dimères en deux monomères, et ce de manière indépendante du système RecA. Au niveau de la structure tridimensionnelle de la protéine XerD, on constate que les régions Ct et Nt forment deux parties bien distinctes. La partie Nt des deux protéines Xer comporte deux hélices  $\alpha$  (résidus 1 à 107) qui interagiraient avec le sillon majeur de l'ADN des sites de liaison pour les protéines XerC et XerD. Ces considérations sur le rôle de cette partie des deux protéines Xer sont extrapolées de la structure de Cre et l'intégrase du phage  $\lambda$ . La partie Ct (à partir du résidu 108) serait nécessaire pour l'interaction entre les protéines et contient les résidus catalytiques nécessaires à la recombinaison (figure 5). Cette partie permettrait la mise en place du résidu tyrosine responsable de la catalyse de la recombinaison. (Spiers and Sherratt, 1999). Même si la catalyse est générée sans que XerC ait besoin de XerD, les deux parties Ct des protéines permettent l'échange des brins d'ADN. Ce serait aussi le domaine Ct qui comporterait des résidus d'acides-aminés (R221 et Q222) permettant cette liaison au niveau du sillon majeur de l'ADN des séquences *dif* et *cer*. On constate au sein de la famille des intégrases  $\lambda$  que la partie Ct de ces intégrases est capable d'adopter des conformations différentes et précises afin de reconnaître leur site d'action (Subramanya *et al*, 1997 ; Ferreira *et al*, 2003).



**Figure 5.** Alignement des parties Ct entre les protéines XerC et XerD chez *Escherichia coli*. Les cylindres représentent les hélices  $\alpha$  et les flèches les feuillet  $\beta$ . On remarquera que les motifs sont conservés. Dans l'article de Ferreira *et al*, 2003, des protéines chimériques avaient été générées par échanges de séquences au niveau des motifs II et de la fin de la séquence qui sont soulignées. Les triplets d'acides-aminés NGH et ESS supposément localisés près des régions réceptrices des deux protéines, sont soulignés en noir. D'autres triplets comme les triplets FQH et RLR soulignés sur cette figure sont impliqués dans des expériences de délétions effectuées dans le même article d'où est tiré cette figure. Les flèches noires montrent des expériences de délétions dans la partie terminale des protéines. Il est important de constater que les résidus RKHRHY sont représentés en gras (avec permission de reproduction, Ferreira *et al*, 2003).

Les expériences menées par Ferreira *et al* avaient pour but de distinguer quelles parties de XerC et de XerD étaient impliquées dans le contrôle de la catalyse lors de l'interaction entre les deux protéines. Les chimères obtenues ont bien montré que la fin de la région Ct de chaque recombinase coordonnait la catalyse et permettait la résolution des dimères. Entre les deux protéines XerC et XerD et d'autres recombinases à tyrosine, le motif II est très conservé, suggérant son importance. Les différences les plus importantes à ce niveau entre XerC et XerD étant situées au niveau des triplets ESS et NHG, qui seraient compris dans la région réceptrice permettant l'interaction entre chaque recombinase. La région de fin des recombinases XerC et XerD notées end (e) sur la figure 5 est importante pour la catalyse de la réaction et la liaison à l'ADN. Dans cette région ce sont les résidus HPRA qui restent identiques, la protéine XerC étant plus longue que XerD de quatre nucléotides (Ferreira *et al*, 2003). Dans l'étude de Ferreira *et al* en 2003, des protéines XerC et XerD chimériques où le motif II et la fin

de la protéine avaient été échangés, avaient été générées. Lorsque la région de fin a été échangée, l'activité catalytique était altérée. En effet, XerC avec la fin de XerD (soit XerC[De]) ne pouvait plus faire de catalyse correctement. Ce qui semble indiquer qu'une interaction normale entre les deux domaines Ct de XerC et XerD est nécessaire au maintien de la jonction d'Holliday dans une conformation permettant l'activation de la catalyse par XerC et l'inhibition de la catalyse par XerD à cette étape de la recombinaison. De plus, XerD avec la fin de XerC (soit XerD[Ce]) ne stimulait pas XerC lors de la catalyse. Ce qui pourrait se traduire par une importance de ces résidus pour l'interaction avec la région réceptrice entre les deux protéines (Ferreira, *et al*, 2003). Les résultats des expériences de délétions menées sur les résidus de 280 à 298 conjugués aux résultats des différentes chimères de cette même étude montrent que les deux régions Ct des protéines XerC et XerD ne sont pas interchangeable. La région Ct de XerC permet l'interaction avec la région réceptrice de XerD, tandis que le motif II de XerC permet l'activité catalytique de XerD. Les mutations générées dans XerD, qui empêchent sa fonction catalytique, soutenant ce modèle précédemment évoqué. Ce sont d'ailleurs les délétions en position 294 à 298 sur XerD qui empêcheraient les interactions entre XerD et la région réceptrice de XerC (Ferreira *et al*, 2003). Les modifications générées en XerD interféreraient avec la capacité de XerC à interagir avec XerD et non l'inverse. En effet XerD est toujours capable d'interagir avec XerC (Ferreira *et al*, 2003).

Au niveau de la séquence *dif*, les deux protéines Xer se lient au niveau du terminus de réplication du chromosome bactérien. Au niveau de la séquence *cer*, les protéines XerC et XerD se lient sur deux séquences plasmidiques spécifiques en aval des séquences supposées pour PepA. Lors de la recombinaison, qu'elle soit chromosomique ou plasmidique, les résidus RKHRHY entrent en action. Les trois résidus RHR génèrent une « zone catalytique » activant le phosphodiester qui permet à la tyrosine (Y) de cliver l'ADN et de générer le premier intermédiaire covalent de la jonction d'Holliday.

Ce premier intermédiaire se situe en 3'-phosphotyrosyl. C'est après que la partie 5'-OH, libre et générée par le premier clivage, fait une attaque nucléophile sur le 3'-phosphodiester cité plus haut. On suppose que c'est la lyse de la boucle  $\beta$  qui protonnerait le groupement 5'-OH pendant l'attaque nucléophile (Redinbo *et al*, 2000 ; Krogh *et al*, 2000). Ces étapes de réaction existent pour les deux protéines XerC et XerD. Que ce soit chez les eubactéries ou les archéobactéries, les protéines Xer sont conservées, mais il existe un à trois gènes *xer* selon les espèces.

Le tableau suivant (tableau 3) donne quelques espèces bactériennes qui ont le gène *xer* en un nombre différent d'exemplaires.

Bactéries	Nombres de gènes <i>xer</i>	Présence d'un homologue de FtsK
<i>Escherichia coli</i>	2	+
<i>Bacillus subtilis</i>	2	+
<i>Helicobacter pylori J99</i>	3	+
<i>Haemophilus influenzae</i>	2	+
<i>Caulobacter crescentus</i>	2	+
<i>Thermotoga maritima</i>	1	-
<i>Proteus mirabilis</i>	2	+
<i>Lactococcus lactis</i>	1	+
<i>Streptococcus suis</i>	1	+

**Tableau 3.** Quelques espèces de bactéries avec le nombre de gènes *xer* présents dans leur génome, ainsi que la présence d'un homologue de la protéine FtsK (adapté d'Hélène Sénéchal).

Dans cette thèse, nous nous intéressons surtout aux Xer impliquées dans la recombinaison *cer* spécifique de ColE1, bien qu'il existe d'autres sites pour XerC et XerD chez *E.coli* et différentes espèces. Chez les entérobactéries, par exemple, on a retrouvé par hybridation avec des sondes contenant les fragments des gènes *xerC* et

*xerD*, des homologies avec les gènes *xerC* et *xerD* chez *Klebsiella pneumoniae*, *Morganella morganni*, *Salmonella typhimurium* et *Shigella flexneri*, et ce au niveau de leurs chromosomes. D'autres entérobactéries comme *Serratia liquefaciens*, ne possèdent que des homologies avec des séquences plasmidiques pour *xerC* et *xerD*. Enfin, *Proteus vulgaris* ATCC 13315 n'a montré d'homologie qu'avec *xerD*, et seulement au niveau d'une séquence plasmidique (Sirois and Szatmari, 1995). De plus dans les bactéries de types streptocoques et lactocoques, il existe un seul *xerS* qui se lie à une séquence qui à la même fonction que la séquence *dif* appelée *dif<sub>sl</sub>* (Le Bourgeois et al, 2007 ; Novillos et al, 2010).

#### **I.4.2. recombinaison en *dif***

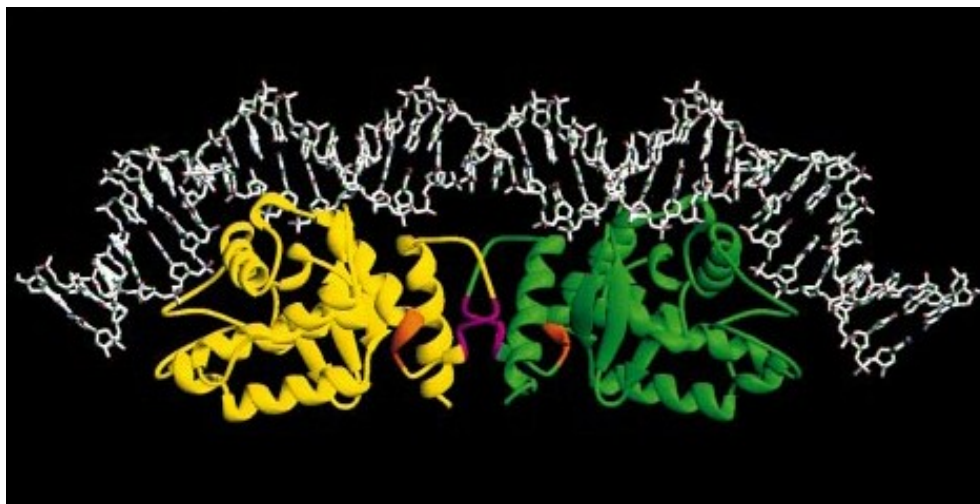
Le site *dif* est la séquence d'ADN qui, chez *E.coli*, est localisée dans la région de terminaison de la réplication du chromosome bactérien. C'est à ce niveau que les dimères du chromosome vont se séparer (Blakely *et al*, 1991 ; De Massy *et al*, 1987).

Le site *dif*, une séquence de 28 pb chez *E.coli*, est situé sur la carte du chromosome à 33,6 min en position 1608 de la carte physique du chromosome bactérien. Il est à noter que cette séquence se trouve entre les régions *terA* et *terC* (Kuempel *et al*, 1991 ; Barre et Sherratt *et al*, 2002), donc le site se trouve proche de la séquence du terminus de réplication. La séquence se décompose en deux demi-sites séparés par un espaceur de 6 pb (voir figure 19).

Pour être active, la séquence *dif* doit être impérativement insérée dans une région très précise, la *dif activity zone* ou DAZ (Perals *et al*, 2000 ; Perals *et al*, 2001). En effet, c'est dans cette zone précise que la recombinaison a lieu.

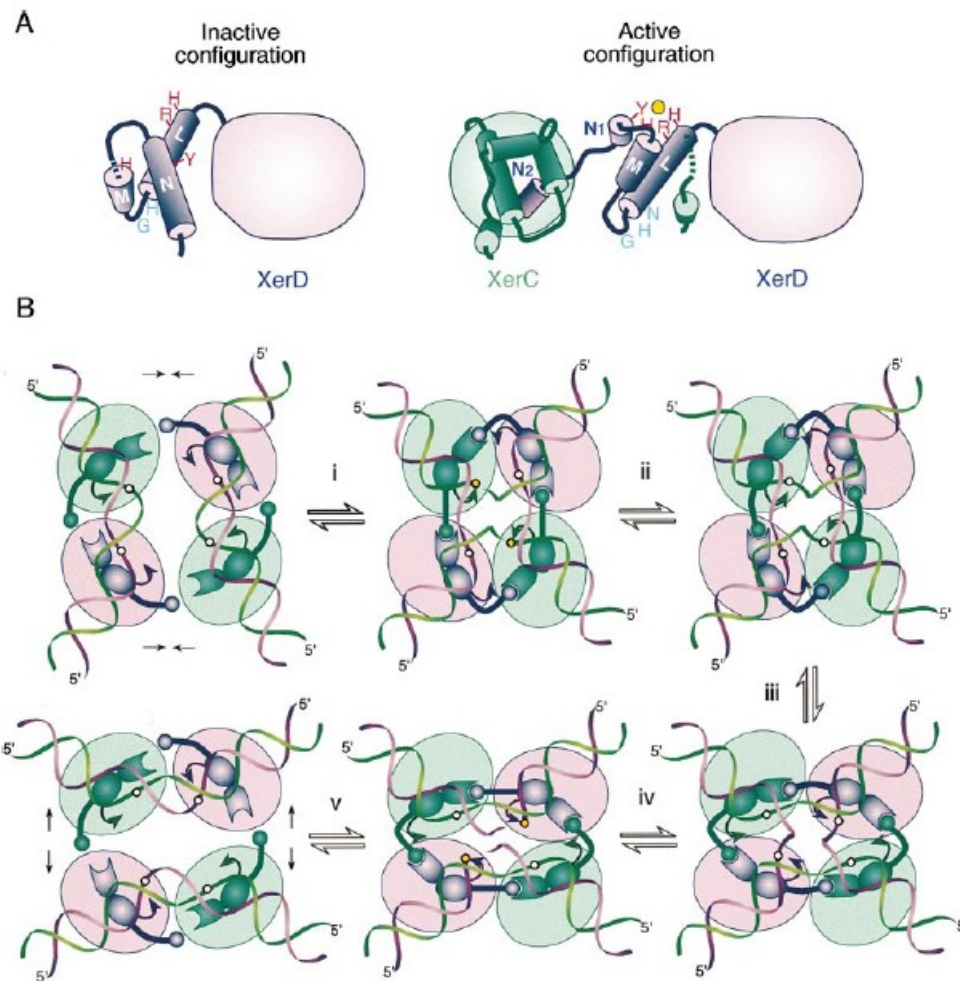
On remarquera que l'origine de réplication *oriC* est diamétralement opposée à *dif*. Les séquences entourant *dif* sont polarisées tout le long de l'axe *oriC-dif* et leurs orientations relatives déterminent le positionnement de la DAZ. De plus, on a montré

que la délétion des séquences entourant *dif* n'entraîne pas de problème pour résoudre les multimères plasmidiques, alors que leurs inversions empêchent la résolution des dimères (Tecklenburg *et al*, 1995 ; Cornet *et al*, 1996 ; Pérals *et al*, 2000). Les séquences déterminantes pour la polarisation sont présentes tout le long d'un large domaine terminal (plus de 200 pb autour de *dif*) et sont hautement répétées. On s'est aperçu que les DAZ étaient polarisés et que chacun des bras *oriC-dif* possédait une polarisation dans la direction opposée que l'on retrouve au niveau de la séquence d'ADN. Les séquences entourant le site *dif* possèdent également cette polarisation (figure 6 et 7).



**Figure 6.** Modèle d'interaction entre le site *dif* et les protéines XerC et XerD. Les régions d'interaction entre les deux protéines sont représentées en orange (résidus 256-258) et en magenta (résidus 263-267) (avec permission de reproduction, Subramanya *et al*, 1997).

Des motifs d'ADN appelés FtsK *orienting polar sequences*, soit KOPS, y ont été identifiés, ils permettent de diriger le mouvement de la FtsK translocase vers la région de terminaison, grâce à la séquence consensus GNGNAGGG, qui est présente sur tout le chromosome bactérien, mais de façon biaisée (Bigot *et al*, 2005 ; Levy *et al*, 2005), notamment autour du réplicore. L'effet des séquences KOPS sur la translocation est dit stochastique, c'est-à-dire qu'il faut deux ou trois séquences KOPS pour avoir un fort effet (Bigot *et al*, 2005 ; Levy *et al*, 2005).

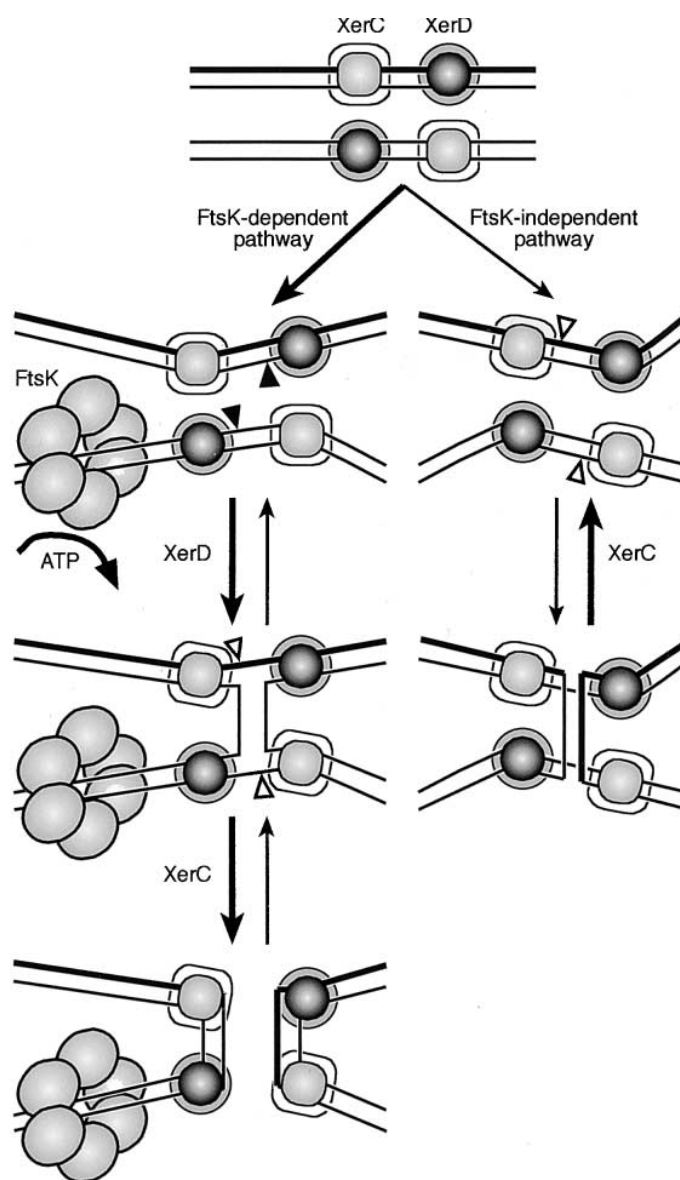


**Figure 7.** Recombinaison médiée par XerC et XerD :

**A)** Lors de la recombinaison la protéine XerD est activée par XerC lorsque leurs deux extrémités Ct sont en contact. Les résidus catalytiques RHYH sont représentés en rouge, le cercle jaune indique l'emplacement que prendra le résidu phosphate lors de la réaction.

**B)** Lors de la première étape en haut à gauche de cette partie, les protéines XerC et XerD sont d'abord liées à leur séquence d'ADN, puis elles entrent en contact l'une avec l'autre, permettant à la première réaction notée i, de se faire, c'est-à-dire le premier échange de brin (coloré en vert) catalysé par la protéine XerC. L'étape ii montre la fin de l'étape précédente où il y avait eu torsion de l'ADN pour aboutir à la formation de la jonction d'Holliday, désormais présente, après attaque du résidu phosphate. L'étape iii montre le changement de conformation des protéines XerC et XerD rapprochant le second brin (en violet) qui subira la catalyse après sa torsion. L'étape iv montre qu'il y a eu attaque nucléophile sur le brin par XerD, à ce niveau, la jonction d'Holliday est résolue et la recombinaison site spécifique terminée. L'étape v montre la dissociation du complexe XerC-XerD, qui sera suivi de la dissociation de ces protéines de l'ADN (avec permission de reproduction, Hallet *et al*, 1999).

Lors de cette recombinaison chromosomique, c'est donc la protéine FtsK qui est la protéine accessoire pour que XerC et XerD finissent par résoudre le dimère généré lors de la réplication, avant la division cellulaire. Dans cette réaction, c'est la protéine XerD qui génère la première attaque sur le dimère pour générer une jonction d'Holliday, c'est XerC qui termine la recombinaison. Cela est permis par au moins un hexamère de FtsK grâce à l'hydrolyse de l'ATP (Aussel *et al*, 2002). Par contre, lorsqu'il n'y a pas de protéine FtsK, comme au niveau du site *cer*, c'est la protéine XerC qui génère la première attaque sur le brin d'ADN (figure 8).



**Figure 8.** Recombinaisons FtsK dépendantes et FtsK indépendantes au niveau du site *dif*. Sur ce schéma, il a été indiqué qu'au moins un hexamère de FtsK est nécessaire pour résoudre les dimères chromosomiques, cette voie étant d'ailleurs majoritaire, sinon le complexe synaptique comprenant les protéines Xer adopte une conformation favorisant un échange de brin par XerC. Si XerC effectue la première attaque, on revient au produit de départ. Dans la résolution des dimères, FtsK peut utiliser l'énergie de l'ATP pour donner une conformation au complexe synaptique permettant l'échange de brin par XerD. À noter que cette réaction pourrait mettre en jeu plusieurs complexes FtsK, qui pourraient se lier à plusieurs bras de cette jonction, les flèches montrent les réactions qui ont le plus de chances de se produire en présence de FtsK. Les sites de clivage pour XerC et XerD sont montrés par des triangles blancs et noirs, respectivement (avec permission de reproduction, Aussel *et al*, 2002).



### I.4.3. recombinaison en *psi*

La séquence *psi* (pSC101 stabilized inheritance) est une séquence plasmidique que l'on trouve sur le plasmide pCS101, entre les positions 6783 et 6810. Les séquences *psi* et *dif* présentent de nombreuses similitudes au niveau des sites de liaison pour XerC (deux nucléotides différents) et XerD (un nucléotide différent).

*Psi* se compose d'une partie principale de 28 pb sur laquelle les recombinaisons XerC et XerD peuvent se fixer pour réaliser la résolution des dimères de pSC101 par formation d'une jonction d'Holliday (Colloms *et al*, 1996 ; Blake *et al*, 1997), la distance séparant les deux sites de liaison pour les recombinaisons est de 6 pb. La séquence *psi* (voir la séquence sur la figure 19), séquence homologue à *dif*, peut même remplacer *dif* sur le chromosome, et permettre de résoudre les dimères chromosomiques au moment de la réplication. Si on insère la séquence *psi* dans un vecteur dérivé de p15A, et si l'on insère cette construction à la place de *dif*, il y a des échanges intermoléculaires entre le chromosome et le vecteur. On est amené à dire que les séquences accessoires sont nécessaires et essentielles à la recombinaison site spécifique, tout comme la séquence *cer* (Cornet *et al*, 1994). Il y a aussi environ 200 pb de séquences pour les protéines accessoires PepA et ArcA, protéines accessoires de la recombinaison site spécifique à *psi*.

Il faut que la séquence *psi* soit directement répétée lorsqu'il y a dimérisation afin que la recombinaison puisse se faire. On remarquera que la protéine ArcA est essentielle pour la recombinaison *in vivo* mais pas *in vitro* (Colloms *et al*, 1998) au niveau du site *psi*.

Lors de cette recombinaison au niveau du site *psi*, le plasmide prend une structure de caténane à quatre nœuds (qui ne se trouve pas résolu *in vitro*) avec le concours des protéines XerC, XerD, PepA et ArcA. Il est à noter que sur *psi*, comme pour *dif*, c'est XerC que génère la jonction d'Holliday et XerD qui la résout (Colloms *et al*, 1996). ArcA a en fait sur la séquence *psi* un rôle similaire à ArgR sur la séquence *cer* (Stirling *et al*, 1998 ; Colloms *et al*, 1998).

Il est également à noter que si PepA se trouve en excès durant des expériences *in vitro*, la protéine peut se lier au site *psi* sans avoir besoin de la protéine ArcA. ArcA aurait sans doute pour rôle de faciliter la liaison et la formation du complexe de recombinaison au niveau du site *psi*.

#### **I.4.4. recombinaison en *cer***

Cette recombinaison est la recombinaison qui est étudiée en détail lors de ce chapitre. En effet, c'est cette recombinaison au niveau du site *cer* qui a lieu lors de la résolution des dimères de ColE1. C'est d'ailleurs après cette dimérisation, que la boîte arginine sera restituée en entier permettant le recrutement de la protéine ArgR sur l'ADN, la liaison se faisant au niveau de la partie Nt d'ArgR.

Mais il n'y a pas que la protéine ArgR qui soit recrutée par la dimérisation du site *cer*, en effet, la protéine PepA et les recombinases XerC et XerD seront recrutées par ce système de recombinaison site spécifique afin de résoudre les dimères. La protéine PepA aurait d'ailleurs des sites de liaison à l'ADN de la séquence *cer*, même si ces derniers sont surtout des sites supposés, les sites Pep1, Pep2 et Pep3, car il semblerait que les sites manqueraient de spécificité.

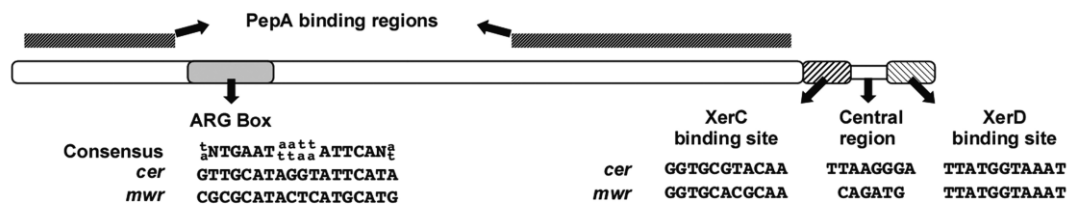
Lors de la recombinaison proprement dite, c'est la protéine XerC qui effectue la première attaque nucléophile sur l'ADN. Cette activation de la recombinase XerC n'est possible que si la protéine XerD est présente et en contact avec la protéine XerC. C'est ensuite grâce à l'attaque nucléophile de XerC sur l'ADN que la jonction d'Holliday se forme, mais XerD ne résout pas cette structure. C'est grâce à une ou plusieurs résolvases non identifiées que la résolution des dimères en monomères se réalise (Colloms *et al*, 1996).

L'objet de cette thèse étant justement de démontrer comment les protéines ArgR et PepA interagissent lors de cette recombinaison site spécifique, j'expliquerai plus en

détails les interactions supposées et hypothèses formulées lors de cette recombinaison pour la formation du complexe ADN-ArgR-PepA, dans les chapitres consacrés à ces deux protéines et aux interactions qu'elles peuvent avoir. Dans toutes les recombinaisons citées, les recombinases à tyrosines XerC et XerD interviennent, seuls les facteurs accessoires changent entièrement ou bien peu. Ces recombinases et les facteurs accessoires qui peuvent s'associer à elles vont être décrits dans les parties suivantes. De plus il a été mis en évidence que lorsque le plasmide ColE1 est monomérique, la protéine FIS réprime le promoteur de la séquence *cer* (Blaby *et al*, 2009).

#### **I.4.5. recombinaison en *mwr* et *nmr***

Cette recombinaison a été identifiée au niveau du site *mwr* contenu dans le plasmide pJHCMW1 découvert dans la bactérie *Klebsiella pneumoniae*. Ce site *mwr* est hautement homologue à *cer*, à tel point que les mêmes protéines accessoires ArgR et PepA sont utilisées lors de la recombinaison site spécifique en *mwr*. Les recombinases impliquées dans cette recombinaison homologue sont les mêmes, à savoir XerC et XerD. Comme en *cer*, il y a formation d'une jonction d'Holliday. La différence majeure entre les plasmides ColE1 et pJHCMW1, portant respectivement *cer* et *mwr*, est que la recombinaison du plasmide possédant *mwr* est gênée par une augmentation de la concentration en NaCl, ce qui n'est pas le cas de ColE1. En effet, la concentration en NaCl influe sur la topologie du plasmide pJHCMW1 en augmentant le nombre de supertours négatifs, plus on augmente la concentration en NaCl, plus il y a apparition de ces supertours, gênant la recombinaison. Ce plasmide est tellement sensible à la concentration en NaCl que l'on est obligé de travailler en milieu plus pauvre que le L broth pour avoir une recombinaison optimum (Trigueros *et al*, 2009). Il est important de dire que la boîte arginine retrouvée au niveau de la séquence *mwr* ne présente pas une aussi bonne affinité pour ArgR que celle présente au niveau de la séquence *cer* (Trigueros *et al*, 2009) (figure 9).



**Figure 9.** Schéma de la structure des sites *cer* et *mwr*. Ces deux sites sont des sites de recombinaison pour les recombinases XerC et XerD et les facteurs accessoires ArgR et PepA. Sous ce schéma, sont indiquées les boîtes arginine des séquences *cer* et *mwr*, ainsi que les sites de liaison pour les protéines XerC et XerD présents sur ces deux mêmes séquences. On remarquera, que les sites de liaison sont de tailles similaires, mais que la séquence *mwr* diffère de *cer*, et l'espaceur entre les sites de liaison pour XerC et XerD possède 8 pb pour *cer* et 6 pb pour *mwr* (avec permission de reproduction, Trigueros, *et al*, 2009).

D'autres systèmes de recombinaison utilisent des séquences proches de *cer*, recrutant les recombinases XerC et XerD. Citons en plus de *mwr*, la séquence *hmr* du plasmide pHS-2 de *Shigella flexneri* qui permet la résolution des multimères. Pour la bactérie *Salmonella typhimurium*, l'homologue de *cer* est nommé *nmr* et se retrouve aussi sur le plasmide NTP16 (figure 10) (Zakova and Szatmari, 1995 ; Réhel *et al*, 1996). Ces deux systèmes recrutent les mêmes protéines que la séquence *cer* lors de leurs recombinaisons sites spécifiques, à savoir les protéines ArgR, PepA, XerC et XerD.

```

cer      1      <---Arg-box--->
          1      TTGCATAGGTATTCATACGGTTAAAATTTATCAGGCGCGATCGCGGCAGT      50
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
nmr      1      ATGCATAGCTATGCAGTGAAGCTGAAAGCGATCCTGACGCATTTTTCGGGT      50
          <---Arg-box--->

cer      51      TTTCGGGTGGTTTGTTGCCATTTTACC-TGTCGTCTGCCGTGATCGCGCT      100
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
nmr      50      TACCCCGGGGAAAACATCTCTTTTTCGGGTGTCTGCTGCAG-AATCGCGTT      100
          SmaI

cer      101.   MluI < left arm > OVERLAP < right arm >
          GAACGCGTTTTAGCGGTGCGTACAATTAAGGGATTATGGTAAATCCACTT      150
          * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
nmr      100   CAGCGCGTTTTGCGCGTGCAGTAA-TGAGACGTTATGGTAAATGTCTTC      149
          < left arm > OVERLAP < right arm >

```

**Figure 10.** Comparaison entre les séquences d'ADN de l'extrémité 3' de 150 pb des séquences *cer* et *nmr*. Les homologies de séquences sont identifiées par des astérisques, les sites de liaison à l'arginine (Arg-box) par des flèches, les sites d'enjambement situés hors de la boîte arginine sont indiqués en gras. De plus, les sites de restriction pour MluI et SmaI sont indiqués (avec permission de reproduction, Zakova and Szatmari, 1995).

### I.4.6. XerC

Chez *E.coli* c'est le gène *xerC* situé sur le chromosome bactérien qui code pour la protéine XerC (298 aa, 33,8 kDa). Il fait partie d'un polycistron à quatre cadres de lecture ouvert, c'est d'ailleurs le troisième gène de l'ensemble (Colloms *et al*, 1990).

Ce polycistron comporte les gènes *dapF*, *yigA*, *xerC*, *yigB*. Le gène *dapF* code la diaminopimelate épimérase, enzyme comprise dans les voies de biosynthèse de la lyse et du peptidoglycane de la membrane (figure 11).

Les gènes *yigA* et *yigB* sont de fonctions inconnues.

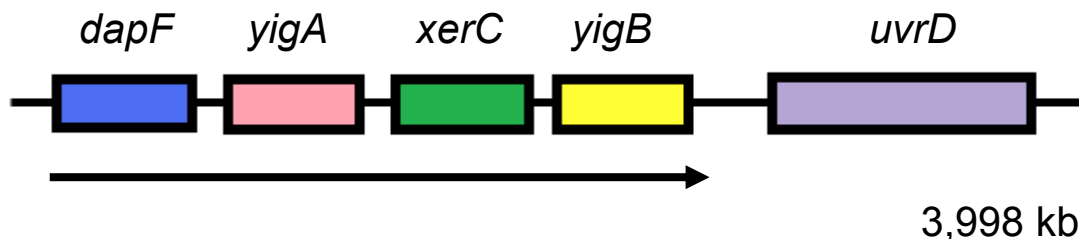
La protéine XerC comprend quatre résidus très conservés permettant la recombinaison site spécifique (Colloms *et al*, 1990).

- L'arginine en position 148 du domaine I
- L'histidine en position 240
- L'arginine en position 243
- La tyrosine catalytique en position 275 dans le domaine II.

Toute mutation en ces emplacements entraîne une perte de l'activité de recombinaison (Arciszewska *et al*, 1995 ; Cornet *et al*, 1997).

XerC est la protéine qui se lierait au sillon majeur de l'ADN par sa partie Nt sur 11 pb, que ce soit sur les sites *cer* ou *dif*. Mais on a remarqué que cette liaison à l'ADN était faible et avait besoin du concours de XerD pour être plus stable. C'est tout d'abord XerC qui clive en premier le brin supérieur de l'ADN permettant la formation de la jonction d'Holliday. C'est ensuite à XerD de terminer la réaction de recombinaison.

Ces données seraient vraies pour le site *cer*, mais pour le site *dif*, ce serait la protéine XerD qui initierait le clivage (Aussel *et al*, 2002 ; Hallet *et al*, 1999).



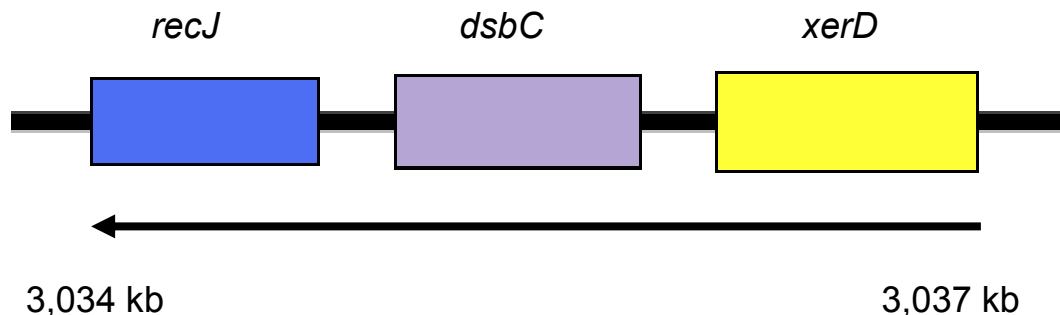
**Figure 11.** Représentation schématique du polycistron contenant le gène *xerC*. Ce polycistron d'*Escherichia coli* comporte quatre ORF, celles de *dapF*, *yigA*, *xerC* et *yigB*. Il y a un autre gène qui suit ce polycistron, le gène *uvrD*.

#### I.4.7. XerD

Chez *E.coli* c'est le gène *xerD* situé sur le chromosome bactérien qui code pour la protéine XerD faisant aussi partie de la famille des recombinases à tyrosines, dont fait partie le phage  $\lambda$  (297 aa, 34 kDa).

Ce gène, préalablement nommé *xprB*, génère un produit permettant chez *E.coli* de faire des recombinaisons sites spécifiques et de résoudre, avec la protéine XerC, la séparation des deux chromosomes fils après la réplication. Comme *xerC*, le gène *xerD* est compris dans une unité de transcription polycistronique.

Cette dernière comporte trois gènes différents (Lovett *et al*, 1991). Le premier (par ordre de transcription) est *xerD*, le second dénommé *dsbC*, code pour une oxydoréductase périplasmique et le dernier *recJ*, code pour une exonucléase (figure 12).



**Figure 12.** Représentation schématique du polycistron contenant *xerD*. Sur ce schéma, les trois ORF codant les gènes *xerD*, *dsbC* et *recJ* y ont été indiquées. Cette organisation des gènes existe chez *Escherichia coli*.

Tout comme XerC, XerD se lierait au sillon majeur de l'ADN par sa partie Nt sur 11 pb, mais y introduit une courbure comprenant un angle de 40 degrés qui autorise XerC à avoir une plus forte liaison avec l'ADN. Si XerC coupe le brin supérieur lors de la recombinaison, XerD coupe le brin inférieur après l'action de XerC (Sherratt *et al*, 1995 ; Colloms *et al*, 1996 ; Arciszewska *et al*, 1995).

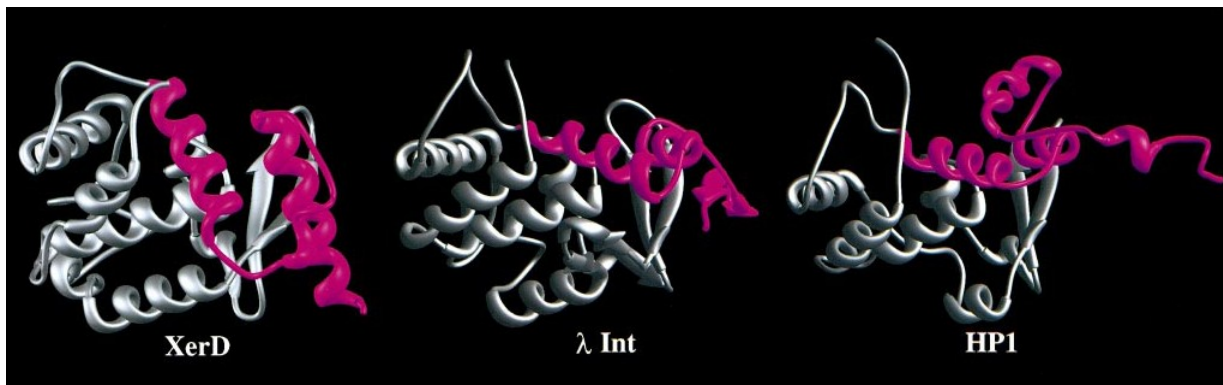
Sa structure, comme XerC, possède deux domaines. Un domaine N-Terminal comprenant les résidus 1 à 107 permettant la liaison à l'ADN, et une partie C-Terminale avec les résidus 108 à 297. Dans cette dernière partie, il y a également présence en position 148 d'une arginine permettant la liaison à l'ADN, même si la partie Nt peut se lier à la molécule sans son intervention.

C'est dans cette partie Ct que les résidus catalytiques sont présents :

- H en position 244
- R en position 247
- tyrosine en position 279.

Lors de la recombinaison au niveau du site *dif*, XerD initierait le clivage lorsque la protéine accessoire FtsK est présente, même en absence de XerC (Hallet *et al*, 1999). Lors de l'interaction entre XerC et XerD, cette dernière interagirait au niveau de sa partie Ct avec XerC, surtout grâce aux acides-aminés compris entre les positions 263 et 268. De plus, c'est XerD qui met un terme à la recombinaison (Spiers *et al*, 1999 ; Neilson *et al*, 1999).

Au niveau de la partie Ct de XerD, on sait que les acides-aminés présentent deux motifs de séquences conservées qui sont également présentes chez les intégrases  $\lambda$  et HP1, même si les séquences entre les trois protéines sont différentes dans ces régions. De plus, dans la protéine XerD l'arrangement des résidus d'acides-aminés permet un mécanisme de clivage en *cis* (figure 13).



**Figure 13.** Structure du domaine Ct de XerD.

La partie en magenta va du domaine II à la partie Ct. C'est cette partie qui présente le plus de différence avec les autres régions similaires des intégrases de la famille  $\lambda$ , c'est-à-dire l'intégrase du phage  $\lambda$  et la protéine HP1 (avec permission de reproduction, Subramanya *et al*, 1997).

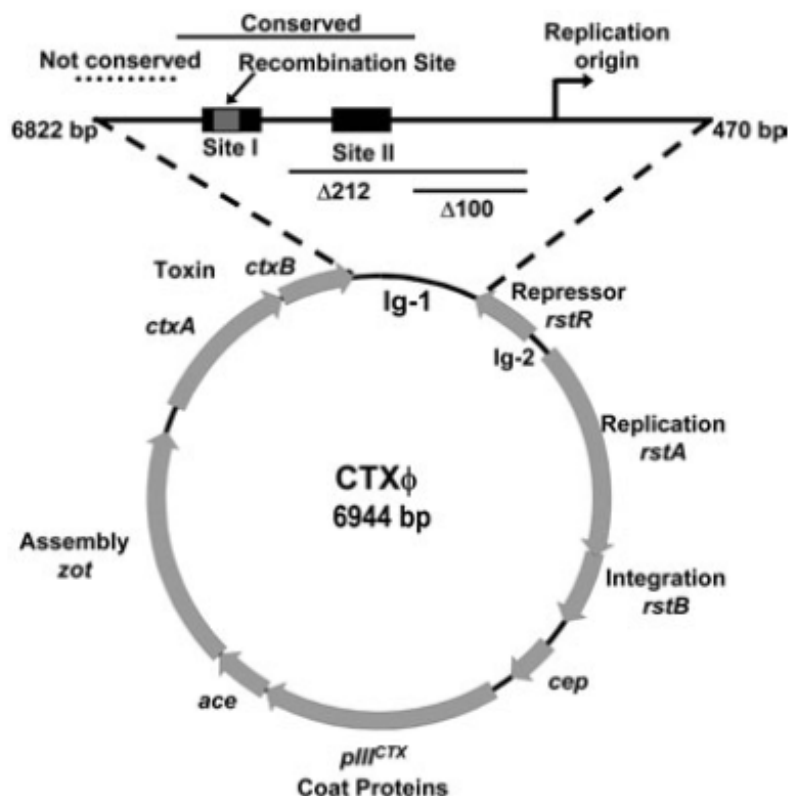
#### I.4.8. XerC et XerD et *Vibrio cholerae*

On s'est aperçu que le phage filamenteux  $\phi$ CTX, également noté CTX $\phi$  (codant la toxine du choléra via le gène *ctxAB*) d'une taille de 6944pb, s'intégrait dans le génome bactérien de *Vibrio cholerae* en utilisant les protéines XerC et XerD de cette bactérie Gram négative (Huber and Waldor, 2002) (figure 14). C'est cette intégration dans le génome du grand chromosome de *Vibrio cholerae* qui rend la bactérie pathogène. Si le site d'intégration du phage n'est pas présent, le génome de  $\phi$ CTX se réplique de manière autonome comme un plasmide et est nommé pCTX. On savait aussi que la protéine RstB codée par le gène d'intégration du phage *rstB* avait peu d'homologie avec les autres facteurs d'intégration connus.

Durant les études sur le phage CTX $\phi$ , on s'est aperçu que son intégration était similaire à celle du phage  $\lambda$ , C'est-à-dire qu'elle a lieu sur un site spécifique avec



recombinaison entre le site bactérien « d'attachement » *attB* et le site phagique *attP*. Cette intégration laisse apparaître deux nouveaux sites, une jonction gauche *attL* et une jonction droite *attR*. L'une des plus importantes particularités de CTX $\phi$  est qu'il ne code pour aucune recombinase contrairement à des phages comme le phage  $\lambda$  alors que sa façon de s'intégrer dans le génome bactérien est d'un premier abord assez similaire. En fait, CTX $\phi$  utilise les recombinases XerC et XerD contenues dans la bactérie qui l'héberge.



**Figure 14.** Schéma du phage CTX $\phi$ . Sur ce schéma, les phases de lecture ouvertes sont représentées en gris. Les gènes *ctxA* et *ctxB* codent pour la synthèse de la toxine cholérique. La région *Ig-1* contient deux sites de liaison pour XerC et XerD qui sont ici représentés en noir, bien qu'il y ait une région privilégiée pour cela, dans le site I, il s'agit de la portion grisée sur ce schéma. Ces deux sites forment d'ailleurs une région conservée, le site du gène *rstB* est également montré (avec permission de reproduction, McLeod and Waldor, 2004). Sur ce schéma, deux annotations  $\Delta 212$  et  $\Delta 100$  ont été soulignées, elles correspondent à deux délétions pratiquées dans une expérience de McLeod and Waldor, montrant que le site II est obligatoire pour que le phage CTX $\phi$  s'intègre.

On sait que le phage s'intègre dans la bactérie sous forme d'ADN circulaire monobrin. En ce qui concerne les XerC et XerD du *V.cholerae*, il y a respectivement 53% et 68% d'homologie avec les XerC et XerD d'*E.coli*.

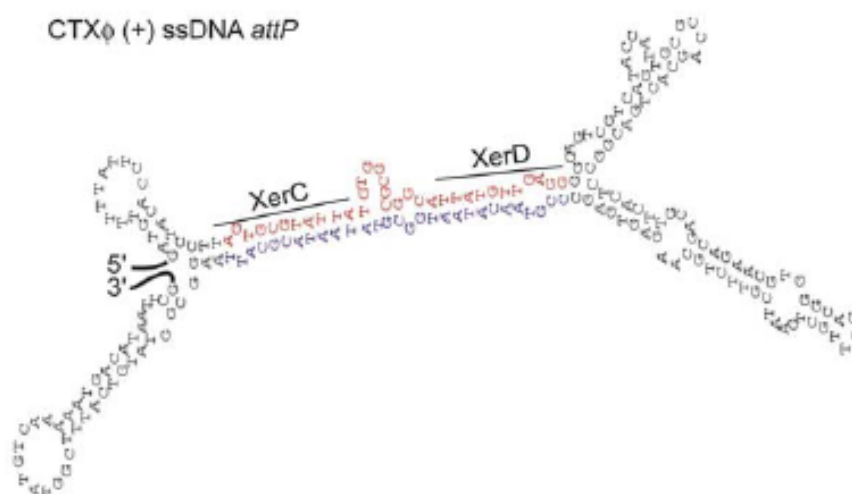
Le site d'intégration *attB* est situé à 180° de l'origine de réplication du chromosome I noté chrI. Or il a été démontré par Huber and Waldor en 2002 qu'*attB* contenait le site *dif* de *V.cholerae*, mais l'intégration du phage se fait sans passer par l'intermédiaire des dimères de chromosome (McLeod and Waldor, 2004). Bizarrement, le phage ne s'intègre pas dans le chromosome II (noté chrII) de *V.cholerae*. En fait, il semblerait que la configuration requise pour les sites de liaison des recombinases XerC et XerD lors de l'intégration du phage soit plus complexe que dans d'autres types de recombinaisons sites spécifiques médiées par recombinases. En effet, les sites I et II sont obligatoires pour l'intégration du phage dans le chromosome. Il y aurait participation d'autres facteurs provenant de la bactérie hôte du phage.

Dans ce système, XerC et XerD se lient de façon adjacente l'une à l'autre et de manière coopérative aux sites *attB* et *attP*, comme pour les sites *cer*, *dif* et *psi*. Si XerC et XerD se lient sans problème sur *attB*, il faut que XerC soit présente pour que la protéine recombinase XerD puisse se lier au site *attP*. *In vitro*, les clivages de séquences *attB* et *attP* montrent que le brin inférieur de l'ADN des sites était en dehors de la région d'échange du brin.

Il est à noter que les deux sites I et II de la séquence *attP* sont dans une orientation inversée l'un par rapport à l'autre et qu'ils sont séparés de 80 pb. Fait tout aussi remarquable, le site II, bien qu'étant capable de lier les protéines XerC et XerD, c'est au niveau du site I que la recombinaison a lieu.

Il semblerait que l'apparente complexité du système s'expliquerait par le fait que c'est la forme monobrin du phage CTX $\phi$  et non la forme bicaténaire, qui s'intègre de manière définitive dans le chromosome bactérien.

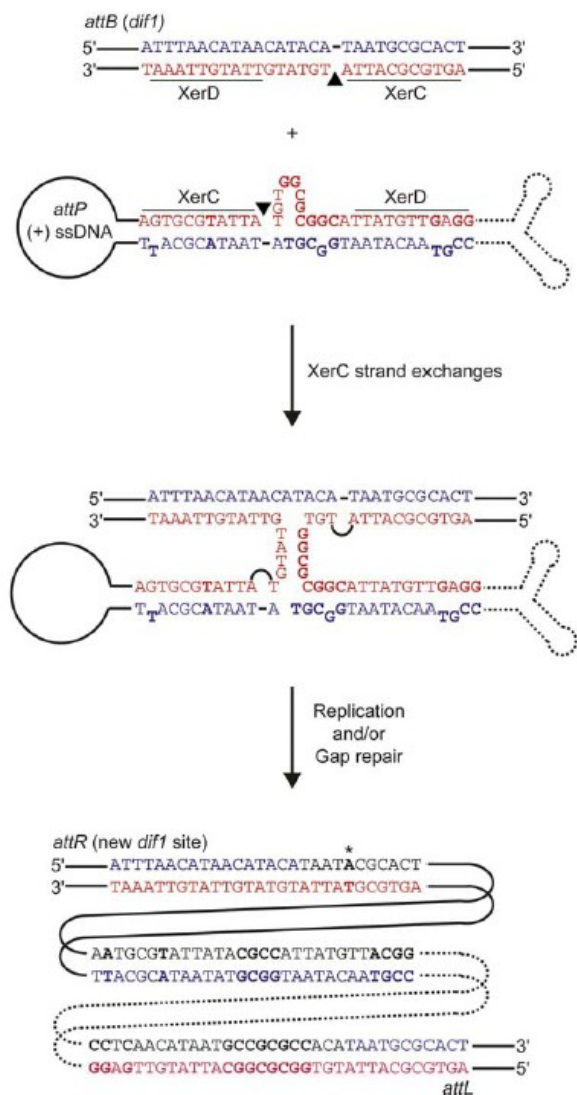
C'est dans les travaux de Val *et al* en 2005 qu'il a été mis en évidence que la forme monobrin pouvait prendre une forme en épingle à cheveux au niveau de la séquence *attP*. En effet, le site I et le site II sont en sens inverse et palindromiques, et lorsque l'ADN se replie en forme d'épingle à cheveux, les sites de liaison pour XerC et XerD sont reconstitués, et surtout fonctionnels (figure 15). En effet, le site reconstitué peut fonctionner avec le site *difI* de la séquence *attB*, mais aussi avec le site *dif* d'*E.coli* (Val *et al*, 2005).



**Figure 15.** Repliement de l'ADN monobrin (soit ssDNA sur la figure) du phage CTXφ, au niveau du site *attP*. Sur ce schéma, le site I et le site II reforment des séquences de liaison pour les protéines XerC et XerD. De plus, cette séquence est tout-à-fait capable d'effectuer une recombinaison site spécifique avec les recombinases XerC et XerD. À noter qu'il s'agit du brin positif noté (+)ssDNA*attP* (avec permission de reproduction, Val *et al*, 2005).

A noter également que le phage monobrin peut être rendu bicaténaire par la machinerie cellulaire, et se conserver à chaque division cellulaire comme un plasmide.

Un modèle d'intégration du phage a été proposé par Val *et al*, et met en jeu le site *difI* du chromosome 1 de *Vibrio cholerae*, et la forme repliée de l'*attP*(+)ssDNA (figure 16).

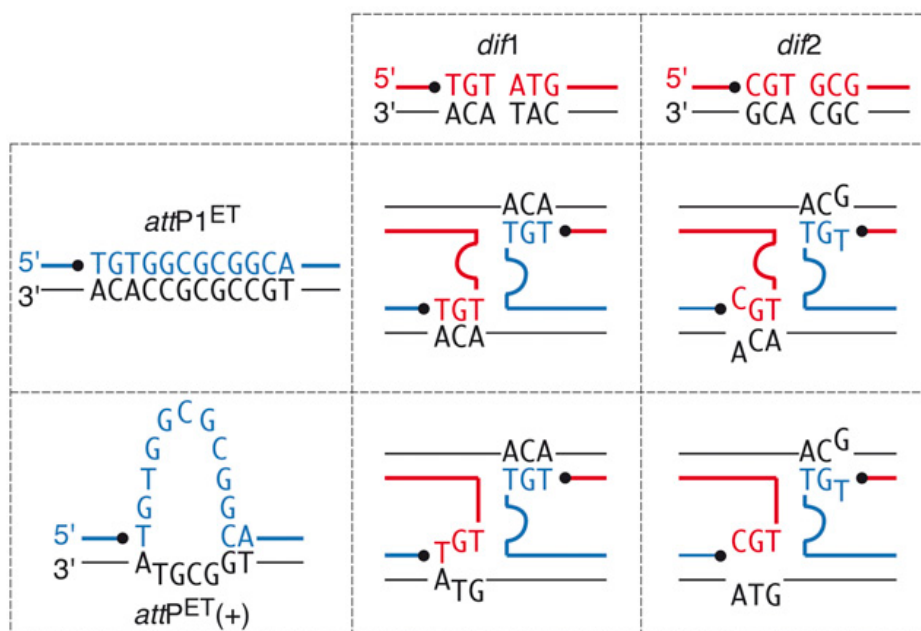


**Figure 16.** Modèle d'intégration du phage dans le chromosome 1. Dans ce modèle seul XerC effectue l'échange de brin même si XerD est capable de se fixer sur le site. Il a été montré qu'ensuite, c'est la réplication ou/et les systèmes de réparation qui terminent l'intégration du phage dans le site *dif1* (avec permission de reproduction, Val *et al*, 2005). On remarque la formation des séquences *attR* et *attL* qui sont apparues de chaque côté de la zone d'intégration du phage CTX $\phi$ .

Dans ce système XerC seul attaque le site *dif1*, la fin de l'intégration se faisant grâce à la réplication ou à des protéines de réparation de l'ADN. C'est ce phénomène qui entraîne la formation des deux séquences *attR* et *attL* propres à l'intégration irréversible du phage. C'est d'ailleurs le site *attR* qui permettra la résolution des dimères du chromosome 1 de par sa séquence permettant la fixation des protéines XerC et XerD de *Vibrio cholerae*. Donc *attR* remplace *dif1*, mais *attL* est incompatible de par sa séquence à la fixation des recombinaisons.

D'après le modèle présenté, il n'y a pas d'intégration dans le site *dif2* du deuxième chromosome de *Vibrio cholerae*. Et ce, par manque d'interaction ne permettant pas la stabilisation de l'échange de brins.

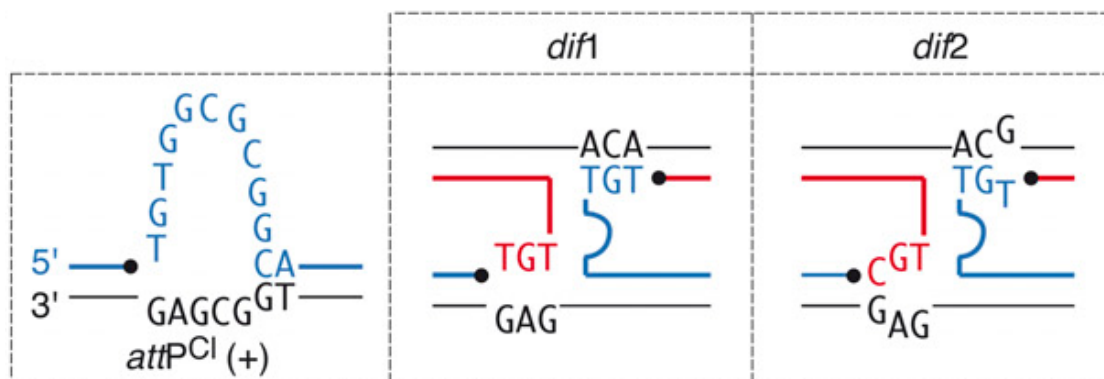
Néanmoins, on a découvert que la version classique de CTX $\phi$  pouvait s'intégrer au niveau de *dif2* dans des souches classiques *Vibrio cholerae* ainsi que dans l'isolat El Tor (souche N16961) qui a la particularité de produire beaucoup de toxine cholérique. Il semblerait que cette intégration dans le site *dif2* soit due à un changement de deux bases dans la région *attP2* (ou *siteII attP*) dans la souche 569B de *Vibrio cholerae*. Dans ce cas, c'est ssDNA $\text{attP}(+)$  qui se recombine, soutenant le modèle d'intégration d'un simple brin du phage (Das *et al*, 2010). On sait aussi que l'appariement permettant une intégration dans le site *dif2* n'est pas parfait, mais il y a eu apparition d'un doublet TG qui permet l'appariement de manière suffisante pour l'intégration du phage (figure 17).



**Figure 17.** Appariement le plus probable des brins du variant El Tor de CTX $\phi$ . Les brins clivés par XerC lors de la résolution des dimères et les sites d'attachement du phage sont montrés respectivement en rouge et en bleu. Les interactions d'appariement sont indiquées par la proximité des bases. Les sites *dif1* et *dif2* sur lesquels les deux sites putatifs d'attachement d'El Tor *attP1<sup>ET</sup>* et *attP<sup>ET</sup>(+)* se lient ont été indiqués (ce dernier provenant du (+)ssDNA de El Tor (Das *et al*, 2010).

Il y a peu d'intermédiaires covalents pour *dif2*/XerC, et il y a un contrôle plus stringant au niveau de *dif2* que de *dif1*.

On a donc mis en évidence l'existence d'un nouveau variant de CTX $\phi$  s'intégrant dans des souches bien particulières (figure 18).



**Figure 18.** Schéma des interactions de type Watson et Crick au niveau des paires de bases entre les sites *dif1* du chromosome 1 et *dif2* du chromosome 2. Ces interactions pourraient stabiliser l'échange de brins catalysé par la tyrosine recombinase XerC entre les régions se chevauchant lors de la résolution des dimères pour les sites de la souche N16961 et le site d'attachement trouvé dans le génome (+) ssDNA du variant classique de CTX $\phi$ . Comme pour la figure précédente, les brins clivés par XerC lors de la résolution des dimères et les sites d'attachement du phage sont montrés respectivement en rouge et en bleu, *attP<sup>Cl</sup>* correspondant à ce que l'on trouve dans le variant classique de CTX $\phi$  (Das *et al*, 2010).

C'est l'intégration dans les séquences *dif1* et *dif2* qui confirme le modèle mis-au-point par Val *et al* en 2005 pour l'intégration du phage CTX $\phi$  dans les diverses souches étudiées de *Vibrio cholerae*.

Une conversion lysogénique aurait donc lieu grâce à CTX $\phi$ , et c'est le mode premier d'acquisition des gènes toxiques chez les souches de choléra (Das *et al*, 2009).

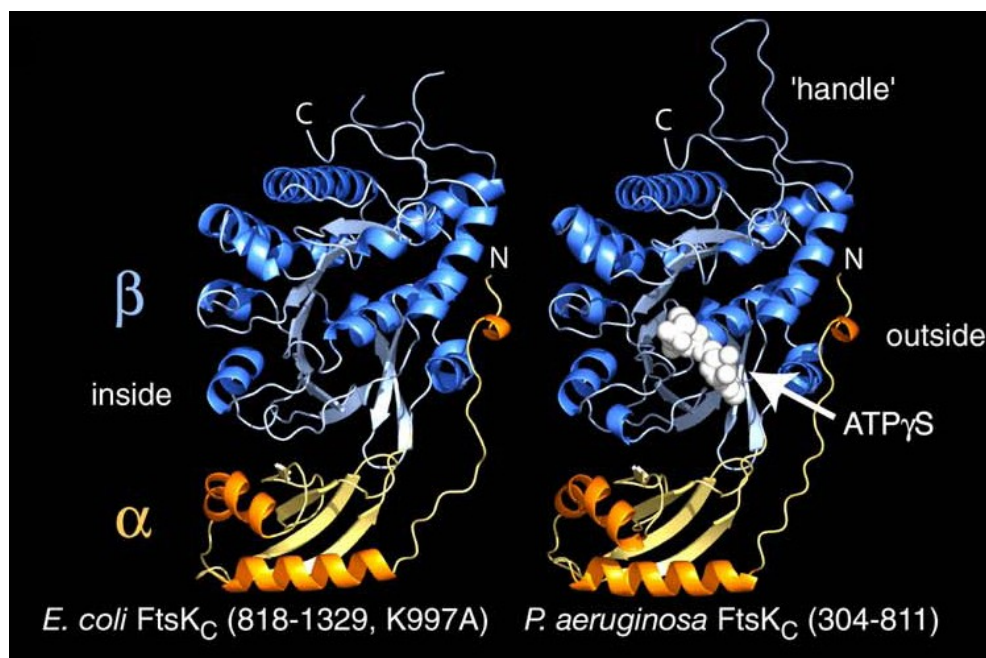
	XerC	espaceur	XerD
<i>dif E.coli</i>	GGTGCGCATAA	-----	TGTATA TTATGTTAAAT
<i>dif B.subtilis</i>	ACTTCCTAGAA	-----	TATATA TTATGTAAACT
<i>dif C.crescentus</i>	AAGATCGACTT	-----	-TGTAAT TTATGTAAAGT
<i>hif H.influenzae</i>	ATTTTCGCATAA	-----	-TATAAA TTATGTTAAAT
R1	GGTGCGCATAA	-----	TGTATA TTATGTTAAAT
<i>psi</i>	GGTGCGCGCAA	-----	GATCCA TTATGTTAAAC
<i>cer</i>	GGTGCGTACAA	-----	TTAAGGGA TTATGGTAAAT
<i>mwr</i>	GGTGACGCAA	-----	CAGATG TTATGGTAAAT
<i>nmr</i>	GGTGCGCGTAA	-----	-TGAGACG TTATGGTAAAT
<i>dif1 attB</i>	AGTGCGCATT	-----	TGTATG TTATGTTAAAT
<i>sitel attP</i>	AGTGCGTATTA	TGTGGCGCGGCA	TTATGTTGAGG
<i>sitelII attP</i>	AGTGCGTATTA	-----	-TACGCCA TTATGTTACGG
Consensus	GGTGCGTACAA		TTATGGTAAAT

**Figure 19.** Alignement des séquences de fixation de XerC (en indigo) et de XerD (en bleu-vert). La première séquence *dif* est celle d'*E.coli*, la seconde séquence du même nom est celle de *B.subtilis*. R1 est en fait la référence du plasmide R1, qui porte une séquence *dif*. La séquence *psi* est celle du plasmide pSC1001, la séquence *cer* est celle présente dans le plasmide ColE1. Les séquences homologues à *cer* nommées *mwr* et *nmr* sont respectivement présentes dans les plasmides pHS-2 et NTP16. La séquence *dif1 attB* est celle du chromosome 1 de *Vibrio cholerae*. Les séquences *sitel attP* et *sitelII attP* sont les séquences présentes dans le phage CTX $\phi$  au niveau d'*attP*, qui forment un repliement en épingle à cheveux lorsque le phage est monocaténaire, permettant la fixation de XerC et XerD. Enfin, la séquence consensus (inspirée de Hayse and Sherratt, 1997) qui montre les bases les plus communément rencontrées dans les différents sites *dif* venant des bactéries gram positives et gram négatives a été indiquée.

### 1.5. La protéine FtsK

La protéine FtsK est une protéine accessoire essentielle à la résolution des dimères chromosomiques. Comme on l'a dit plus haut avec la séquence *dif*, cette

résolution des dimères générés par la réplication a lieu avant la division cellulaire. Cette protéine est membranaire, essentielle et multifonctionnelle. Cette protéine de 1329 aa peut être divisée en trois domaines, une partie N-terminal membranaire de 200-279 aa (noté FtsK<sub>N</sub>) comprenant 4 domaines transmembranaires situés au niveau du septum bactérien au moment de la division cellulaire (Draper *et al*, 1998), une partie dite de liaison, FtsK<sub>L</sub>, de 650 aa riche en proline et glutamine, participant à la septation (Dubarry *et al*, 2010). Et enfin, une partie C-terminale de 500 aa avec une fonction AAA ATPase (FtsK<sub>C</sub>) (Yu *et al*, 1998 ; Barre *et al*, 2000 ; Aussel *et al*, 2002 ; Graham *et al*, 2005). Cette partie FtsK<sub>C</sub> contient trois domaines, un  $\alpha$  unique à la protéine, un domaine  $\beta$  faisant penser à un domaine de la protéine RecA (figure 20) (Massey *et al*, 2006), et un domaine  $\gamma$  qui se lit aux séquences KOPS (Löwe *et al*, 2008 ; Nolivos *et al*, 2010).

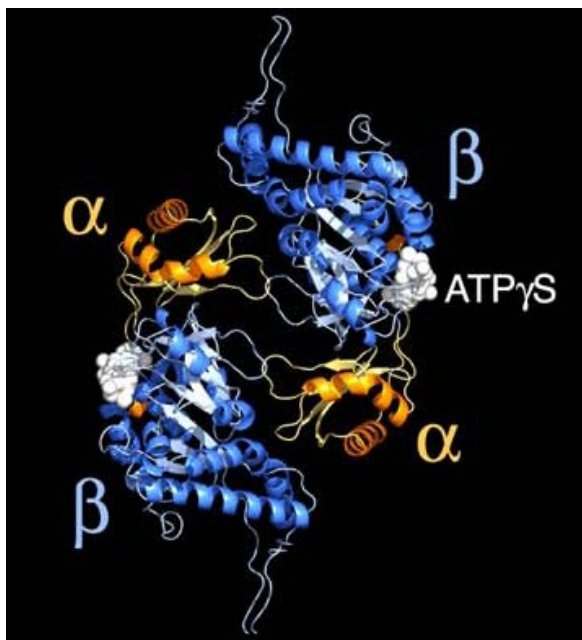


**Figure 20.** Structure 3D de la sous-unité FtsK chez *E.coli* et *P.aeruginosa*. Les domaines  $\alpha$  et  $\beta$  sont respectivement colorés en orange et en bleu. À titre de comparaison, la protéine FtsK de *P.aeruginosa* a été mise à côté de celle d'*E.coli*, en effet, il y a peu de différences structurales entre les deux. Pour la protéine de *P.aeruginosa*, le site de fixation de l'ATP a été indiqué (avec permission de reproduction, Massey *et al*, 2006).

Si l'on sait que cette protéine est essentielle pour le bon déroulement de la ségrégation du chromosome, étant donné que c'est la protéine accessoire de la



recombinaison médiée par les protéines XerC et D, c'est la partie Ct qui permettrait aux Xer de réaliser leur travail (Liu *et al*, 1998 ; Yu *et al*, 1998 ; Recchia *et al*, 1999 ; Steiner *et al*, 1999). De plus, cette protéine forme un hexamère pour être fonctionnelle. Cet hexamère forme un anneau à travers lequel l'ADN est censé passer lors de la ségrégation de l'ADN (figure 22).



**Figure 21.** Représentation d'un dimère de deux sous-unités de la protéine FtsK de *P.aeruginosa*. Cette structure est également valable pour la protéine venant d'*E.coli*. Cette structure montre les deux sous-unités tête-bêche. Mais on ne la retrouve pas dans l'hexamère fonctionnel de FtsK *in vivo*. Sur ce pur produit expérimental *in vitro*, on peut voir les sites de fixation de l'ATP (avec permission de reproduction, Massey *et al*, 2006).

On considère généralement que FtsK positionne les régions du terminus de réplication du chromosome au milieu des cellules, formant une synapse avec les sites *dif*, qui sont rapprochés après réplication et ce au niveau du septum, ce qui permettrait de générer une catalyse par XerC et XerD FtsK dépendante (Capiaux *et al*, 2002 ; Corre and Louarn, 2002 ; Bigot *et al*, 2004). Cette action favorisant l'activité de XerC et XerD serait ATP dépendante (Aussel *et al*, 2002).

Dans le système de recombinaison dont on parle, on a constaté que *dif* permettait de diriger la protéine FtsK (qui agit en tant que translocase) pour qu'elle agisse dans la bonne orientation (Pease *et al*, 2005). C'est ensuite la partie Ct de FtsK qui est impliquée dans la recombinaison, en permettant à XerD de catalyser la formation de la jonction d'Holliday, aussi bien *in vivo* qu'*in vitro* (Aussel *et al*, 2002).

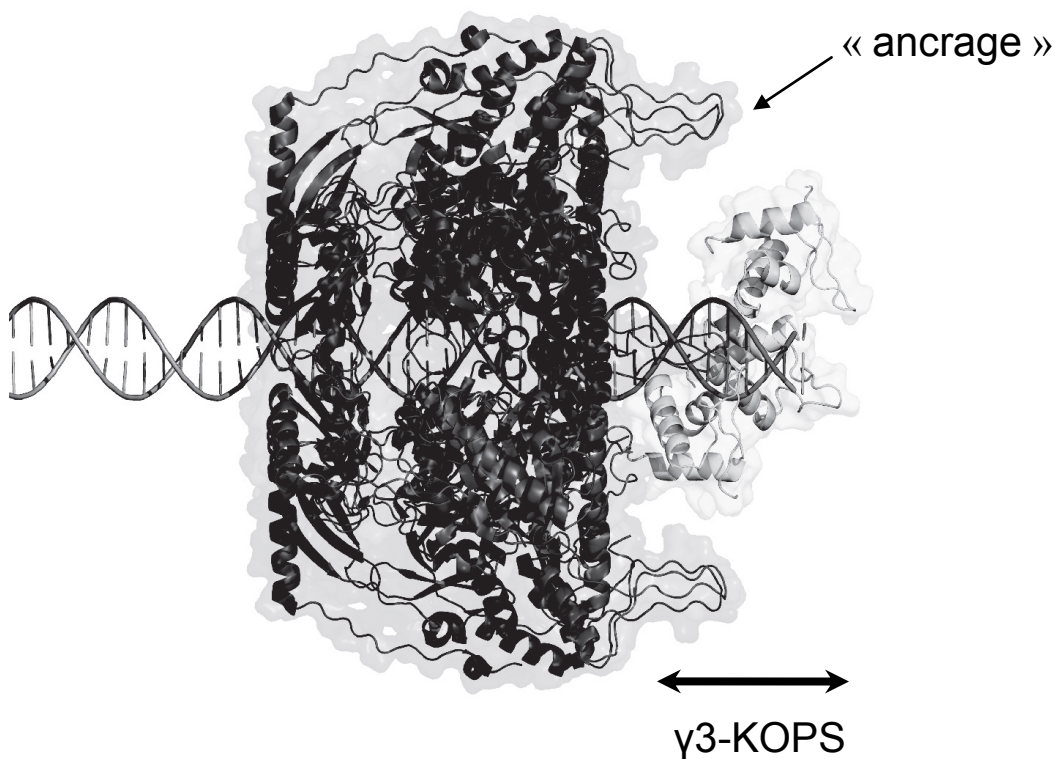
En 2002, deux voies alternatives, une voie FtsK dépendante et une voie FtsK indépendante, ont été proposées par Aussel *et al* (2002) pour la recombinaison au niveau du site *dif*. L'une est initiée par XerC et l'autre par XerD. Lors de la recombinaison, FtsK « *switch* » l'activité des deux recombinases dans le complexe synaptique formé par FtsK, XerC et XerD. Ce qui fera que la recombinaison va suivre un chemin sur lequel XerD va médier le premier échange de brins pour former l'intermédiaire typique de la jonction d'Holliday. Ensuite, XerC finalise la recombinaison, séparant le dimère en deux monomères.

Cette précédente description correspond à la voie FtsK dépendante. En l'absence de FtsK, les intermédiaires générés par les jonctions d'Holliday sont générés par des échanges de brins dus à XerC et la résolution des dimères est avortée. Les jonctions d'Holliday reviennent cinétiquement vers le substrat de départ, et XerD ne peut pas intervenir pour former des monomères.

La partie Ct de FtsK aurait donc un rôle de translocase favorisant la synapse au niveau du site *dif* et entraîne un échange de brins par la protéine XerD après formation de la dite synapse. Cette activation de XerD dans les échanges de brins de la jonction d'Holliday ne peut se faire qu'avec hydrolyse de l'ATP, phénomène permis par la partie ATPase de FtsK (Massey *et al*, 2004). De plus le domaine  $\gamma$  de FtsK active la recombinaison au niveau du site *dif* par les protéines XerC et XerD (Nolivos *et al*, 2010).

Mais on sait que c'est *in vivo* que FtsK (avec sa fonction translocase) semble devenir active au niveau du septum, et ce tardivement durant la division cellulaire (appelée aussi cytokinèse) en s'assemblant sur l'ADN chromosomique qui a échoué dans sa ségrégation, soit à cause de la formation des dimères citée plus haut, soit à cause de problèmes dans la décaténation du chromosome, soit à cause d'une réplication délayée. Dans tous les cas de figures, la translocation par FtsK est guidée par les séquences KOPS, et FtsK va interagir avec deux recombinases à tyrosine, XerC et XerD au niveau du site *dif* (Kennedy *et al*, 2008 ; Arciszemska *et al*, 2006).

En 2009, il a été démontré que lorsque l'on fabriquait des protéines FtsK tronquées ayant perdu leur domaine transmembranaire, elles permettaient toujours la résolution des dimères chromosomiques. Mais seulement si ces protéines sont connectées à un peptide du septum bactérien. Ce qui semble prouver que FtsK n'a pas besoin de transporter l'ADN à travers un pore formé par son domaine membranaire (Dubarry and Barre, 2010).



**Figure 22.** Représentation simplifiée de l'hexamère formé par FtsK au moment de la ségrégation. Ici sont surtout représentées de manière visible deux des sous-unités de cet hexamère. La séquence d'ADN représentée sur cette figure est une séquence KOPS qui représente le complexe d'initiation de la translocation vers le site *dif*. Si le complexe formé par FtsK est hexamérique, la partie  $\gamma$  qui se lie à la séquence KOPS est ici représentée de manière trimérique (pour plus de clarté) et notée  $\gamma$ 3-KOPS. Une zone d'ancrage venant de FtsK est également montrée (avec permission de reproduction, adapté de Graham *et al*, 2010).

## I.6. La protéine ArcA

La protéine ArcA (238 aa) est une protéine de type régulateur de réponse (*response regulator*, soit RR) faisant partie de la sous-famille des protéines OmpR/PhoB (Toro-Roman *et al*, 2005). Elle est rendue fonctionnelle par la protéine ArcB par phosphorylation (il est à noter que ArcB s'autophosphoryle) (Luchi and Lin, 1992 ; Luchi and Lin, 1993 ; Luchi and Lin, 1996).

ArcA se lie aux promoteurs de très nombreux gènes, et permet la régulation de plus de 30 opérons impliqués dans la régulation redox chez *E.coli* et plus récemment, on a vu qu'ArcA est directement impliquée dans la régulation d'au moins 55 autres opérons impliqués dans de nombreuses fonctions cellulaires.

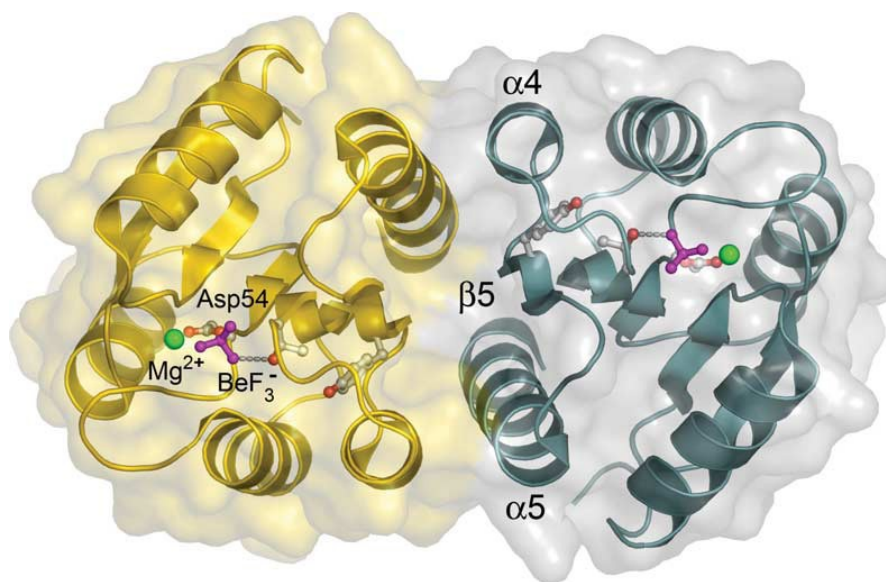
On s'est aperçu que plus de 9% des ORF (Open Reading Frames, soit phase de lecture ouverte) d'*E.coli* sont affectés directement ou indirectement par ArcA-P (Toro-Roman *et al*, 2005).

ArcA phosphorylée (ArcA-P) permet généralement de réprimer les gènes dont la transcription est inutile sous condition d'anaérobiose, et d'activer d'autres gènes requis lors de l'anaérobiose (Chao *et al*, 1997 ; Luchi and Lin, 1988 ; Tardat and Touati, 1993 ; Cotter and Gunsalus, 1992 ; Sawyer and Suppmann, 1992).

Il a été prouvé qu'ArcA-P permet la recombinaison au site *psi* où elle joue un rôle de protéine accessoire (Colloms *et al*, 1998). Elle y permet la formation d'un complexe synaptique, et se lie à 100 pb en amont des séquences échangées lors de la recombinaison. ArcA-P doit se multimériser pour permettre la recombinaison au niveau du site *psi*. En effet, elle forme des octamères pouvant se lier à l'ADN grâce à sa partie Ct. (Jeon *et al*, 2001 ; Toro-Roman *et al*, 2005). Cette multimérisation avait été notamment supposée par des expériences de protection à la DNaseI sur les sites de liaison à ArcA-P montrant que les zones protégées contre la digestion sont plus importantes que ce qu'une seule unité ne pourrait recouvrir (Bregu *et al*, 2002).

De plus, la partie Ct aurait un rôle dans la multimérisation et l'oligomérisation de la protéine entière.

En ce qui concerne la partie Nt de ArcA, elle a été récemment analysée, avec un analogue du groupement Phosphoryle (qui normalement est fixé sur ArcA par ArcB), le  $\text{BeF}_3^-$  qui a permis de mettre en évidence la dimérisation d'ArcA. Ce dimère est symétrique par un point et sa formation est médiée par les hélices et feuillettes  $\alpha 4$ - $\beta 5$ - $\alpha 5$ . Dans cette même région, on trouve une partie hydrophobe et tout un ensemble de ponts salins. Dans cette région, l'analogue au phosphoryle se lie à l'Asp54 de la protéine, mais permet aussi la liaison hydrogène et les ponts salins avec Thr81 et Lys103, et ce respectivement (figure 23).

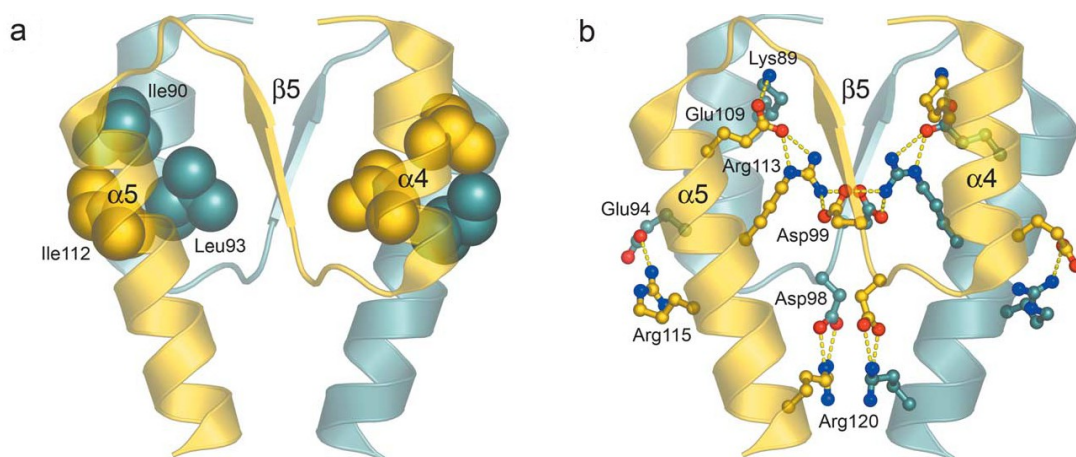


**Figure 23.** Schéma d'un dimère de protéine ArcA. La partie Nt est ici mise en évidence. L'analogue au phosphoryle, le  $\text{BeF}_3^-$  aussi, ainsi que l'ion catalytique  $\text{Mg}^{2+}$ . Chaque dimère est représenté de couleur différente et le point de symétrie est situé sur la face où se trouve l'axe formé par  $\alpha 4$ - $\beta 5$ - $\alpha 5$ . Ici, l'Asp54, sur lequel  $\text{BeF}_3^-$  se fixe également a été indiqué (avec permission de reproduction, Toro-Roman *et al*, 2005).

Au niveau des dimères, l'hélice  $\alpha 4$  d'un monomère se lie avec l'hélice  $\alpha 5$  de l'autre monomère, et ce grâce à un ensemble d'acides-aminés hydrophobes constitués

de l'Ile90 (dans l'hélice  $\alpha 4$ ), la Leu93 ( $\alpha 4$ ) et de l'Ile112 ( $\alpha 5$ ). Sont également à prendre en compte des portions des chaînes aliphatiques des acides-aminés Lys89 ( $\alpha 4$ ), Glu109 ( $\alpha 5$ ) et Arg113 ( $\alpha 5$ ).

Pour ce qui est des ponts salins qui existent dans cette région, ils se forment entre les acides-aminés Lys89 ( $\alpha 4$ )-Glu109 ( $\alpha 5$ ), Asp99 ( $\beta 5$ )-Arg113 ( $\alpha 5$ ), Glu94 ( $\alpha 4$ )-Arg115 ( $\alpha 5$ ) et Asp98 ( $\alpha 4$ -boucle  $\beta 5$ )-Arg120 ( $\alpha 5$ ) (Toro-Roman *et al*, 2005) (figure 24).



**Figure 24.** Représentation de la région de liaison entre les deux monomères d'ArcA.

- a)** Représentation des acides-aminés hydrophobes, leur couleur change selon le fait qu'ils viennent d'un monomère ou de l'autre. Ces acides-aminés permettent de rapprocher les hélices  $\alpha 4$  et  $\alpha 5$  et comprennent Ile90, la Leu93 et l'Ile112.
- b)** Représentation des ponts salins qui permettent la stabilisation de l'interface entre les deux monomères. Beaucoup d'acides-aminés provenant de chaque monomère sont impliqués. La liste est la suivante, Lys89 ( $\alpha 4$ )-Glu109 ( $\alpha 5$ ), Asp99 ( $\beta 5$ )-Arg113( $\alpha 5$ ), Glu94 ( $\alpha 4$ )-Arg115 ( $\alpha 5$ ) et Asp98 ( $\alpha 4$ -boucle  $\beta 5$ )-Arg120 ( $\alpha 5$ ) et met en œuvre les acides-aminés des deux hélices  $\alpha$  et du feuillet  $\beta$  (avec la permission de reproduction, Toro-Roman *et al*, 2005).

Il est remarquable que la structure  $\alpha 4$ - $\beta 5$ - $\alpha 5$  soit très conservée dans la sous-famille des protéines OmpR/PhoB.

## **I.7. La protéine PepA**

PepA a été caractérisée pour la première fois en 1970 par l'équipe de Vogt, chez *Escherichia coli*.

### **I.7.1. Découverte**

Elle a été découverte dans les années 1970. Cette protéine cytoplasmique, codée par le gène *pepA*, fait partie de la famille des leuaminopeptidases, large famille présente chez les bactéries, mais aussi chez les plantes et les mammifères. Elle permet le clivage des peptides et polypeptides en position N-terminale (Nt), tant chez les eucaryotes que chez les procaryotes, surtout au niveau des acides-aminés leucine et méthionine (Cuypers *et al*, 1982), mais elle permet également la recombinaison au niveau du site *cer*. PepA avait été précédemment appelée XerB et CarP (Bartling and Weiler, 1992 ; Burley *et al*, 1992).

La découverte de PepA vient du fait que l'équipe de Vogt voulait connaître quelle enzyme était responsable de la lyse de polypeptides au niveau du résidu méthionine, car seulement 42% des protéines de la cellule conservaient cette méthionine (2 % de protéines solubles, 40 % de protéines insolubles). Or toutes les protéines sont générées avec un résidu méthionine au début de leur séquence lors de la traduction des ARNm, et 50 % des protéines extraites des cellules débutaient leurs séquences avec un résidu sérine ou un résidu alanine. Il est important de remarquer que les eucaryotes ont également leurs protéines commençant par un résidu méthionine, d'où l'importance de la protéine dans les reines procaryotes et eucaryotes.

Partant de ce constat, l'équipe de Vogt rechercha une enzyme susceptible de cliver spécifiquement une protéine dans sa partie Nt si cette dernière contenait une méthionine. Lorsqu'une telle protéine fut découverte, on constata que non seulement elle pouvait cliver le dérivé méthionine, mais également le dérivé leucine (Vogt, 1970). Par contre, ce n'est qu'en 1995 que l'on détermina que les gènes *pepA* et *carP*

formaient un seul et même gène, désormais identifié comme *pepA* (Charlier, D. *et al*, 1995b).

### I.7.2. Caractéristiques

Pour PepA, lorsque la fonction aminopeptidase est utilisée, les peptides générés le sont souvent à cause de la dégradation de protéines cellulaires ou bien ils sont générés à partir d'autres peptides qui sont d'origine exogène, et qui vont servir de nutriment à la cellule.

PepA est une protéine de 55,3 KiloDaltons (kDa) d'une taille de 503 acides-aminés (aa). C'est une protéine dépendante aux ions  $Mn^{2+}$  qui se fixent sur la partie C-terminale (Ct) de la protéine PepA (ils servent de cofacteurs). Il n'y a qu'un seul ion  $Mn^{2+}$  qui se fixe sur cette partie, rendant la protéine active. On a constaté que l'ion  $Zn^{2+}$  a la capacité inverse, c'est-à-dire qu'il inhibe l'activité peptidase de la protéine. La protéine PepA est constituée de trois parties principales.

Un domaine N-terminal allant des acides-aminés 1 à 166 qui permet la liaison à l'ADN. Une région centrale de type hélice  $\alpha$  des acides-aminée 167 à 192. Une région C-terminale allant des acides-aminés 193 à 503 qui génère la fonction aminopeptidase.

Pour être fonctionnelle, la protéine PepA doit s'assembler en un complexe hexamérique de 330 kDa au total (figure 25), sous cette forme, le complexe fait 330 kDa et peut également se lier à l'ADN, notamment lors de rétrocontrôles négatifs de son propre gène. Si sa fonction aminopeptidase était bien connue, son rôle dans d'autres fonctions biologiques était inconnu jusque dans les années 1980. En effet, c'est grâce à l'étude du système de recombinaison du plasmide ColE1 au niveau de la séquence *cer*, que la nature du gène *xerB* fut décelée.



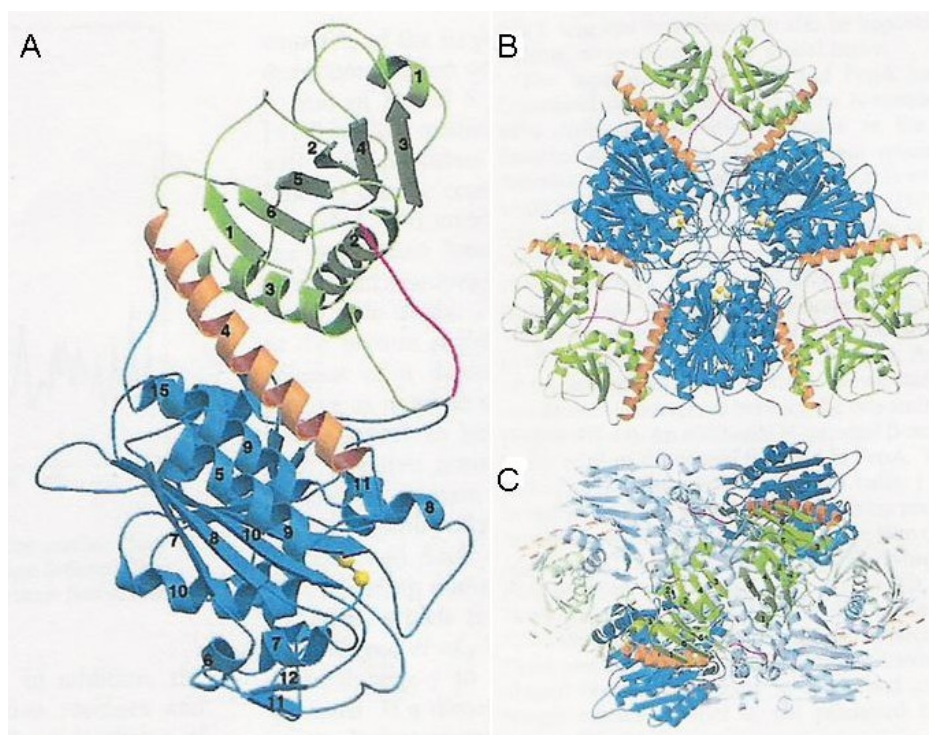
Dans ces études, ce gène chromosomique d'*E.coli*, permettant de résoudre les dimères plasmidiques, a été inséré dans des versions mutantes et non fonctionnelles du gène *pepA* des cellules de *S.typhimirium*.

Les résultats furent les suivants, *xerB* complémentait les bactéries *S.typhimirium*, et après séquençage de *xerB*, on découvrit qu'il y avait de fortes homologues avec *pepA* de *S.typhimirium* (Stirling *et al*, 1989). D'autres homologues, sur la séquence protéique cette fois, ont été trouvées entre la protéine codée par *xerB* et l'aminopeptidase bovine (appelée BLLAP pour Bovine Lens Leucine Aminopeptidase, mais qui ne semble pas impliquée dans des liaisons à l'ADN), soit 22% en partie Nt et 52% en partie Ct, où 7 résidus du site actif sont conservés. Tous ces événements ont conduit à renommer *xerB* d'*E.coli pepA*.

De plus il a été démontré que cette fonction aminopeptidase n'est pas utilisée lors de la recombinaison. En effet, si on mute un des 7 résidus conservés cités plus haut, il y a perte de l'activité de clivage. Plusieurs mutants ayant perdu cette activité ont été analysés pour voir leur activité de résolution de multimères. Tous ont été capables de générer des monomères. Donc, PepA agit au sein d'un complexe protéique permettant aux recombinases de type Xer de résoudre les problèmes de dimérisation. Mais PepA est aussi connue pour d'autres répressions comme celle de l'opéron *carAB* par liaison spécifique de sa partie Nt au niveau du promoteur P1 (appelé carP1) de ce dernier. Dans ce système, c'est le gène *carP* (régulateur de l'opéron *carAB*), qui après séquençage, s'est révélé être le gène *pepA* (Charlier *et al*, 1995b ; Roovers *et al*, 1988).

Il a été dit dans Reijns *et al* (2005) que la région Nt de PepA permet la recombinaison en *cer* et *psi* grâce à des résidus d'acide-aminés basiques (souvent et surtout lysine et arginine) situés à la surface de PepA, formant une sorte de « chemin » sur la dite surface. De plus, ces mutations empêchent la répression de l'opéron *carAB*, montrant que ce sont les mêmes résidus aux mêmes emplacements qui sont responsables des deux phénomènes. Ce sont des substitutions qui ont permis de connaître l'emplacement de près de 16 résidus permettant la recombinaison, toutes ont

été localisées dans ce « chemin » cité plus haut. Aucune de ces mutations n'a gêné la fonction peptidase de la protéine PepA hexamérisée. D'autre part, toujours dans le même article, de nombreuses mutations ont été générées dans la partie Ct de PepA entre les résidus 190 et 503. Certaines ont montré une forte baisse de l'activité peptidase en plus d'une incapacité à permettre la recombinaison.



**Figure 25.** Représentation de la sous-unité de PepA et de son hexamère.

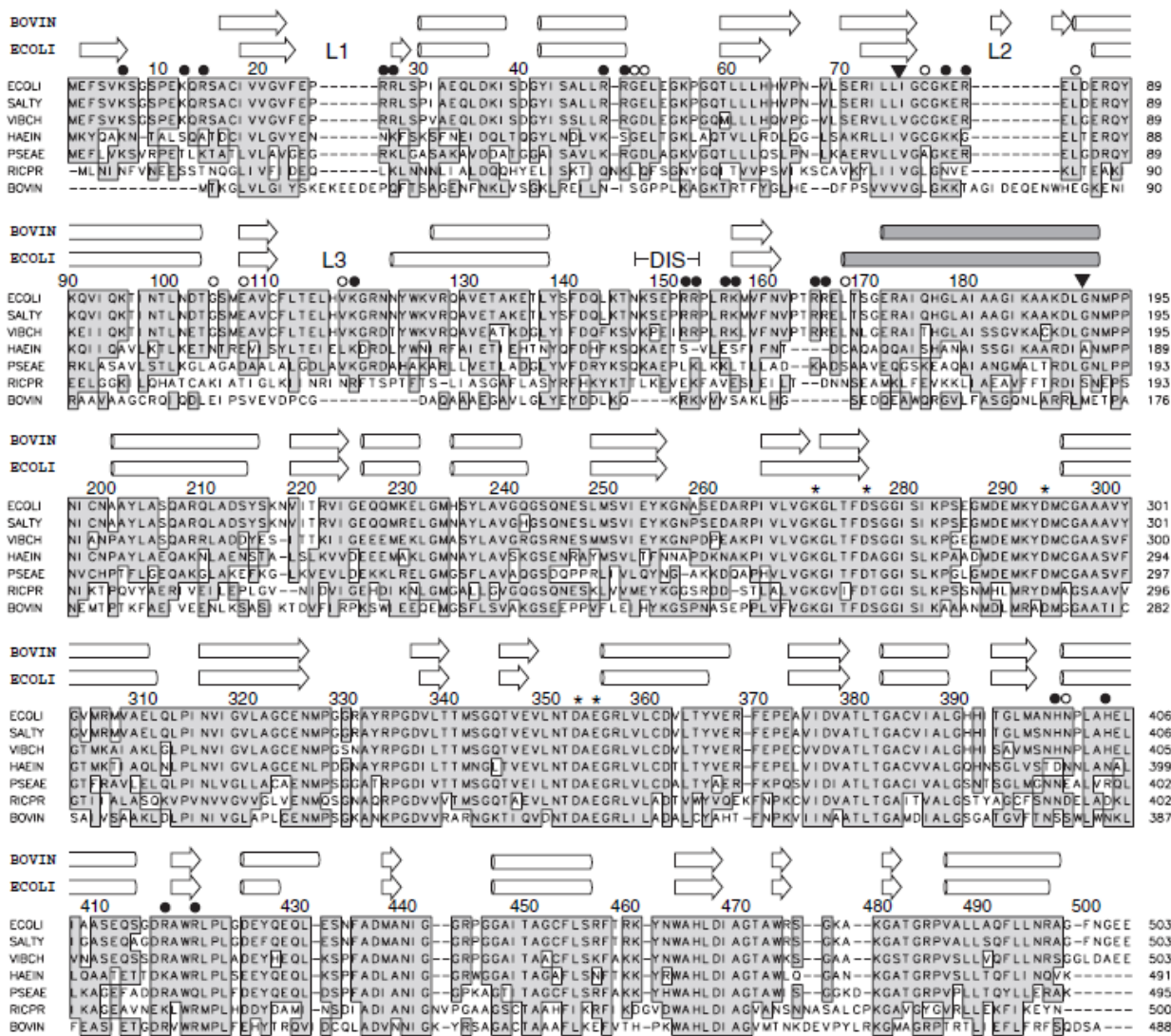
**A)** Protéine PepA avec son site pour  $Mn^{2+}$ , une région centrale formée d'une hélice  $\alpha$ , représentée en orange, lie les deux régions Nt (en vert) et Ct (en bleu). La région représentée en Violet est appelée région désordonnée car on ne connaît pas exactement sa structure spatiale.

**B)** Hexamère de PepA, les parties Nt sont en périphérie du complexe avec une symétrie de type 32.

**C)** Autre représentation du complexe hexamérique PepA, les parties Nt en périphérie sont en contact deux par deux (avec permission de reproduction, Sträter *et al*, 1999).

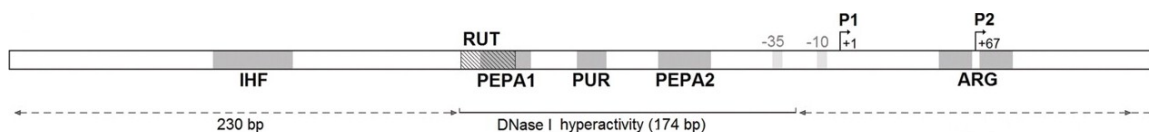
Outre la stimulation de la recombinaison au niveau des sites *cer*, les hexamères PepA permettent la recombinaison au niveau des sites *psi*. Lors de la recombinaison, il y aurait liaison sur des séquences accessoires de 180 pb, adjacentes à celles sur lesquelles les recombinases XerC et XerD agissent (il s'agit de sites de 20 à 30 pb). Au

niveau de la séquence *cer*, les séquences de liaisons décrites pour la protéine PepA sont présentes de part et d'autres de la boîte arginine de 18 pb.



Il existe de nombreux homologues pour la protéine PepA, eux aussi sont impliqués dans la régulation de la transcription des gènes en plus de leur fonction aminopeptidase. De plus, les protéines PepA issues de *V.cholerae* et de *S.typhimurium* sont capables de compléter des *E.coli* déficientes en PepA (figure 26).

### 1.7.3. *CarAB*



**Figure 27.** Schéma de la région de contrôle de l'opéron *carAB* comportant les promoteurs *carP1* et *carP2*, sur lesquels se lie la protéine PepA au niveau des séquences PEPA1 et PEPA2. La zone sur laquelle la DNase I possède une hyperactivité contient des sites de liaison pour les protéines PurR et PepA. D'autres sites de liaisons pour des protéines comme IHF, PUR et ARG ont également été indiqués (avec permission de reproduction, Minh *et al*, 2009).

L'opéron *carAB* code les gènes de la carbamoylphosphate synthétase, qui catalysent une étape commune à la biosynthèse de l'arginine et des pyrimidines (Charlier *et al*, 1995).

Chez *E.coli* et *S.typhimurium* la liaison de PepA se fait en amont du promoteur *carP1* au niveau des sites PEPA1 et PEPA2. On a démontré grâce à des expériences employant la DNaseI, que la protéine PepA se lie à des régions riches en résidus A et T. Au niveau de cet opéron, PepA exerce une régulation négative sur ses deux sites PEPA1 et PEPA2.

Il y a deux promoteurs sur cette région de l'opéron, les promoteurs P1 et P2. Le promoteur P1 fixant la protéine PepA et le promoteur P2 fixant la protéine ArgR (figure 27).

On a connaissance que d'autres facteurs comme l'intégration host factor (IHF), l'UMP-kinase (PyrH), la purine repressor (PurR) et le régulateur de l'utilisation de la pyrimidine (RutR) sont également requis pour une modulation totale de l'activité du promoteur (Minh *et al*, 2009).

Il faut signaler que la régulation du promoteur P1 est complexe et fait intervenir deux mécanismes différents et apparemment indépendants. Le premier serait un mécanisme sensible aux nucléotides UTP de manière réitérative et influencerait sur la transcription. Le second mécanisme serait un mécanisme protéine dépendant, contrôlant l'initiation de la transcription. Il requiert la participation d'au moins trois protéines multifonctionnelles, ce sont les protéines PepA, IHF (qui est une protéine proche des histones) et PyrH (qui se présente aussi sous la forme d'hexamères).

Toutes ces protéines voient leurs sites de liaison en amont des promoteurs P1 et P2. De plus PepA et IHF modulent le promoteur P1, tandis que la protéine PyrH servirait de senseur détectant des changements dans la quantité de pyrimidine intracellulaire (Charlier *et al*, 2000). C'est au niveau de P1 que l'on a une séquence GATC 106 paires de bases en amont de l'origine de transcription P1 qui se trouve protégée de la Dam méthylase lorsqu'il y a présence de résidus pyrimidines

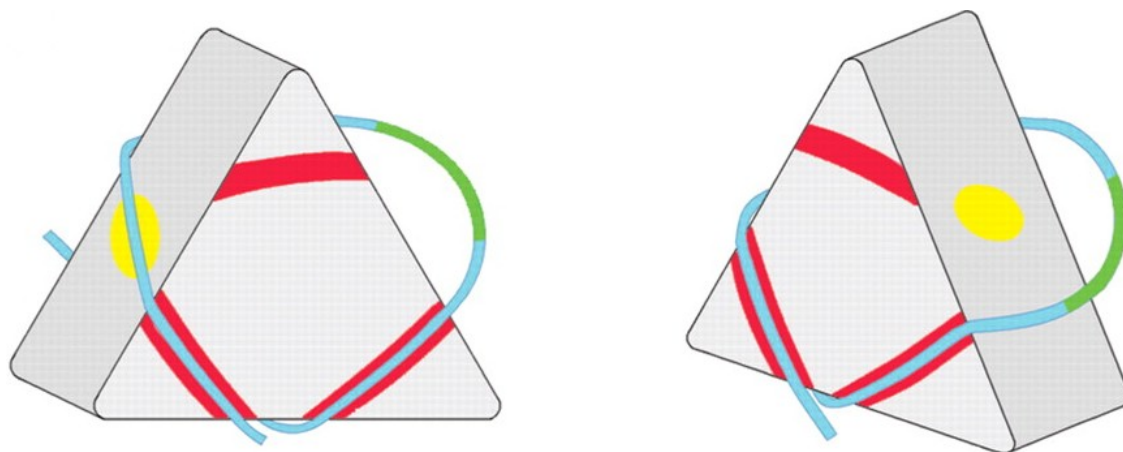
Il a été mis en évidence dans l'article de Charlier *et al* en 2000, que la plupart des mutations empêchant la liaison de la protéine PepA à l'ADN de la région promotrice de l'opéron *carAB*, se trouvent surtout dans la partie Nt et parfois dans la fin de la partie Ct. Parmi les mutations générées, la plupart ne permettait pas l'autorégulation du gène de la protéine PepA, malgré sa capacité à toujours se lier au promoteur de *carAB*. Parmi ces mutations conservant la liaison à l'ADN, on avait une perte de la répression de *carAB*.

Citons également parmi ces mutations le double mutant LMP15 (c'est le numéro pour les mutants de la série LMP) qui se caractérise par deux substitutions, A211T et D212 dans l'hélice  $\alpha 5$ . On citera aussi le mutant LMP25 et la mutation

A450V dans l'hélice  $\alpha$ 14, ainsi que LMP28 et la mutation E53K, juste après la première hélice  $\alpha$ .

Selon un récent article de l'équipe de Mr. Charlier en 2009 où l'on utilise la technique de microscopie par force atomique (Minh *et al*, 2009), lors de la liaison de PepA à l'ADN, il y a enroulement de l'ADN autour d'un hexamère de PepA.

Sur le promoteur de *carAB*, la liaison d'un hexamère de PepA entraîne un « emballage » de 235 pb avec condensation de l'ADN. Et ce, dans des proportions plus importantes en théorie que sur un seul hexamère, ce qui suggère qu'il y a des zones d'interaction dites serrées (partie Nt et sillon Ct de PepA) connectées par des boucles à faibles contacts. Ces boucles pourraient accueillir d'autres facteurs de régulations, dont la protéine PurR, car la liaison de PepA seule sur P1 n'est pas suffisante pour assurer le contrôle de la région, ce qui suggère qu'il y a contact protéine à protéine (figure 28) (Charlier *et al*, 2000).



**Figure 28.** Schéma interaction de PepA avec *carAB* (avec permission de reproduction, Minh *et al*, 2009). Ici l'ADN fait contact avec deux sillons Ct (en jaune, l'un n'est pas visible) et le passage de liaison pour l'ADN adjacent (en rouge, partie Nt de la protéine). Le troisième sillon de l'ADN (en vert) ne fait pas contact, et laisse le site de liaison de la protéine PurR libre.

Il est également important de remarquer que la protéine ArgR est capable de se lier sur la séquence P2 des promoteurs cités plus haut, grâce à une boîte arginine au sein de cette séquence. Elle permet aussi une régulation négative de *carAB*.

## **I.8. La protéine ArgR**

C'est une protéine qui a été découverte dans les années 1950 chez *E.coli* (Scher *et al*, 1957).

### **I.8.1. Découverte**

A la suite d'expériences sur la formation de l'acétyl-ornithinase, enzyme permettant la synthèse de l'ornithine par conversion de l'acéthylornithine, on a observé chez des mutants, dont la formation de l'acéthylornithine était bloquée, un temps d'attente avant la reprise de leur croissance lorsqu'on les plaçait d'un milieu complémenté en arginine à un milieu complémenté en acéthylornithine. Ce temps de latence a tout d'abord suggéré que l'acétyl-ornithinase était fabriquée en réponse à ce changement de milieu. Mais des investigations plus poussées ont montré que lors de la culture sur milieu supplémenté en arginine, l'enzyme l'acétyl-ornithinase était inhibée. Le temps d'attente reflétait en fait la remise en route de la machinerie enzymatique de la voie de synthèse de l'arginine.

D'autres études menées sur l'ornithine transcarbamylase (OTCase) ont montré qu'en présence d'arginine, l'enzyme n'était plus produite. Une fois l'arginine retirée du milieu, la synthèse de l'OTCase ne durait qu'au maximum une génération, ce qui laissait supposer que le rétablissement de la formation d'OTCase entraînait à nouveau la formation d'arginine, donc l'existence d'un feedback réprimant la formation d'OTCase. C'est avec l'utilisation de mutants déficients pour la formation d'ornithine

chez *Escherichia coli*, que l'on a montré que l'arginine contrôlait bien la synthèse d'OTCase (Maas *et al*, 1994), et non l'ornithine.

D'autres études par Vogel ont montré que sur les mutants bloqués avant la formation d'acétylornithine, un milieu supplémenté en acétylornithine limitait la synthèse en arginine, et limitait également la répression de la voie de synthèse. Mais le répresseur de cette voie utilisant l'arginine comme corépresseur restait encore à trouver.

Pour ce faire, on a utilisé de la canavanine, un analogue de l'arginine, qui entre en compétition avec elle et qui de ce fait, inhibe la croissance cellulaire. On a donc recherché des mutants qui étaient résistants à la canavanine, on en a trouvé dans les souches K12 d'*E.coli*. Ces mutants présentaient de hauts niveaux d'OTCase et n'étaient pas réprimés par l'arginine. D'autres enzymes de la voie de synthèse de l'arginine n'étaient plus réprimés. Les mutations furent cartographiées et toutes localisées sur un site du chromosome bactérien qui fut nommé *argR*. D'autres souches présentant la même altération de la répression par l'arginine furent trouvées et étudiées. C'est alors que le modèle de l'opéron présenté par Jacques Monod en 1961 éclaira les résultats et que le terme de régulon fut inventé, caractérisant un groupe des gènes apparentés au niveau métabolique, mais répartis sur tout le chromosome, et sous le contrôle d'un seul gène. Ce fut le cas pour *argR*.

La nature d'ArgR fut précisée grâce à des études sur des mutants sensibles à la température, mais aussi par double marquage, où une fraction protéique était présente chez les souches sauvages, mais pas les souches mutantes. On détermina ensuite que l'arginine était corépresseur de l'ArgR, qui diminuait la transcription des gènes. Le gène de l'ArgR ne fut cloné qu'en 1980 par Eckhardt (Maas *et al*, 1994).



## I.8.2. Description de la protéine

L'ArgR dont le nom signifie *Arginine repressor*, est une protéine de 17 kDa faisant 156 acides-aminés, codée par le gène chromosomique *argR*. Cette protéine se lie à l'ADN par sa partie Nt basique (acides-aminés 1 à 71), et interagirait avec les protéines et ses propres sous-unités grâce à sa partie Ct acide (acides-aminés 82 à 156), de plus, cette protéine possède théoriquement six hélices  $\alpha$  et six feuillets  $\beta$  (figure 29) (Burke *et al*, 1994 ; Tian and Maas, 1994 ; Van Duyne *et al*, 1996, Sénéchal *et al*, 2010).

On remarquera que la région centrale d'ArgR comprenant les résidus d'acides-aminés 72 à 81 est sensible aux protéases, mais ne semble pas intervenir dans la répression des gènes du régulon, ni dans d'autres phénomènes attribuables à ArgR.

```
MRSSAKQEELVKAFKALLKEEFSSQGEIVAALQEQGFDNINQSKVSRMLTKF
GAVRTRNAKMEMVYCLPAELGVPTTSSPLKNLVLDIDYNDVAVVVIHTSPGAA
QLIARLLDSLGAEGILGTIAGDDTIFTTPANGFTVKDLYEAILELFDQEL
```

**Figure 29.** Séquence de la protéine ArgR d'*E.coli*. Elle contient 156 acides-aminés

### I.8.2.1. Partie Nt

Comme indiqué plus haut, la partie Nt est responsable des interactions ADN/protéine, mais cette fonction fut d'abord décelée grâce à des expériences de mutagenèse dirigées sur les résidus 44 et 47 de la protéine au niveau de la troisième hélice  $\alpha$  (Tian and Maas, 1994 ; Sunnerhagen *et al*, 1997). En effet, des mutations de ces résidus empêchèrent la liaison de la protéine ArgR à l'ADN. Des études ont ensuite confirmé que c'est la partie Nt seule sans le reste de la protéine, qui aurait un effet de répression sur les gènes, bien que dans ce cas précis, la répression soit bien moindre (Grandori *et al*, 1995). Au niveau de la structure de cette partie de la protéine, on aurait

les trois premières hélices  $\alpha$  et les deux premiers feuillets  $\beta$  (Sunnerhagen *et al*, 1997; Sénéchal *et al*, 2010).

### **I.8.2.2.Partie Ct**

Cette partie Ct de la protéine a été décrite comme la partie responsable de la multimérisation d'ArgR, car elle suffit à former un hexamère (Van Duyne *et al*, 1996). Puis des expériences de diffraction aux rayons X démontrèrent que Ct se liait à la L-arginine. De plus, la structure tridimensionnelle de cette partie, montre que l'on retrouve les trois dernières hélices  $\alpha$  et les feuillets  $\beta$  numérotés de 3 à 6 (Sénéchal *et al*, 2010). Ce serait d'ailleurs les résidus 128 et 129 de la protéine ArgR, ainsi que les résidus 105 et 123, qui permettraient la liaison de la L-arginine à la protéine (Burke *et al*, 1994 ; Tian and Maas, 1994 ; Van Duyne *et al*, 1996). Les résidus 105 et 123 seraient également impliqués dans l'oligomérisation de la protéine. Cette partie serait aussi impliquée dans la formation de complexes avec d'autres protéines, en particulier PepA.

### **I.8.3. Fonction**

Comme son nom l'indique, ArgR permet la régulation de la voie de synthèse de l'arginine à partir du glutamate, par répression des gènes de cette voie après fixation d'ArgR sur le promoteur de l'opéron. Par ce phénomène, ArgR régule l'expression de plus d'une douzaine de gènes (Glansdorff, 1996 ; Caldara *et al*, 2006). Ces gènes sont organisés en neuf unités transcriptionnelles qui sont toutes réprimées par ArgR.

En présence de L-arginine, la protéine, alors sous forme hexamérique (on parle d'homo-hexamère) comme la protéine PepA, devient active et peut se lier à l'ADN. Il faut savoir que la formation de cette hexamère passe d'abord par la formation de trimères d'ArgR qui vont ensuite se dimériser (il s'agit de liaisons faibles, non-

covalentes). Il a récemment été démontré que c'était l'arginine en position 110 qui permettait cette trimérisation d'ArgR en interagissant avec l'Asp128 (Strawn *et al*, 2010).

Lorsqu'il y a liaison de la L-arginine à la protéine, la formation des trimères, puis des hexamères, pourra se faire, ainsi que la liaison des multimères d'ArgR sur l'ADN (en fait la liaison de la L-arginine rend la protéine ArgR active). À l'heure actuelle, on a décrit une vingtaine de gènes, regroupés en opérons ou non, qui sont régulés par la protéine ArgR. Ces gènes sont pour la plupart impliqués dans la voie de synthèse et le transport de l'arginine dans les cellules d'*Escherichia coli*.

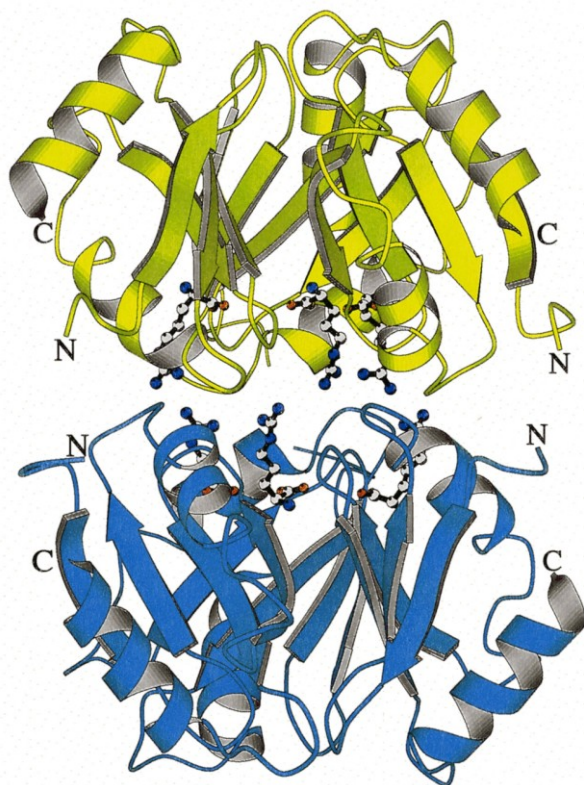
Tous ces gènes font partie du système de régulon décrit dans les années 1960 et qui a été mis-à-jour en 2006 (tableau 4) (Caldara *et al*, 2006).

On a en effet constaté que sans L-arginine, les trimères s'alignent les uns derrière les autres pour protéger la région hydrophobe d'ArgR. Ce sont d'ailleurs des interactions hydrophobes qui permettent la formation des trimères (Van Duyne *et al*, 1996).

Gènes	Enzymes codées
<i>argA</i>	NAGase
<i>argB</i>	N-Acétyleglutamase 5-phosphotransférase
<i>argC</i>	N-Acétyleglutamase 5-phosphate reductase
<i>argD</i>	N-Acétyleglutamase 5-aminotransférase
<i>argE</i>	N-Acétyleornithinase
<i>argF</i>	OTCase
<i>argI</i>	OTCase
<i>argG</i>	Argininosuccinate synthétase
<i>argH</i>	Argininosuccinate lyase
<i>carA</i>	CPSase, sous-unité A
<i>carB</i>	CPSase, sous-unité B
<i>artP</i>	Sous-unité du transporteur ABC de l'arginine
<i>artI</i>	Sous-unité du transporteur ABC de l'arginine
<i>artQ</i>	Sous-unité du transporteur ABC de l'arginine
<i>artM</i>	Sous-unité du transporteur ABC de l'arginine
<i>artJ</i>	Sous-unité du transporteur ABC de l'arginine
<i>hisJ</i>	Sous-unité du transporteur ABC de l'arginine
<i>hisQ</i>	Sous-unité du transporteur ABC de l'arginine
<i>hisM</i>	Sous-unité du transporteur ABC de l'histidine et du transporteur ABC lysine/arginine/ornithine
<i>hisP</i>	Sous-unité du transporteur ABC de l'histidine et du transporteur ABC lysine/arginine/ornithine

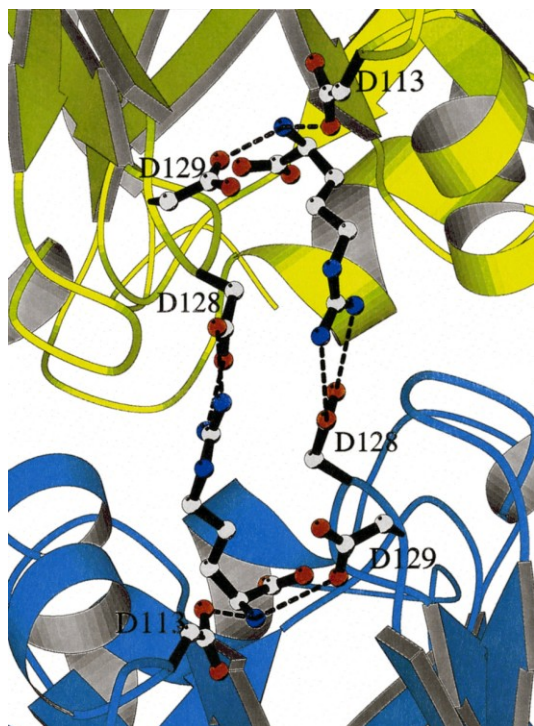
**Tableau 4.** Liste de gènes faisant partie du régulon ArgR. Tous sont réprimés par ArgR et sont en relation avec la voie de synthèse de l'arginine, les gènes *argB*, *argC*, *argE*, *argF* et *argH* font partie du même opéron. Il en va de même avec les gènes *carA* et *carB* qui font partie de l'opéron *carAB*. Tous ces gènes sont présents chez *E.coli* de la souche K-12 (avec permission de reproduction, inspiré de Caldara *et al*, 2006).

Si la L-arginine est présente (au moins six molécules), cette dernière va se lier entre deux trimères justement au niveau de la région hydrophobe, générant 10 liaisons hydrogènes et stabilisant la forme hexamérique d'ArgR (Van Duyne *et al*, 1996). De plus, il a été avancé que la molécule de L-arginine va se lier à un trimère d'ArgR en s'intercalant entre l'Arg110 et l'Asp128, remplaçant l'interaction d'Asp128 avec l'Arg110. Puis il y aura interaction de cette L-arginine avec Gln106, Asp113, Thr124 et Ala126, entraînant un changement de conformation de la sous-unité puis du trimère, permettant l'hexamérisation de la protéine et son activité (Strawn *et al*, 2010). De plus, les molécules de L-arginine d'un trimère vont interagir avec celles de l'autre trimère, renforçant la stabilité de l'hexamère. L'interaction entre la molécule de L-arginine et une sous-unité d'ArgR se fait également au niveau des acides-aminés Asp113, Asp128 et Asp129, contribuant à stabiliser le tout (Figure 30 et 31) (Van Duyne *et al*, 1996).



**Figure 30.** Parties Ct de ArgR (avec permission de reproduction, Van Duyne *et al*, 1996). Ici, on peut voir deux trimères de Ct tête-bêche liés par des interactions faibles.

C'est la partie Ct de la protéine ArgR qui est responsable de l'hexamérisation de la protéine selon une symétrie de type 32. On peut également voir dans chaque face intérieure, les molécules de L-arginine liées.



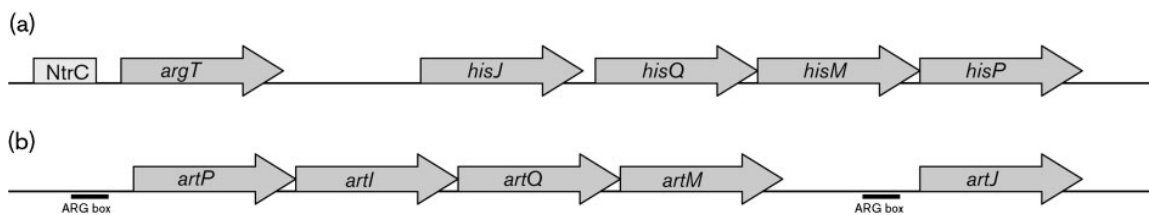
**Figure 31.** Schéma agrandi de la zone d'interaction entre les molécules de L-arginine et d'ArgR.

On peut voir les molécules de L-arginine interagissant avec les chaînes latérales des acides-aminés Asp113, Asp128, Asp129.

Les deux sous-unités se faisant face sont de deux couleurs différentes (Van Duyne *et al*, 1996).

Cette liaison à l'ADN de l'hexamère d'ArgR a lieu sur un site spécifique, la boîte arginine, qui est composée d'une répétition palindromique de 18 pb séparée de 3 pb, boîte pour laquelle ArgR possède une forte affinité. Cette séquence est riche en adénines et thymine, et se retrouve fortement dégénérée, la séquence consensus étant TNTGAATWWWATTTCANW. Elle a été décrite grâce à des analyses sur les sites régulant la biosynthèse de l'arginine (Maas, 1994). Au cours des dernières décennies, de nombreux gènes réprimés par ArgR ont été décrits. Tous sont impliqués dans la voie de synthèse de l'arginine ou dans le transport de l'arginine (quand l'arginine est présent dans le milieu.). Beaucoup de ces gènes sont organisés en clusters et opérons, comme les opérons *argECBH*, opéron contenant des gènes de la voie de synthèse de l'arginine à partir du glutamate, et *carAB*, contenant des gènes codants les sous-unités A et B de la CPSase. On a récemment découvert des clusters pour les gènes codant des sous-unités des systèmes de transport pour l'arginine et l'histidine. Les clusters sont les suivants, *artPIQM-artJ*, permettant la transcription des gènes codant un système de transport de l'arginine dans la cellule et *argT-hisJQMP* qui permet la transcription des gènes codant un système de transport de l'arginine, de la lysine, de l'ornithine et de l'histidine dans la cellule (figure 32 et 33). Il est important de remarquer que ces deux

systèmes fonctionnent grâce à l'hydrolyse de l'ATP (Caldara *et al*, 2006 ; Caldara *et al*, 2007).

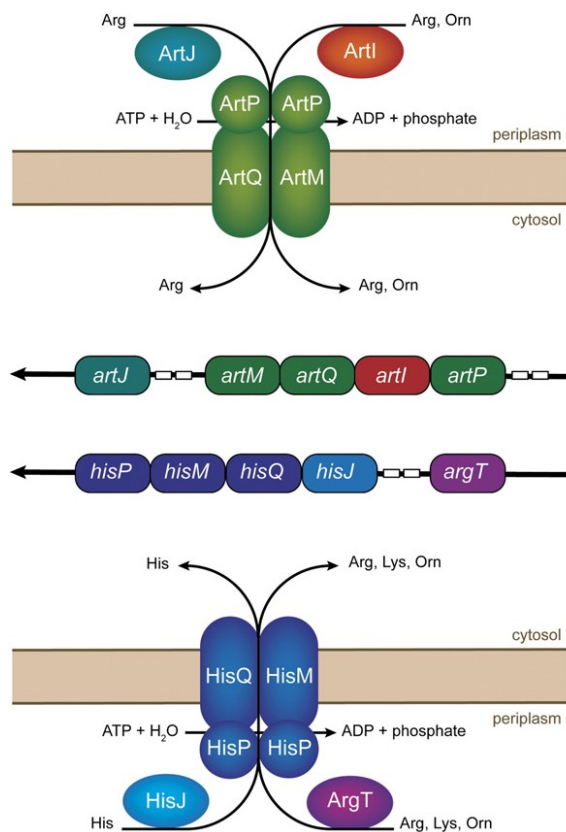


**Figure 32.** Représentation des opérons *argT-hisJQMP* et *artPIQM-artJ* permettant respectivement la fabrication des systèmes de transport pour l'arginine et le système de transport pour l'arginine, la lysine, l'ornithine et l'histidine.

**A)** Schéma représentant l'organisation de l'opéron *argT-hisJQMP*.

**B)** Schéma représentant l'organisation de l'opéron *artPIQM-artJ*, les deux boîtes arginine comprises dans cet opéron y sont soulignées.

Tous ces systèmes de transport sont ATP dépendants (avec permission de reproduction, Caldara *et al*, 2006).



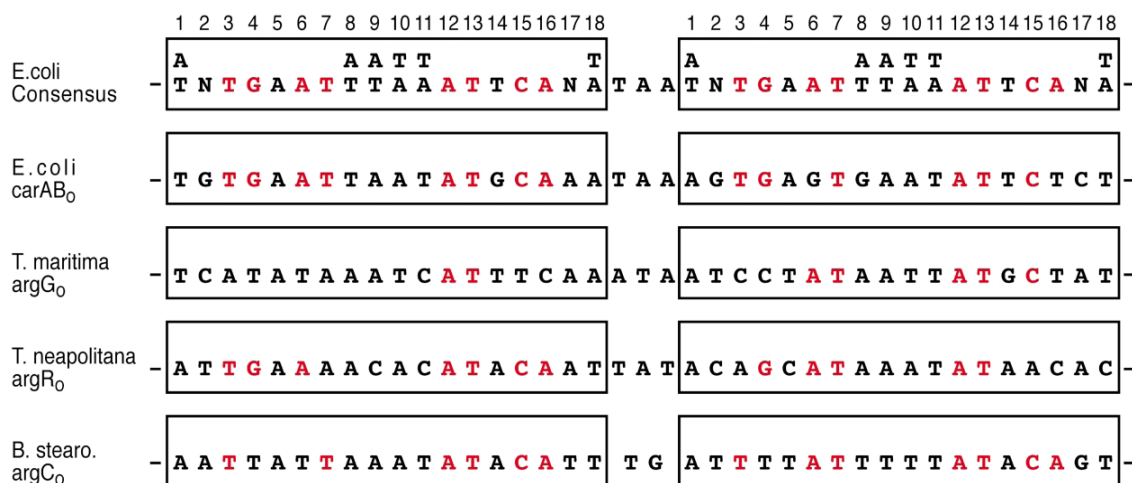
**Figure 33.** Schémas représentant les systèmes de transports générés respectivement par les opérons *artPIQM-artJ* et *argT-hisJQMP* (on revoit d'ailleurs l'organisation des opérons cités ci-dessus). Ces deux systèmes sont ATPasiques et permettent le transport des molécules d'arginine, de lysine, d'histidine et d'ornithine du périplasm au cytosol de la cellule. Les rectangles blancs situés sur les deux opérons correspondent aux boîtes arginine (sachant que deux rectangles à la suite correspondent à une seule boîte) (avec permission de reproduction, Caldara *et al*, 2007).

Il est à noter que les hexamères font 102 kDa, et que la protéine ArgR régule sa propre expression par fixation sur son propre promoteur (Lim *et al*, 1987), puisque la

boîte arginine (la séquence de liaison d'ArgR à l'ADN) se retrouve dans la région promotrice du gène *argR*, plus précisément dans la région contenant les promoteurs P1 et P2, contrôlant la transcription du gène (figure 34).



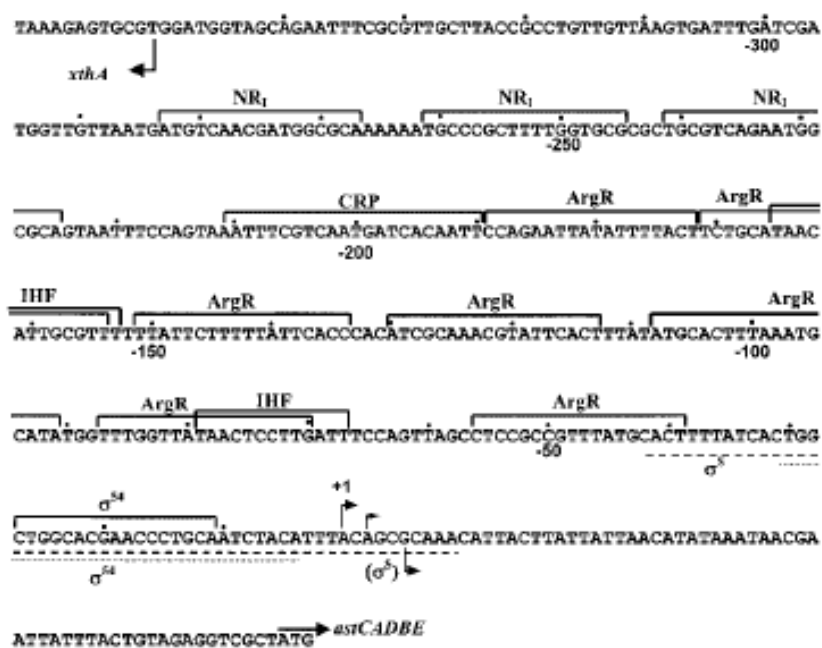
**Figure 34.** Séquence nucléotidique du gène *argR* avec les promoteurs P1 (en orange) et P2 (en bleu). Les régions -10 et -35 de P1 se superposent avec la boîte arginine (qui est fortement dégénérée et soulignée sur cette figure) dont les séquences sont soulignées. Le codon Méthionine a été écrit en minuscules et en vert. Le codon stop a été écrit en minuscule et en rouge (inspiré de Lim *et al*, 1987, séquence obtenue sur le site EcoCyc).



**Figure 35.** La boîte arginine. Séquence consensus d'*E.coli* que l'on retrouve dans les différentes souches de cette bactérie ainsi que celle pour l'opéron *carAB* et celles d'autres micro-organismes, dont *Thermotoga maritima*, *Thermotoga neapolitana* et *Bacillus stearothermophilus* (avec permission de reproduction, Morin *et al*, 2003). Le nombre de bases présent dans chaque demie-boîte arginine est également, de 18. C'est sur ce type de séquences que la partie Nt des hexamères d'ArgR peut se fixer.



Dans le système de régulon ArgR, il y a toujours une boîte arginine au niveau des promoteurs de chaque unité transcriptionnelle de ce même régulon (Cunin *et al*, 1983). Cette boîte arginine est très conservée chez les bactéries, même si elle se présente de manière dégénérée, parfois au sein du même organisme. On retrouve donc cette boîte autant chez des bactéries gram négatives, comme *E.coli*, *S.typhimurium*, *T.maritima*, *T.neopolitana*, mais aussi chez des bactéries gram positives comme *B.stearothermophilus* et *B.subtilis* (figure 35). De plus, il a été mis en évidence que l'ArgR est un co-activateur de l'opéron *astCADBE* (figure 36) (Kupakis *et al*, 2002) au sein d'*E.coli* dans la voie de catabolisme de l'arginine donnant du glutamate et du succinate après la dégradation de l'arginine par la voix AST (Schneider *et al*, 1998). Cette activation a lieu quand l'arginine est une des seules sources d'azote disponible pour la bactérie, et lorsque la L-arginine c'est liée à ArgR. Trois boîtes arginine ont été identifiées dans le promoteur de cet opéron et trois hexamères d'ArgR peuvent se lier au promoteur, du moins *in vitro* (Kupakis *et al*, 2002). La liaison des hexamères courberait l'ADN et faciliterait les interactions entre NR<sub>1</sub> et E  $\sigma^{54}$ . Il est de plus possible qu'un activateur non identifié soit présent et qu'ArgR amplifie son rôle (Kupakis *et al*, 2002).



**Figure 36.** Promoteur de l'opéron *astCDABE*. La séquence du promoteur *ast* est ici montrée avec les sites de liaisons de nombreux facteurs. Le site de départ de la transcription est montré en +1 précédé part une suite de points. Les sites de liaisons des protéines ArgR, NR<sub>1</sub>, CRP, IHF et des sous-unités  $\sigma^{54}$  et  $\sigma^S$  de la polymérase sont indiqués (avec permission de reproduction, Kupakis *et al*, 2002).

Si l'on sait que la forme hexamérique de ArgR se lie à l'ADN et permet de réprimer les gènes du régulon, on sait aussi que les trimères peuvent également se lier à l'ADN, mais sans pouvoir réprimer les gènes de ce même régulon, ni générer la recombinaison site spécifique en *cer*. Lorsque la protéine ArgR rentre en contact avec sa boîte grâce à sa partie Nt, l'ADN s'enroule autour de la protéine ArgR en faisant quatre tours grâce à son sillon majeur. On s'est également aperçu qu'ArgR peut se lier sur les deux sillons mineurs externes de l'ADN sur quatre bases notées AT/TA (Wang *et al*, 1998), et ce sur la partie Nt de la protéine. On a remarqué que la partie Nt d'ArgR possède une structure semblable à celle des protéines de la grande famille des hélices-boucles-hélices ailées (soit *winged helix-turn-helix* ou wHTH). Cette famille est caractérisée par la présence dans leur structure de trois hélices  $\alpha$ , dont deux avec le motif hélice-boucle-hélice qui se trouve encadré par feuillet  $\beta$  (qui forment les fameuses ailes). Pour ArgR, lors du contact avec l'ADN de la boîte arginine, c'est la troisième hélice  $\alpha$  avec les feuillet  $\beta$  qui permet ledit contact. Ce qui entraîne une courbure de l'ADN de l'ordre de  $65^\circ$  (Tian *et al*, 1992 ; Sunnerhagen *et al*, 1997), et finalement la répression des gènes de la voie de synthèse de l'arginine. Si des protéines homologues à ArgR d'*E.coli*, telle AhrC, ont déjà été cristallisées, ArgR d'*E.coli* ne l'a été que partiellement. On remarquera que d'autres bactéries, possèdent des homologues à ArgR. Notamment les bactéries *S.typhimurium*, *B.stearothermophilus*, *H. influenza*, *T. maritima*, *T. neopolitana* et aussi *B. subtilis* qui possède l'homologue AhrC, et dont la structure en Ct possède des similitudes avec celle d'ArgR (Van Duyne *et al*, 1996). D'autre part, les ArgR de *T.neopolitana* ont montré qu'elles pouvaient se lier à seulement une demie-boîte arginine de manière assez forte car il n'a qu'une perte d'affinité de deux à six fois moindre comparée à l'affinité de la protéine ArgR avec sa boîte entière. (Morin *et al*, 2003). Ce qui a entraîné la distinction des répresseurs en trois classes.

La première classe est dépendante de la L-arginine pour être fonctionnelle et requiert la boîte arginine en entier pour se lier à l'ADN (l'espaceur est composé des bases A et T) et empêcher la RNAPolymérase de se lier à l'ADN. ArgR d'*E.coli* fait partie de cette classe.

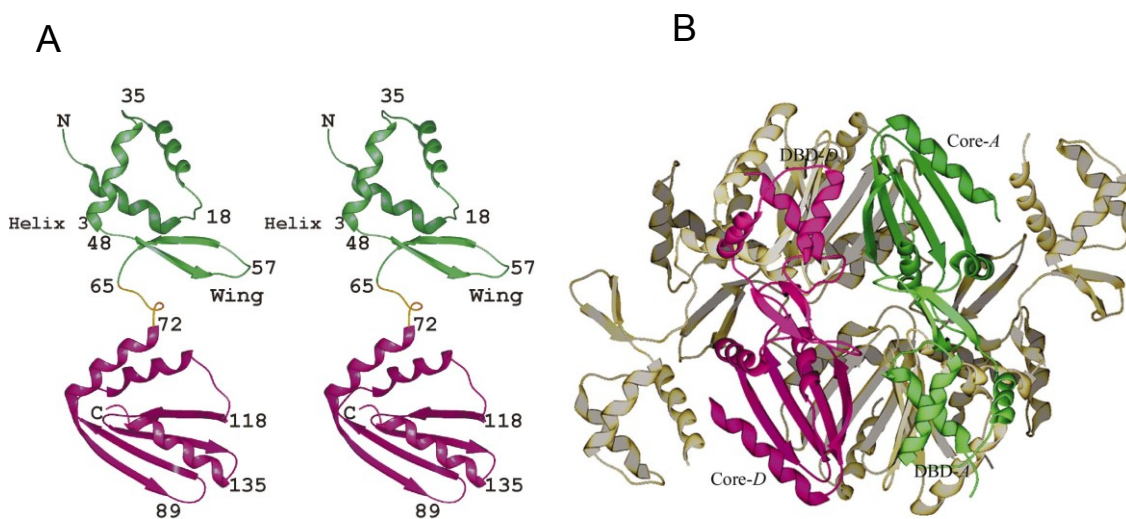
La seconde classe comprend des répresseurs issus des bactéries gram positives telles *B.subtilis* (dans ce cas, AhrC). Ils ne présentent que 25 à 30 % d'homologies avec la classe I, mais ils peuvent s'hexamériser en présence de L-arginine et présentent des structures tridimensionnelles proche de la classe I. De plus leur boîte arginine est organisée similairement à celles de la classe I.

La troisième classe est surtout représentée par les ArgR de *Thermotoga maritima* et *Thermotoga neopolitana*. Ces répresseurs peuvent se lier à une demie-boîte arginine et n'ont pas besoin de la L-arginine pour être fonctionnels. De plus, ils peuvent se lier facilement aux boîtes des autres classes.

### **I.9. La protéine AhrC**

La bactérie *Bacillus subtilis* (gram positive) possède un homologue à ArgR : la protéine AhrC. Sa découverte fait suite à la résistance à la L-arginine hydroxamate (un analogue à la L-arginine) de certaines bactéries de cette espèce toujours capables de surproduire la L-arginine. Cette résistance a été localisée sur quatre loci, l'un d'entre eux fut identifié comme l'homologue à *argR* et fut nommé arginine hydroxamate résistant mutant soit *ahrC* (Harwood *et al*, 1977). AhrC comprend au total 27 % d'identité avec ArgR, mais ce pourcentage s'élève à 35% dans la partie Ct des deux protéines (North *et al*, 1989) et à 28% dans la partie Nt. Comme ArgR, AhrC fonctionne, sous forme d'hexamères, comme un répresseur de la voie de biosynthèse de la L-arginine permise par les gènes *argC* et *argG*, en interagissant avec des séquences palindromiques similaires à la boîte arginine (Smith *et al*, 1989). Encore comme ArgR, AhrC a besoin de la L-arginine pour fonctionner. Les palindromes cités plus haut sont retrouvés en amont de deux unités de transcription regroupant les sept gènes du régulon de l'arginine. Il est à noter que la protéine AhrC joue un rôle multifonctionnel dans *B. subtilis* en servant de répresseur, mais également d'activateur de gènes impliqués dans le catabolisme (Van Duyne *et al*, 1996). Pour *E.coli* les palindromes sont séparés de 3 pb. Bien que la boîte arginine soit similaire, chez *B. subtilis* (GATGAATAA-

AAATg/tCAAg/t) les boîtes arginine se voient séparées de 11 pb au niveau de promoteur du gène *argC*. Il y a une autre boîte en aval, et ce dans la séquence codante du gène *argC*. De plus cette protéine fonctionne également comme un activateur des gènes *rocA* et *rocD* permettant la catabolise de la L-arginine. Tout comme ArgR, AhrC s'hexamérise en sous-unités de 149 acides-aminés en présence de L-arginine (Garnett *et al*, 2008). Le domaine Ct sans L-arginine montre une symétrie de type 32 sous forme cristallisée (figure 37).



**Figure 37.** La protéine AhrC, sous-unité et hexamère.

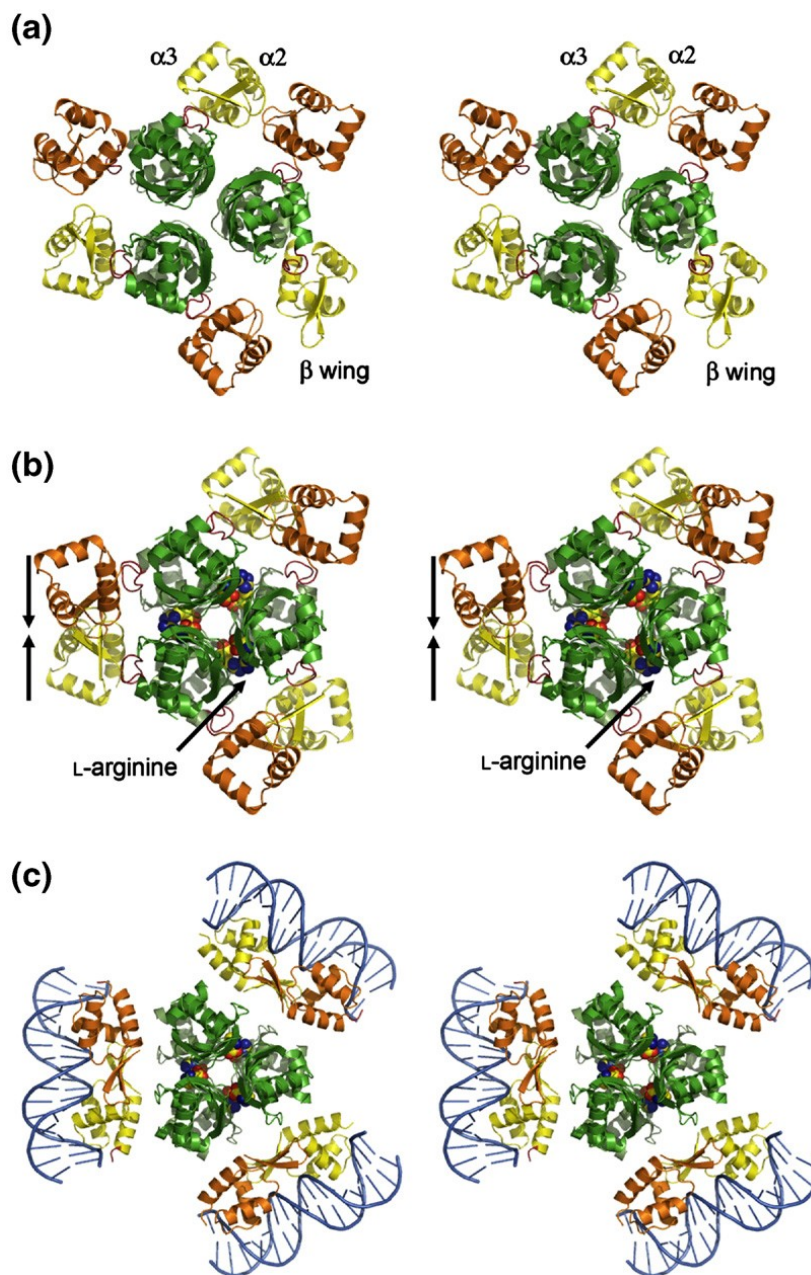
**A)** Structure monomérique de la protéine AhrC de *B.subtilis*. Ces deux représentations montrent en fait quelques différences minimales de structure tridimensionnelle. La partie Ct, qui permet l'oligomérisation, est colorée en magenta. La partie Nt, qui permet la liaison à l'ADN (*DNA binding domain*, soit DBD) est colorée en vert.

**B)** Structure hexamérique de la protéine AhrC de *B.subtilis*. Ici, deux sous-unités différentes sont représentées d'une couleur différente, vert pour l'une et magenta pour l'autre, on remarquera le fait que les parties Nt sont rejetées en périphérie de cet hexamère (avec permission de reproduction, Dennis *et al*, 2002).

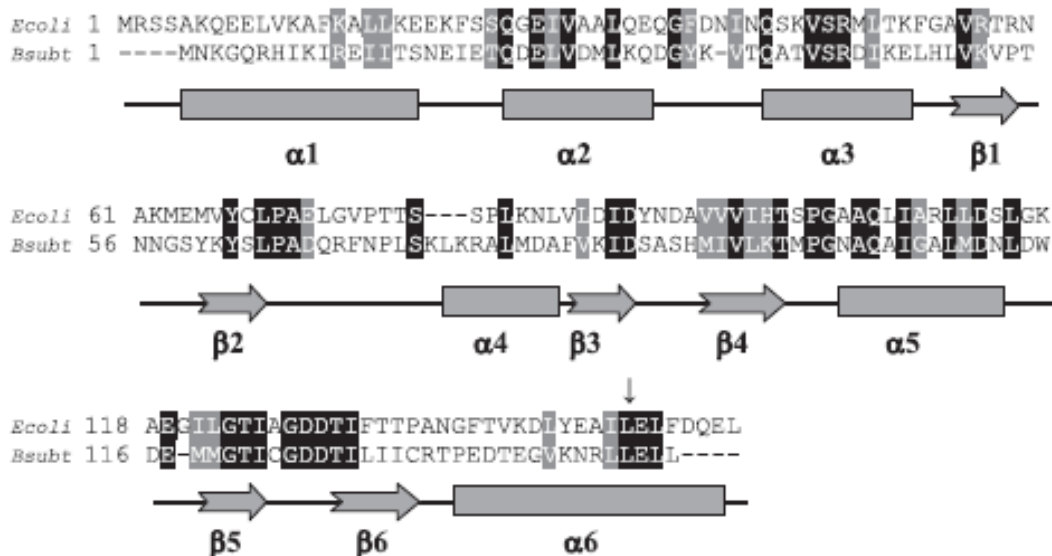
Les deux protéines homologues vont se lier à des opérateurs entre 38 et 47 pbs formés de deux palindromes imparfaits de 18 pbs. Pour les promoteurs des gènes de catabolise de l'arginine, on trouve une boîte adjacente directement au site de départ de la transcription de *rocA* et *rocD*. On suppose que la forme hexamérique d'AhrC doit être assez proche de celle d'ArgR, car AhrC pouvant compléter des *E.coli* est déficiente en ArgR (Stirling *et al*, 1988b).

En effet, lors d'expériences mettant en jeu des plasmides avec deux séquences *cer*, la protéine AhrC a permis de réaliser une recombinaison au niveau de cette séquence *cer* après complémentation, reprenant le rôle de la protéine ArgR (Stirling *et al*, 1988b). Pourtant, l'inverse n'est pas vrai, ArgR ne peut compléter des bactéries *B.subtilis*. Fait intéressant, AhrC présente une séquence plus courte qu'ArgR, 149 acides-aminés pour AhrC et 156 acides-aminés pour ArgR, la différence de longueur se répercutant surtout dans les parties Nt et Ct (figure 39).

Une récente étude du complexe hexamérique d'AhrC suggère qu'avec la liaison de la L-arginine au niveau de la partie Ct de la protéine, il y a un changement de conformation de l'hexamère. En effet, les parties Nt d'un des trimères vont subir une flexion de 15°, rendant la symétrie 32 imparfaite, mais surtout permettant aux deux parties Nt (qui sont des DBD, soit *DNA Binding Domain*) de se rapprocher et de favoriser la liaison avec la boîte arginine (figure 38) (Garnett, *et al*, 2010).



**Figure 38.** Différents modèles pour la liaison de l'hexamère d'AhrC sur la boîte arginine. Les domaines Ct sont en vert, les domaines Nt sont colorés en jaune pour le trimère du dessus, et en orange pour le trimère du dessous, la partie flexible est colorée en rouge. (a) Modèle de l'apo-AhrC, les parties Nt sont en périphérie de l'hexamère, qui montre une symétrie 32 imparfaite. (b) Modèle de l'holo-AhrC, c'est ici que la liaison de la L-arginine entraîne la flexion d'un des trimères, rapprochant les deux domaines Nt (flèches), la partie subissant la flexion est en rouge. (c) Modèle de l'holo-AhrC lié à trois ARG-boxes. Si l'hexamère se lie sur un de ses sites, il se peut qu'il y ait un « réarrangement » des domaines Nt pour reconnaître le site de liaison en les mettant dans la bonne position (avec permission de reproduction, Garnett *et al*, 2010).

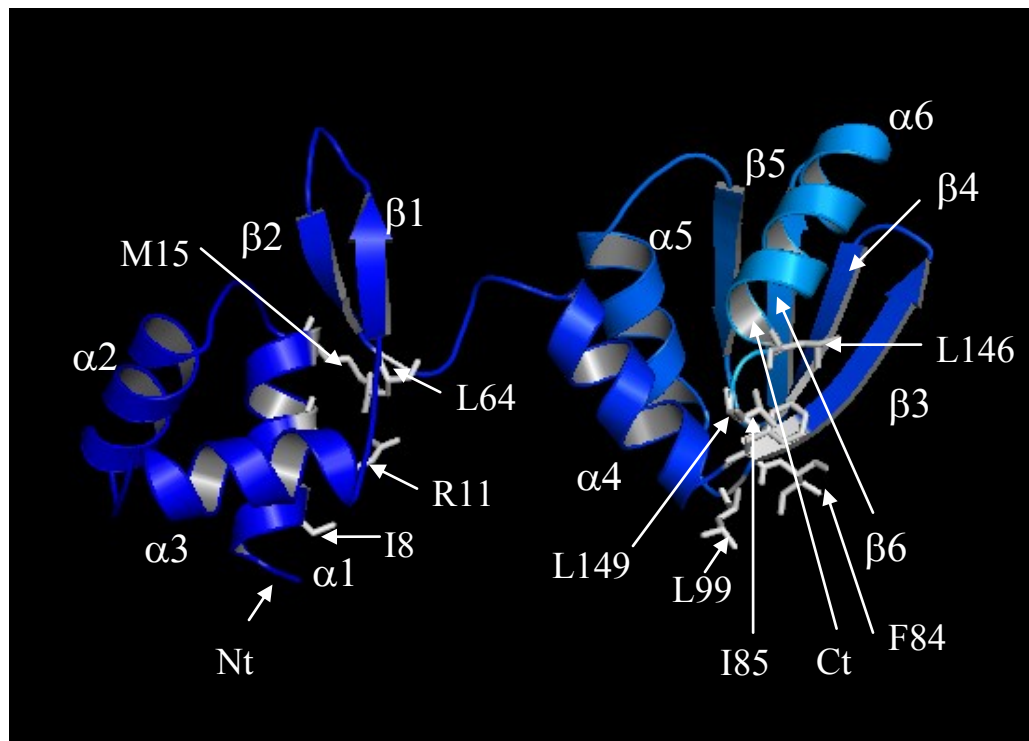


**Figure 39.** Alignement de séquences entre les protéines ArgR d'*Escherichia coli* et AhrC de *Bacillus subtilis*. Les acides-aminés identiques sont surlignés en noir, tandis que les acides-aminés de fonction similaire sont surlignés en gris. Les positions des hélices  $\alpha$  et des feuillets  $\beta$  sont indiquées sous l'alignement de séquences (avec permission de reproduction, Sénéchal *et al*, 2010).

### I.10. La protéine ArgR de *Bacillus stearothermophilus*

La bactérie *Bacillus stearothermophilus* possède également une protéine homologue à AhrC et ArgR d'*E.coli*, une protéine également appelée ArgR, mais notée ArgRBst. Comme les protéines ArgR d'*E.coli* et AhrC, cette protéine forme des hexamères à partir de six sous-unités d'ArgRBst. De plus, comme ArgR, elle a besoin elle aussi de la L-arginine pour être fonctionnelle (Ni *et al*, 1999). Cette L-arginine se fixe aussi au niveau de la partie Ct de la protéine et permet de rendre l'hexamère fonctionnel d'une façon comparable, c'est-à-dire par interaction entre la molécule de L-arginine et les sous-unités du trimère. Comme pour l'ArgR d'*E.coli*, l'ArgRBst forme un dimère de trimères pour être fonctionnelle, et lorsque l'on regarde la structure tridimensionnelle d'ArgRBst, qui a été entièrement cristallisée, on s'aperçoit que malgré une plus courte longueur (149 acides-aminés pour ArgRBst contre 156 pour ArgR), la structure et l'organisation des parties Ct sont comparables et similaires

(figure 40). De plus, les similitudes ne sont pas limitées à ArgR, mais lorsque l'on compare les hexamères d'ArgRBst et d'AhrC, ils sont presque superposables, leurs tailles et leurs structures spatiales sont similaires.



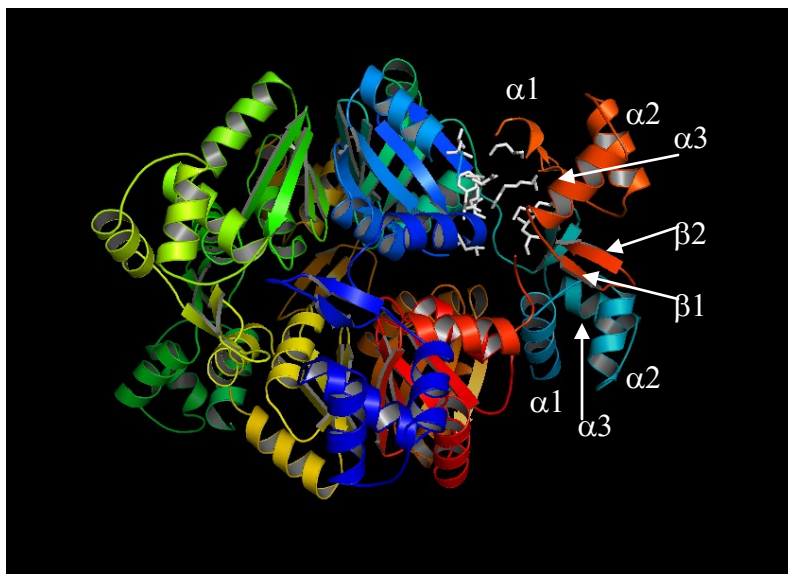
**Figure 40.** Représentation tridimensionnelle d'un monomère d'ArgRBst. Les hélices  $\alpha$  et les feuillets  $\beta$  sont représentés en bleu. Les parties Nt et Ct sont également visibles ainsi que les hélices  $\alpha$  et les feuillets  $\beta$ . Les acides-aminés qui sont considérés responsables de l'interaction entre deux sous-unités et permettant la stabilisation de l'hexamère sont indiqués (adapté de Ni *et al*, 1999; logiciel utilisé PyMOL).

Ce qui a été mis en évidence pour la protéine ArgRBst, c'est que les interactions entre les deux trimères ne sont pas limitées aux parties Ct où se fixe la L-arginine. En effet, une interaction entre les parties Ct et Nt de la protéine hexamérisée a été mise en évidence (Ni *et al*, 1999). Cette interaction se fait entre la partie Ct d'une des sous-unités, et la partie Nt d'une sous-unité située sur l'autre trimère. Ce qui stabilise l'hexamère formé (figure 41 et 42).

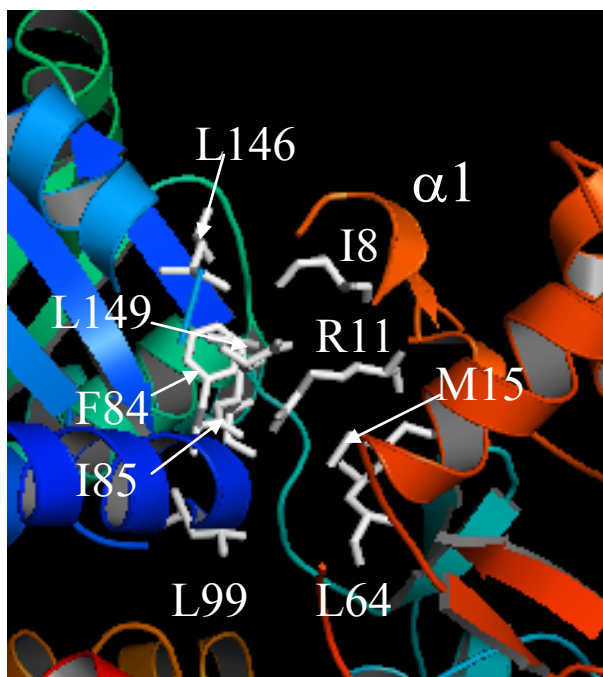
Même si seulement la partie Ct d'ArgR a été cristallisée, étant donné que les structures tridimensionnelles d'ArgRBst, d'AhrC et ArgR sont semblables entre les parties Ct, et que les structures tridimensionnelles complètes d'ArgRBst et AhrC sont



également similaires, il a été constaté que la structure Nt d'ArgR est également aussi similaire à celle de ses homologues. Il y a donc probablement interaction entre la partie Nt et Ct d'ArgR comme pour ArgRBst.



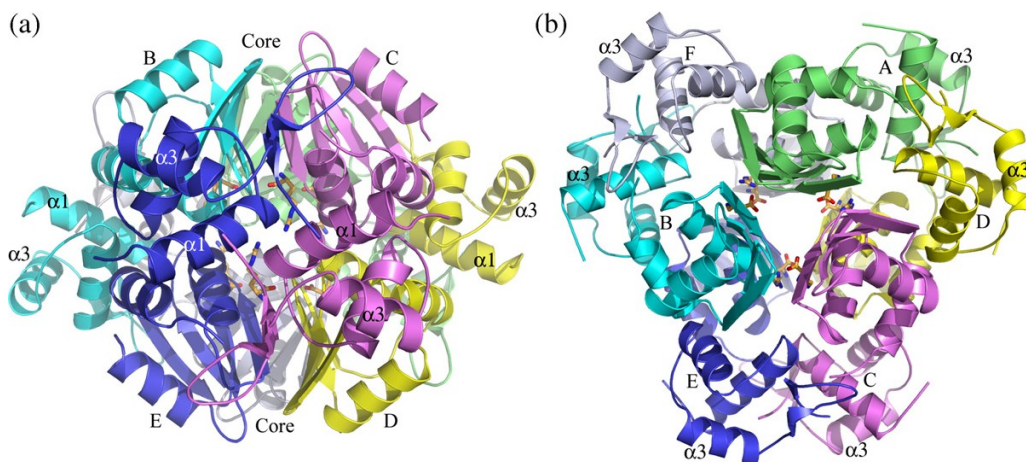
**Figure 41.** Représentation de l'hexamère d'ArgRBst. On peut voir que cet hexamère a une structure similaire à celle d'AhrC. Sur cette représentation tridimensionnelle de la molécule, on peut deviner les résidus Ct qui interagissent avec les résidus Nt d'une autre sous-unité d'ArgRBst (tous représentés en blanc). Chaque sous-unité est représentée par une couleur différente (adapté de Ni *et al*, 1999; logiciel utilisé PyMOL).



**Figure 42.** Représentation tridimensionnelle de l'interface entre la partie Ct et Nt d'ArgRBst. Sur ce schéma, on peut voir en bleu la partie Ct et en orange, la partie Nt. Les résidus L146, L149, F84 et L99 de la partie Ct interagiraient avec les résidus I8, R11, M15 et L64 de l'hélice  $\alpha 1$  de la partie Nt provenant d'une sous-unité située sur l'autre trimère (adapté de Ni *et al*, 1999 logiciel utilisé PyMOL).

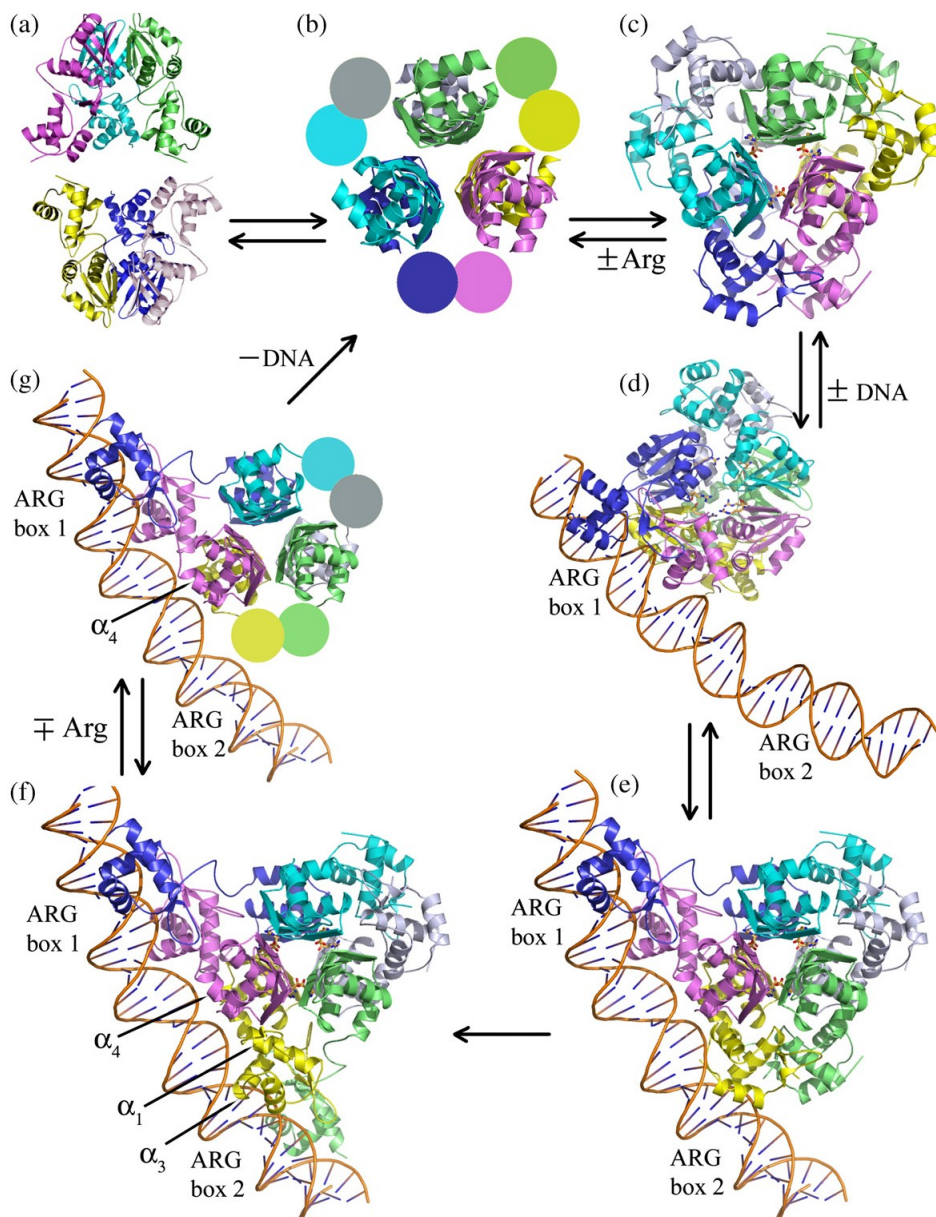
### I.11. La protéine ArgR de *Mycobacterium tuberculosis*

La protéine ArgR provenant de la bactérie *Mycobacterium tuberculosis* est également capable de se multimériser et de se lier à la L-arginine. Comme pour les protéines AhrC, ArgRBst et ArgR, la sous-unité forme des dimères de trimères, donnant des hexamères fonctionnels capables de se lier à l'ADN au niveau d'une boîte arginine régulant la voie de synthèse de l'arginine. Comme pour les hexamères d'AhrC et d'ArgRBst, la partie Ct est proche de la partie Nt pour les hexamères d'ArgRMtb non liés à l'ADN. En effet l'hélice  $\alpha_2$  de la partie Nt (appelée aussi DBD, soit *DNA-Binding-Domain*) est proche de la partie Ct dans cet hexamère (figure 43) (Cherney *et al*, 2010).



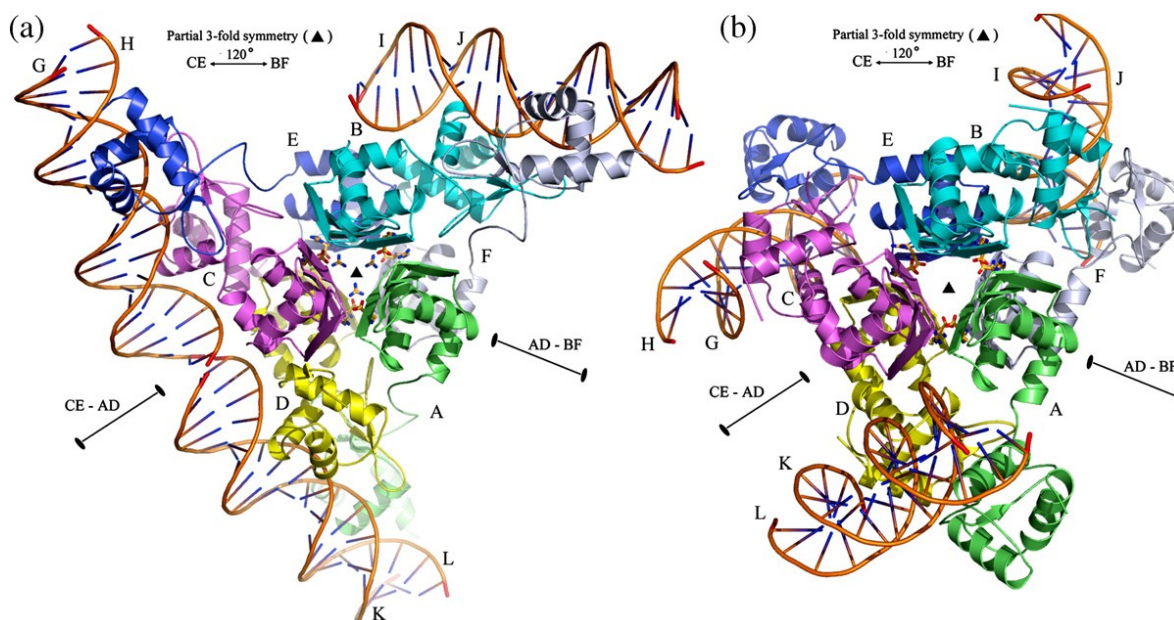
**Figure 43.** Représentation tridimensionnelle de l'hexamère d'ArgR de *Mycobacterium tuberculosis*. En (a), l'hexamère « de côté » et les deux trimères liés par leurs parties Ct ont été indiqués. En (b), l'hexamère d'ArgRMtb a été visualisé de « dessus ». Les molécules de L-arginine sont placées au centre et les parties Nt (DBD) sont rejetées en périphérie (avec permission de reproduction, Cherney *et al*, 2010).

Lorsque l'hexamère se fixe sur l'ADN au niveau des hélices  $\alpha_1$  et 3 de la partie Nt, il y a un changement de conformation de l'hexamère qui est flexible justement au niveau de cette partie Nt qui va se lier à l'ADN (figure 44) (Cherney *et al*, 2010). La liaison entre l'ADN et la protéine suivrait des étapes précises (Cherney *et al*, 2010). Il est fort probable que les hexamères d'ArgRBst, AhrC et ArgR se lient à l'ADN selon un mécanisme similaire.



**Figure 44.** Modèle de liaison de l'hexamère d'ArgR de *Mycobacterium tuberculosis* à sa boîte arginine. Les parties (a) et (b) montrent des apo-trimères contenant la partie Ct de l'hexamère. Les positions supposées des parties Nt sont représentées par des cercles colorés. En c), l'hexamère lié à la L-arginine a été représenté. En d), le complexe commence à se former, avec la liaison d'un DBD sur la ARG box 1, les autres DBD n'ont pas encore subi de changement de conformation. En e) le domaine Nt adjacent à celui qui s'est déjà lié à la ARG box 1 se lie à son tour sur la ARG box 1. Deux autres DBD (en jaune et vert), sont désormais positionnés pour se lier à la ARG box 2. En f), les DBD qui étaient positionnés se lient à la ARG box 2, avec changement de conformation. Les hélices  $\alpha_1$ ,  $\alpha_3$ ,  $\alpha_4$  sont montrées, la 1 et la 4 étant rapprochées l'une de l'autre, la 3 étant en contact avec l'ADN. En g) un désassemblage du complexe lorsqu'il y a perte de la L-arginine y a été montré (avec permission de reproduction, Cherney *et al*, 2010).

Le modèle présenté explique bien la liaison à une boîte arginine complète. Dans le même article (Cherney *et al*, 2010), un deuxième modèle avait été présenté, mais il n'y avait pas de liaison à la boîte arginine en entier, ce qui aurait entraîné des liaisons trop faibles à l'ADN. Il est fort probable qu'il en soit de même avec les autres ArgR et AhrC lors de la résolution des dimères plasmidiques (figure 45).



**Figure 45.** Représentations des complexes ADN-*MtbArgR* possibles. En a) la structure appelée complexe II, présentée dans la figure 44 y a été schématisée. Les deux DBD (*DNA binding domain*) en gris et turquoise sont capables d'interagir avec une boîte arginine. En b) la structure appelée complexe I a été modélisée, la symétrie est similaire, mais ce complexe serait instable et il n'y a pas de liaison à une boîte arginine entière (avec permission de reproduction, Cherney *et al*, 2010).

### I.12. Modèles d'interaction ArgR-PepA dans le système Xer/*cer*

Étant donné que l'on sait que les protéines ArgR et PepA sont les protéines accessoires impliquées dans la recombinaison en *cer*, la façon dont elles interagissent est le sujet de nombreux modèles suggérés au cours du temps.

Mais c'est en 1997 que les travaux de Claudia Alén *et al* ont établi que les protéines PepA et ArgR interagissaient sur le site *cer*. Avant ces travaux, on ne faisait

que supposer une interaction car la recombinaison en *cer* se faisait seulement en présence de ces deux protéines accessoires. C'est tout d'abord à l'aide de migrations sur gels de chloroquine (après traitement à la topoisomérase 1) d'un plasmide noté pSDC15 en présence des protéines accessoires qu'il a été démontré que PepA et ArgR se liaient à l'ADN. En effet, dans cette expérience, il y avait augmentation du nombre de supertours du plasmide pSDC15. Il a ensuite été démontré que cette augmentation de supertours (environ 3 supertours de plus en présence d'ArgR et PepA) advenait seulement quand la séquence *cer* était directement répétée. De plus il ne s'agissait pas d'un enroulement de type solénoïde, mais toujours plectonémique (Alén *et al*, 1997).

Toujours dans la même étude d'Alén *et al*, des expériences de protection à la DNase I sur la séquence *cer* ont montré que la protéine ArgR se fixait sur la partie correspondant à la boîte arginine (induisant une courbure de 65°), mais qu'en présence de PepA, il y avait une modification du schéma de protection de l'ADN contenant la séquence *cer*, démontrant ainsi que PepA se fixait sur une partie de cette séquence. Par contre, cette étude n'avait pas permis de savoir combien de protéine PepA se fixaient sur la séquence *cer*. Elle avait généré deux modèles possibles pour le complexe ADN-ArgR-PepA : un modèle où une seule molécule hexamérique de PepA interagit avec l'hexamère d'ArgR et l'ADN, et un autre modèle théorique où il y avait deux hexamères de PepA dans le complexe qui interagit avec l'ADN, mais toujours un seul hexamère d'ArgR.

C'est après cette étude de 1997 qu'un des premiers modèles cristallographiques datant de 1999 (Sträter *et al*, 1999) a été mis au point.

C'est à la même époque et au début des années 2000 (Paul and Summers, 2004 ; Gourlay and Colloms, 2004) que l'on s'est servi des protéines PepA et ArgR pour mettre au point des expériences qui semblaient indiquer que la protéine Cre du bactériophage P1 avait besoin d'ArgR (et de PepA) comme les protéines XerC et XerD, lors de la recombinaison au niveau du site *loxP*.

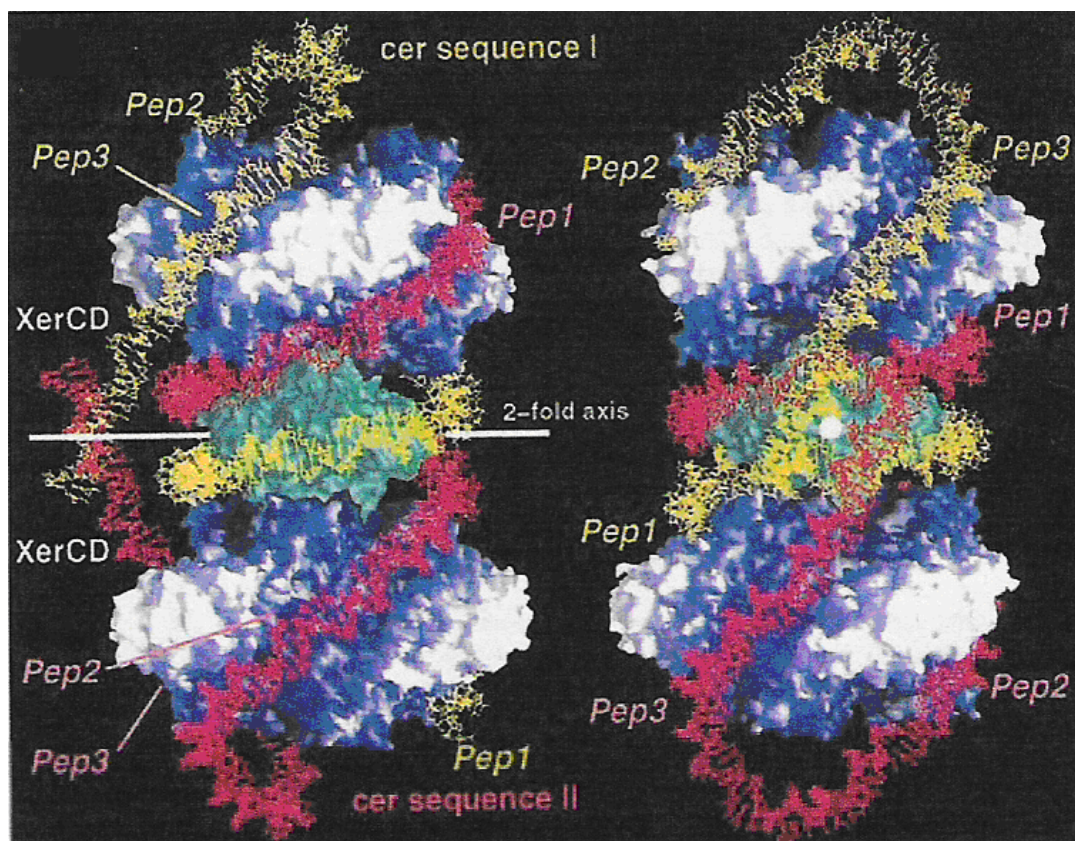
En effet, on trouve une séquence très analogue à la boîte arginine au niveau de *loxP*, cette séquence, très similaire à celle retrouvée dans la séquence *cer*, semblait indiquer qu'ArgR serait recruté pour l'action de la recombinaise Cre.

Des constructions comprenant les sites de liaison pour ArgR et PepA, et le site de liaison pour Cre semblèrent aller dans ce sens, puisqu'il y avait formation d'un complexe de recombinaison avec ces constructions lorsqu'elles étaient directement répétées, lesdites constructions étant ensuite recombinées par Cre. Dans ce système, Cre remplaçait XerC et XerD. Mais en fait la formation du complexe ArgR-PepA et des *four-noded* caténanes qui en résultaient était due aux séquences venant de *cer*.

En effet, les résolutions et recombinaisons observées étaient dues à Cre seule, aucun qui n'avait pas besoin des protéines ArgR et PepA pour effectuer la recombinaison. Le fait que la recombinaison due à Cre soit ArgR indépendante malgré la présence d'une séquence proche d'une boîte arginine, a été confirmé en 2008 par l'équipe de Sean D. Colloms (MacDonald *et al*, 2008). La séquence *loxP* dépourvue de tout ajout de séquence provenant de *cer* ne faisait pas de recombinaison avec la protéine Cre en présence d'ArgR et PepA. L'une des données importantes est que dans cette étude, un mutant de Cre, nommé Cre1, a besoin de PepA, lorsque l'on met ce mutant en présence d'un plasmide sans séquence provenant de *cer*, il n'y a pas de recombinaison. Mieux encore, Cre n'a besoin que de *loxP* pour fonctionner (MacDonald *et al*, 2008).

Toujours au début des années 2000, il a été démontré que l'on avait des problèmes de résolution des dimères plasmidiques, ou de plasmides contenant deux séquences *cer*, lorsque l'on avait des mutations au sein de PepA, soit dans la région centrale, soit dans la région Ct de PepA (Charlier *et al*, 2000). Cela nous a amené à penser que la région Ct était importante dans la recombinaison *cer* spécifique, sans doute par contact avec les hexamères d'ArgR. Le modèle de Sträter, déjà décrit par le passé, représente une structure tridimensionnelle du complexe ADN-protéines avec un hexamère d'ArgR et deux hexamères de PepA. Dans ce modèle, l'hexamère d'ArgR

occupe la position centrale et l'ADN des deux copies de *cer* vient s'enrouler autour de ce complexe, sachant que l'hexamère d'ArgR est le premier à se fixer sur l'ADN dans ces modèles. C'est seulement après que les protéines XerC et XerD, viennent se joindre au complexe pour générer la recombinaison (figure 46). En effet, l'enroulement de l'ADN autour du complexe a suffisamment rapproché les sites de liaison pour que les recombinases XerC et XerD puissent effectuer la recombinaison site spécifique en *cer*.



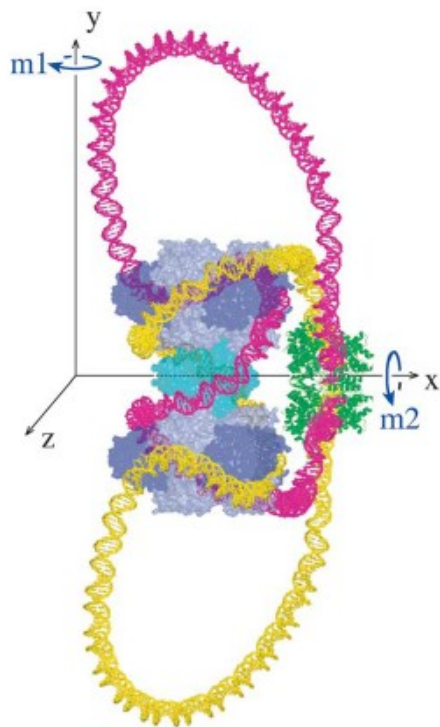
**Figure 46.** Complexe ADN/protéines ArgR-PepA (avec permission de reproduction, Sträter *et al.*, 1999), premier modèle. Il y a deux hexamères de PepA et un seul d'ArgR. Ce complexe a été représenté sous l'angle de l'axe 2-fold et également perpendiculairement à cet axe. Sur cette figure, l'hexamère d'ArgR est représenté en vert et les deux hexamères de PepA sont représentés en bleu et blanc. Chaque copie des gènes du dimère plasmidique (en l'occurrence ColE1) est représentée d'une couleur différente, rouge pour l'une et jaune pour l'autre. L'ADN se lie aux sillons Ct de PepA, et dans ce cas, il n'est pas très fortement en contact avec les domaines Nt. Les trois sillons de chaque hexamère de PepA sont occupés par de l'ADN. Il s'agirait de l'ADN des séquences Pep1, Pep2 et Pep3 détectées sur la séquence *cer*.

Il s'agit du plus ancien des modèles d'interaction entre les hexamères des protéines PepA et ArgR, il a été obtenu par cristallographie au rayon X (Sträter *et al*, 1999). Les deux protéines accessoires ArgR et PepA permettent une recombinaison exclusivement intramoléculaire des dimères des plasmides (figure 46) (voir des constructions contenant deux sites *cer* tel le plasmide pCS210).

Ce modèle a été mis au point grâce aux expériences d'empreintes à la DNase I. Un des sites de liaison présumé de PepA se superpose à une boucle d'ADN de 60 pb au niveau d'une zone d'hyperactivité de la DNase I (lorsqu'il n'y a pas de protection), l'autre, séparé de 10 pb, se superpose au site de liaison de XerC dans le site *cer* (Alén *et al*, 1997). À cette époque, on ignore si PepA possède une séquence de liaison qui lui est propre, ou si cette protéine se lie à des régions riches en bases A et T, c'est le cas de la région de 60 pb citée plus haut. Dans ce modèle datant de 1999, on ne peut pas trancher entre les deux modes d'interaction, il est possible que les deux soient présents lors de la liaison de l'hexamère de PepA à l'ADN (Alén *et al*, 1997 ; Sträter *et al*, 1999). De plus, des changements de conformation de PepA ne seraient pas à exclure lors de la liaison de l'hexamère de PepA à l'ADN, surtout au niveau N-terminal, domaine qui est justement sensé se lier à l'ADN. Le problème principal de ce modèle est que les sites réels de liaisons soient mal connus au niveau de l'ADN, même si on estime avoir de fortes présomptions.

Un autre modèle, plus récent et assez validé en 2005, suggère un enroulement différent de l'ADN autour des complexes. Il a d'abord été généré par calcul statistique de type Tangle (figure 47).

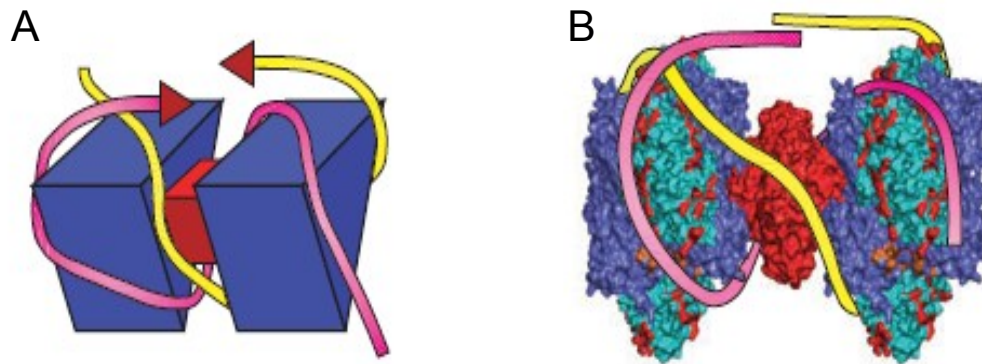




**Figure 47.** Nouvelle modélisation du complexe ADN-PepA-ArgR utilisant la méthode de calcul dite Tangle. Ici les protéines XerC et XerD sont représentées en vert, prêtes à effectuer la recombinaison (avec permission de reproduction, Vazquez *et al*, 2005).

La structure de l'enroulement de l'ADN autour du complexe présentée ici a été choisie parmi d'autres structures proposées dans l'article de Vazquez *et al*, 2005, pour les raisons suivantes : c'est celle qui propose le moins de contraintes de torsion de l'ADN autour du complexe protéique, et le moins de contraintes thermodynamiques.

Ce modèle obtenu par calcul, est désormais appelé modèle de Sträter modifié. Le modèle de Sträter modifié, fait partie des différents modèles d'interaction possibles pour ArgR-PepA. Il a la particularité de simplifier le trajet de l'ADN sur le complexe protéique de telle manière que l'ADN n'occupe désormais que deux des sillons Ct de chaque hexamère de PepA. De plus, il est compatible avec le modèle de Cherney *et al*, 2010, pour les interactions d'ArgR avec la boîte arginine (figure 48).



**Figure 48.** Seconde modélisation du complexe ADN-PepA-ArgR utilisant la méthode de calcul dite Tangle désormais appelée modèle de Sträter modifié (avec permission de reproduction, Reijns *et al*, 2005).

**A)** Version schématique du modèle avec polygones triangulaires.

**B)** Version modélisée en 3D du Sträter modifié.

Dans les deux figures, les deux copies des gènes d'ADN plasmidique sont de deux couleurs différentes, l'une est représentée en jaune, l'autre en rose.

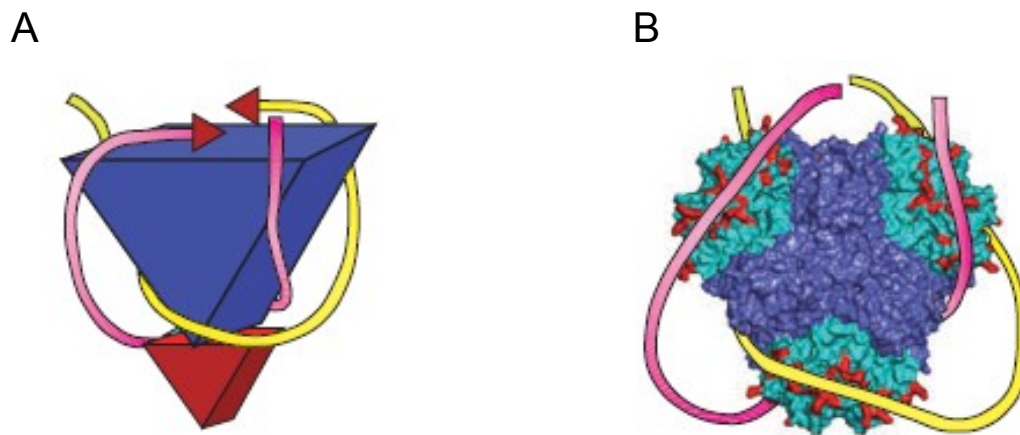
Dans ce modèle, l'ADN se lie au domaine Nt de PepA et passe à travers le sillon Ct de la protéine.

En résumé, dans les deux modèles de Sträter (original et modifié), ArgR hexamérisé va se lier à 2 hexamères PepA : seule la façon dont l'ADN s'associe au complexe change.

Un troisième modèle datant de 2005 montre une modélisation où il n'y a qu'un seul hexamère de PepA et un d'ArgR qui interagissent ensemble (figure 49).

Ce modèle a été mis-au-point par l'étude de mutations sur la partie Nt de PepA et a mis en évidence l'existence d'un chemin de résidus d'acides aminés basiques majoritairement constitué d'arginines et de lysines sur la surface de l'hexamère de PepA. Quand ce chemin était modifié par substitution d'acides-aminés, il y avait perte de l'activité de recombinaison au niveau de la séquence *cer*. Or cette organisation à la surface de PepA était jusque-là inconnue. De plus, ces mutations ont permis de constater que les acides-aminés susceptibles de permettre la liaison à l'ADN en Nt sont plus nombreux que dans le modèle de Sträter et s'étendent sur une plus grande région (Reijns *et al*, 2005).

Ce qui a conduit à revoir la manière dont le complexe synaptique se forme.



**Figure 49.** Modélisation du complexe formé par les protéines accessoires ArgR/PepA (avec permission de reproduction, Reijns *et al*, 2005).

**A)** Complexe synaptique *cer* avec en bas en rouge, l'hexamère ArgR, et en bleu l'hexamère PepA. Les brins d'ADN qui vont être recombinés sont représentés en rose et en jaune. Dans ce modèle, il n'y a qu'un hexamère de chaque protéine. Les deux parties de l'ADN qui se croisent sont en contact avec les deux faces de PepA.

**B)** Modèle 3D du complexe PepA.

Deux croisements de la molécule d'ADN dimérisée sont maintenus par deux des faces de l'hexamère de PepA. L'ADN se lierait en partie Nt de la protéine PepA, donc de l'hexamère. Les parties en rouge sur l'hexamère de PepA correspondent au chemin d'acides-aminés en Nt sur lequel l'ADN se lierait.

Ce modèle n'interdirait pas la présence de deux hexamères d'ArgR.

Ce nouveau modèle prend donc en compte les nouvelles données, et indique que l'ADN se lie le plus autour de l'hexamère de PepA. Il est même postulé que le déterminant majeur de la structure synaptique du complexe est PepA. Dans ce modèle à un seul hexamère de PepA, l'ADN passe par toute la partie Nt déterminée dans l'étude de 2005, délimitant un chemin par lequel passe l'ADN de la séquence *cer*, et par seulement deux sillons C-terminaux de cet hexamère.

On remarquera que ce modèle se base sur les positions des chemins d'ADN dans la partie Nt qui flanque chaque sillon Ct de PepA. Les sillons Ct sont flanqués par un autre chemin d'acides-aminés situés en Nt liant l'ADN dans une boucle au-dessus du sillon (figure 49).

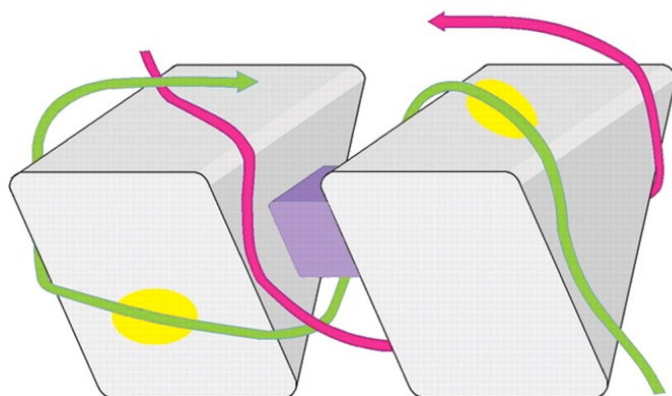
Il est important de signaler que la connaissance de ce chemin d'acides-aminés a été permise par la génération d'une suite de mutations qui empêchèrent la liaison de la protéine à l'ADN. Il y aura deux croisements d'ADN piégés sur les deux faces de l'hexamère de PepA. Les recombinaisons XerC et XerD vont se trouver piégés sur la troisième face.

Dans ce modèle, l'hexamère d'ArgR ne sert qu'à plier l'ADN dans la boucle qui passe par-dessus. Cela permet de voir pourquoi PepA peut définir la structure de la synapse et sa topologie (et de se lier seul à l'ADN si les hexamères sont surabondants). Mais il faut bien savoir que cette structure a du mal à être compatible avec le système Cre-XerCD pour des raisons purement thermodynamiques (Reijns *et al*, 2005). De plus, les parties d'ADN et de PepA se liant les unes aux autres sont différentes entre le modèle de Sträter modifié et le modèle présenté au-dessus figure 49. Mais ces deux modèles sont compatibles avec des expériences d'empreintes à la DNaseI (Alén *et al*, 1997).

Parmi tous les modèles d'interaction présentés, il s'est avéré que de nombreuses études et expériences montrent un complexe avec deux hexamères de PepA, souvent grâce à des études et expériences avec la DNaseI comme mentionné précédemment.

Cependant, avec l'étude menée par Minh en 2009 (Minh *et al*, 2009), le modèle de Sträter modifié semblerait privilégié. En effet il y a un très fort raccourcissement de l'ADN dans le complexe PepA-2*cer* de 800 pb visualisé par microscopie par force atomique. Ce qui semble indiquer un passage compliqué de l'ADN plus compatible avec deux hexamères de PepA (figure 50).

Lorsque l'on regarde la séquence *cer*, on ne perçoit pas de séquences bien précises pour la liaison des hexamères de PepA, même si on en a suggéré (les séquences Pep1, Pep2, Pep3) (Sträter *et al*, 1999).



**Figure 50.** Modèle modifié de Sträter *et al* proposant deux sites *cer* (vert et magenta) assemblés par 2 hexamères de PepA. L'ADN entre en contact avec la boucle Ct (en jaune) et les parties Nt adjacentes. L'hexamère ArgR est représenté en mauve. Il a été avancé que de toute façon PepA et sa liaison à l'ADN sont nécessaires à la résolution des multimères.

D'après Minh *et al*(2009), la façon dont PepA se lie aux séquences *cer* peut faire de ce complexe ADN/ protéines un mauvais substrat pour la recombinaison XerCD. Le surenroulement induit par le complexe serait une condition supplémentaire à la résolution des dimères de ColE1 (avec permission de reproduction, Minh *et al*, 2009).

On sait également que si PepA se trouve fortement en excès, les protéines peuvent se lier au site *cer* sans intervention d'ArgR. Ces conditions sont souvent présentes seulement *in vitro*.

Si plusieurs modèles existent pour décrire l'interaction entre les protéines ArgR et PepA, il reste à connaître les résidus permettant au complexe ArgR-PepA de se former. Si certains résidus ont déjà été proposés pour la protéine PepA, certains travaux semblent les infirmer (Reijns *et al*, 2005).

### I.13. Projet de thèse

Des travaux effectués dans notre laboratoire montrent que certains résidus de la partie Ct de la protéine ArgR seraient impliqués dans cette interaction entre ArgR et PepA. Les résidus et régions importants de PepA pour cette interaction protéine à protéine restent encore à bien décrire. On s'efforcera au cours de cette thèse d'élucider certaines de ces questions, et donc de connaître certains des résidus d'ArgR en Ct impliqués dans cette interaction entre les deux hexamères.

L'une des questions principales qui n'a pas encore de réponse claire au sujet de l'interaction ArgR-PepA est la suivante : où se fait l'interaction entre les hexamères d'ArgR et les hexamères de PepA ?

En effet, quant on regarde les dix dernières années, beaucoup d'expériences et résultats suggèrent qu'il y a un complexe ADN-protéine entre les séquences *cer* et les protéines ArgR et PepA, mais aussi un contact entre ArgR et PepA pour permettre la résolution des dimères plasmidiques avec formation d'un complexe stable *in vivo*. Ce contact se ferait au niveau des parties C-terminale de chaque hexamère, et ce de manière spécifique entre ArgR et PepA. De plus, les deux sites *cer* présents à cause de la dimérisation du plasmide sont toujours alignés parallèlement l'un par rapport à l'autre, soutenant donc une association stable entre les hexamères d'ArgR et de PepA.

Durant mon projet de thèse, je m'efforcerais donc de répondre aux différents points suivants, savoir s'il y a une association directe entre ArgR et PepA et caractériser les régions d'ArgR impliquées dans cette interaction, et d'en connaître les résidus d'acides-aminés situés en Ct si possible. On va donc chercher à étudier et identifier les résidus d'acides-aminés impliqués dans la formation du complexe des protéines accessoires ArgR et PepA. De précédentes expériences menées dans notre laboratoire ont montré que des mutants de la protéine ArgR, générés soit par mutagenèse dirigée, soit par insertion du transposon Tn4430 dans la sixième hélice  $\alpha$ , respectivement les mutants ArgR149 et ArgR5aa, étaient incapables de permettre une recombinaison

correcte. Ces deux mutants avaient justement subi des mutations au niveau de la région Ct de la protéine. Le mutant ArgR149 présente la séquence AIL en Ct, et le mutant ArgR5aa présente la séquence AILGVPLLELFDQEL en Ct, la séquence en Ct d'ArgRWt étant AILELFDQEL. De plus, les protéines ArgR149 et ArgR5aa, respectivement générées par les plasmides pHSArgR149 et pHSArgR5aa dans des études précédant cette thèse, sont toujours capables de se lier à la séquence Arg-box (interaction visualisée grâce au dosage  $\beta$ -galactosidase, à la recombinaison *in vivo* dans des cellules DS956 pourvues du plasmide pCS210 et par des retards sur gel) (Sénéchal et al, 2010).

Un deuxième point à élucider : montrer quels mutants d'ArgR C-terminaux ne peuvent interagir correctement avec PepA.

Un troisième et dernier point est de tenter de cerner quelle région de PepA est impliquée dans la formation du complexe ArgR-PepA lors de la résolution des dimères plasmidiques.

**CHAPITRE II**  
**MATÉRIELS ET MÉTHODES**



## CHAPITRE II. MATÉRIEL ET MÉTHODES

### II.1. Le matériel

#### II.1.1.Plasmides

Plasmide	Taille	Caractéristiques	Résistance	Référence
pDB169	5,2 kbs	vecteur de surexpression, promoteur <i>tac</i> , produit la protéine ArgRWt de 17 kDa	Ap <sup>R</sup>	Lim <i>et al</i> 1987
pDBArgR149	4,7 kbs	vecteur de surexpression, promoteur <i>tac</i> , produit la protéine ArgR149 de 17 kDa	Ap <sup>R</sup>	Ce travail
pDBArgR5aa	4,7 kbs	vecteur de surexpression, promoteur <i>tac</i> , produit la protéine ArgR5aa	Ap <sup>R</sup>	Ce travail
pGS38	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgRWt	Ap <sup>R</sup>	Stirling <i>et al</i> 1988b
pHSArgR149	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgR149	Ap <sup>R</sup>	Sénéchal <i>et al</i> 2010
pHSArgR5aa	3,8 kbs	promoteur <i>lac</i> produit la protéine ArgR5aa	Ap <sup>R</sup>	Sénéchal <i>et al</i> 2010
pHSArgR150	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgR150	Ap <sup>R</sup>	Ce travail
pHSArgR151	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgR151	Ap <sup>R</sup>	Ce travail
pHSArgRE150A	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgRE150A	Ap <sup>R</sup>	Ce travail
pHSArgRL149A	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgRL149A	Ap <sup>R</sup>	Ce travail
pHSArgRL151A	3,8 kbs	promoteur <i>lac</i> , produit la protéine ArgRL151A	Ap <sup>R</sup>	Ce travail

**Tableau 5.** Liste des plasmides produisant les protéines sans marquage. Il est à remarquer que le plasmide pGS38 est un dérivé du pUC19. Tous les plasmides intitulés pHS sont dérivés du pGS38.

Plasmide	Taille	Caractéristiques	Résistance	Référence
pCS126	4,15 kbs	vecteur de surexpression, promoteur <i>tac</i> , produit la protéine PepA de 55 kDa	Ap <sup>R</sup>	Stirling <i>et al</i> 1989
pNCK1	4,5 kbs	possède les séquences <i>cer</i> et <i>nmr</i> , séparées par le gène de résistance à la kanamycine	Ap <sup>R</sup>	Zakova <i>et al</i> 1995
pCD34	3,3 kbs	possède deux séquences <i>cer</i> , utilisées pour les amplifications de <i>cer</i> par PCR	Ap <sup>R</sup>	Communication Dr SZATMARI
pCS210	7,6 kbs	possède le gène <i>lacZ</i> flanqué de deux séquences <i>cer</i>	Tet <sup>R</sup>	Stirling <i>et al</i> 1989

**Tableau 6.** Liste des plasmides n'étant pas dérivés de pDB169 et de pGS38 et produisant des protéines sans marquage.

Plasmide	Taille	Caractéristiques	Résistance	Référence
pQE31ArgRWt	3,9 kbs	vecteur de surexpression, produit la protéine ArgRWt, fusionnée à une queue 6 HIS	Ap <sup>R</sup>	Ce travail
pQE31ArgR149	3,9 kbs	vecteur de surexpression, produit la protéine ArgR149, fusionnée à une queue 6 HIS	Ap <sup>R</sup>	Ce travail
pQE31ArgR5aa	3,9 kbs	vecteur de surexpression, produit la protéine ArgR5aa, fusionnée à une queue 6 HIS	Ap <sup>R</sup>	Ce travail
pQE31ArgRL149A	3,8 kbs	vecteur de surexpression, produit la protéine ArgRL149A, fusionnée à une queue 6 HIS.	Ap <sup>R</sup>	Ce travail
pQE31ArgRL151A	3,8 kbs	vecteur de surexpression, produit la protéine ArgRL151A fusionnée, à une queue 6 HIS.	Ap <sup>R</sup>	Ce travail

**Tableau 7.** Liste des plasmides produisant les protéines fusionnée à une queue 6-HIS.

pMAL-C2PepANt	7,2 kbs	vecteur de surexpression, produit la protéine PepANt-MBP	Ap <sup>R</sup>	Ce travail
pMAL-C2PepACt	7,7 kbs	vecteur de surexpression, produit la protéine PepACt-MBP	Ap <sup>R</sup>	Ce travail

**Tableau 8.** Liste des plasmides produisant les protéines fusionnées au MBP.

### II.1.2. Souches d'*Escherichia coli* utilisées

Durant les expériences menées dans le cadre de cette thèse, on a utilisé de nombreuses souches. Toutes ces souches possèdent des génotypes différents, bien que toutes soient des *Escherichia coli*, de type K12. Certaines de ces souches sont déficientes en ArgR, d'autres en PepA, certaines sont même déficientes pour les deux protéines.

Ces souches déficientes sont généralement utilisées pour la surexpression des protéines en vue de leur purification.

Les différentes souches utilisées lors de mes expériences sont présentées dans le tableau suivant.

Souches	Caractéristiques	Référence
DH5 $\alpha$	<i>F' <math>\phi</math>80dlacZ<math>\Delta</math>M15 <math>\Delta</math>(lacZYA-argF) U169 endA1 recA1 hsdR17(r<sup>-</sup><sub>K12</sub>m<sup>+</sup><sub>K12</sub>) deoR thi-1 supE44<math>\lambda</math> gyrA96 relA1</i>	Grant <i>et al</i> 1998
DS955	<i>argR :: Tp<sup>r</sup> pepA :: Tn5 redF lacZ<math>\Delta</math>M15 argE3<math>\Delta</math>(gpt-proA)62 his-4 leuB6 thr-1 thi-1 ara-14 lacY1 galK2 mtl-1 xyl-5 kdg K51 suppE44 rpsL31 tsx-1</i>	Flinn <i>et al</i> 1989
DS956	<i>argR :: Tp<sup>r</sup> redF lacZ<math>\Delta</math>M15 argE3<math>\Delta</math>(gpt-proA)62 his-4 leuB6 thr-1 thi-1 ara-14 lacY1 galK2 mtl-1 xyl-5 kdg K51 suppE44 rpsL31 tsx-1</i>	Flinn <i>et al</i> 1989
DS957	<i>pepA :: Tn5 redF lacZ<math>\Delta</math>M15 argE3<math>\Delta</math>(gpt-proA)62 his-4 leuB6 thr-1 thi-1 ara-14 lacY1 galK2 mtl-1 xyl-5 kdg K51 suppE44 rpsL31 tsx-1</i>	Stirling <i>et al</i> 1988
EC146 $\lambda$ ( $\lambda$ AZ-7)	<i>argD argR argA :: lacZ</i>	Eckhardt 1980
T7 Express	<i>fhuA2 lacZ :: T7 gene1[lon] ompT gal sulA11 R(mcr-73 :: miniTn10—Tet<sup>S</sup>)2[dcm] R(zgb-210 :: Tn10—Tet<sup>S</sup>)endA1 <math>\Delta</math>(mcr-mrr)114 :: IS10</i>	NEB
Turbo	<i>F'proA<sup>+</sup>B<sup>+</sup><math>\lambda</math><math>\alpha</math><math>\chi</math>Z<math>\Delta</math>M15/fhuA2 <math>\Delta</math>(lac-proAB) glnV zgb-210 :: Tn10(Tet<sup>R</sup>) endA1 thi-1 <math>\Delta</math>(hsdS-mcrB)5</i>	NEB

**Tableau 9.** Liste des souches de cellules utilisées dans les expériences.

Lors des expériences impliquant le système de recombinaison reproduit *in vitro*, les protéines XerC et XerD utilisées provenaient de la bactérie *Proteus mirabilis*. De plus l'une a été fusionnée avec 6-HIS, l'autre avec MBP (Maltose Binding Protein, dont le poids moléculaire est de 40 kDa). Le tableau suivant montre les plasmides utilisés pour générer ces protéines.

Protéine	Plasmide vecteur	Souche bactérienne	fusion
XerCPm	pMAL-C2	T7 Express	MBP
XerDPm	pMVD72	T7 Express	6-HIS

**Tableau 10.** Liste des protéines Xer utilisées. On décrit ici, les plasmides et souches utilisées pour produire les protéines XerC et XerD (Villion and Szatmari, 2003).

### **II.1.3. Milieu de culture Luria-Bertani (LB)**

Le milieu LB que nous utilisons a une formule pour 1 litre qui est la suivante : Tryptone 10g, extrait de levure 5g et chlorure de sodium 5g.

Toutes ces solutions sont autoclavées pour stérilisation. Les géloses préparées avec l'agar 1,5 % sont coulées dans des boîtes de pétri avec le ou les antibiotiques de sélections rajouté(s) dans la gélose non-polymérisée si nécessaire aux concentrations finales appropriées.

### **II.1.4. Liste des antibiotiques les plus utilisés**

Ampicilline, la solution stock est concentrée à 100 mg/mL et la concentration d'utilisation est de 100 µg/mL.

Tétracycline, la solution stock est concentrée à 12,5 mg/mL et diluée dans 50% d'éthanol. La concentration d'utilisation est de 6,25 µg/mL.

### **II.1.5. Géloses TAXI**

Les géloses TAXI sont des géloses LB avec agar dans lesquelles on met de la Tétracycline (T), de l'ampicilline (A), de l'X-gal (X) et de l'IPTG (I). Les concentrations des antibiotiques sont celles d'utilisation précédemment citée.

La concentration d'X-Gal est de 40 µg/mL. La concentration d'IPTG est de 1 mM au final dans le milieu.

### **II.1.6.Milieu MacConkey Agar**

Sa formulation approximative pour un litre est la suivante : Peptone 17g, Peptone de protéase 3g, lactose 10g, sels biliaires n°3 1,5g, chlorure de sodium 5g, gélose 13,5g, rouge neutre 0,03g, cristal violet 0,001g.

Pour préparer des gélules, on dissout 50 g de poudre (composée des produits précédemment cités) pour 1 L. On chauffe ensuite le mélange en l'agitant de temps en temps, puis on la fait bouillir une minute afin de dissoudre la poudre. Puis, on autoclave 20 min à 121°C pour stériliser le milieu. Ensuite si nécessaire, on coulera les gélules avec un antibiotique de sélection.

### **II.1.7.Fabrication de milieu auto-inductible (Studier, 2005)**

La synthèse des protéines est induite par un milieu auto-inductible liquide contenant, pour un volume total de 100 mL : 4,76 Grammes de Terrific Broth seront dissout dans 100 mL d'eau distillée.

La formule du Terrific Broth pour un litre est la suivante : Bacto Tryptone 12 g, extrait de levure Bacto 24 g, phosphate de potassium dibasique 9,4g, phosphate de potassium monobasique 2,2 g. Le reste de la solution auto-inductible est composée de 0,05% de glucose, 0,2% de lactose, 0,4% de glycérol et 2 mM de MgSO<sub>4</sub>.

### **II.1.8.Fabrication du tampon pour cellules compétentes**

Il s'agit d'un tampon calcium-HEPES d'un volume total de 100 mL, que l'on garde à 4°C et qu'on utilise froid lorsque l'on rend les *Escherichia coli* compétentes à la transformation. Il est composé de 60 mM de CaCl<sub>2</sub>, 15 % de glycérol et 10 mM de HEPES pH 7.

### **II.1.9. Tampon de liaison pour retard sur gel**

Le tampon de liaison ADN-protéine est composé de la manière suivante : pour une concentration stock de 4X, 80 mM Tris-HCl pH 7,5, 1 mM EDTA, 200 mM NaCl, 20% glycérol et une concentration de L-arginine d'au moins 4 mM. Lors de la migration, la concentration finale de tampon sera de 1X. C'est cette composition qui a été utilisée pour la migration des séquences *cer*. Pour la migration de la boîte arginine, 10mg/mL de BSA ont été rajoutés à cette même solution. De plus, 2 ng d'ADN et 1 µg de poly dIdC ont été rajoutés dans chaque échantillon avant migration.

### **II.1.10. Tampons TM**

Le tampon TM généralement utilisé dans les dialyses est composé de la manière suivante, Tris-MgCl<sub>2</sub>, 10 mM Tris-HCl pH=7,5, 10 mM MgCl<sub>2</sub>. Il existe aussi un tampon TM à haute concentration en sel pour resuspendre les protéines ArgR et ses différentes formes mutantes. Sa composition est la suivante, 20 mM Tris-HCl pH=7,5, 20 mM MgCl<sub>2</sub>, 200 mM NaCl et 50 % de glycérol (il est appelé TM spécial).

### **II.1.11. Tampons pour la purification de PepA employée pour les expériences de topologie**

Après induction de la protéine PepA, le tampon de lavage se compose de 10 mM Tris-HCl pH 7,5, 10 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol. Le premier tampon de resuspension est de la composition suivante, 50 mM Tris-HCl pH 8,2, 1 M NaCl, 10 mM MgCl<sub>2</sub>, 0,1 mM EDTA, 0,1 mM DTT avant lyse par sonication. Le second tampon de resuspension est composé de 20 mM Tris-HCl à pH 8,2, 200 mM KCl, 1 mM MgOAc, 0,1 mM EDTA. Le tampon de dialyse contient 20 mM Tris-HCl à pH 8,2, 10 mM KCl, 1 mM MgOAc, 0,1 mM EDTA. Le troisième tampon de resuspension

quant à lui, contient 20 mM Tris-HCl à pH 8,2, 200 mM KCl, 1 mM MgOAc, 0,1 mM EDTA, 50% glycérol.

#### **II.1.12. Tampon pour le *crosslinking* au glutaraldéhyde**

Le tampon contient 50 mM de triéthanolamine pH 8,0, 150 mM de KCl. un pourcentage de glutaraldéhyde qui va de 0,005% à 0,08% a été rajouté ensuite dans les solutions de tampon (les pourcentages suivants, 0,005%, 0,01%, 0,02%, 0,04% et 0,08% ont été calculés).

#### **II.1.13. Tampon pour les migrations SDS-PAGE**

Ce tampon Tris-glycine-SDS est composé de 980 mM glycine, 250 mM, 0,5% SDS, pour un stock de 5X.

#### **II.1.14. Tampons pour les protéines de fusion 6-HIS**

Ces tampons proviennent du kit Qiagen tm, (*Qiaexpressionist handbook*). Il y a tout d'abord, le tampon de lyse, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl et 10 mM imidazole, pH 8, puis le tampon de lavage, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl et 20 mM imidazole, pH 8. Normalement, le Tampon d'élution est de composition suivante, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl et 250 mM imidazole, pH 8. Mais pour les éluions des protéines ArgR, le taux d'imidazole a été modifié plusieurs fois comme expliqué plus bas, pour éliminer des impuretés/protéines liées non spécifiquement à la résine.

Lors des purifications des différentes protéines ArgR mutantes ou sauvages, on a utilisé trois concentrations différentes d'imidazole, 100 mM d'imidazole, 120 mM



d'imidazole et 150 mM d'imidazole, les protéines ArgR sortant à 150 mM d'imidazole. Chaque élution a eu lieu 4 fois.

#### **II.1.15. Tampons pour la purification des protéines fusionnée au MBP**

Tampon pour colonne (200 mM NaCl, 1 mM EDTA, 20 mM Tris pH 7,5, 300 microlitre de PMSF à une concentration initiale de 10 mg/mL). Tampon de lavage n°1 (Tris EDTA 10 mM pH 8,0, NaCl 1M). Tampon d'élution (tampon pour colonne avec 10 mM de maltose), on récupère 1 mL d'élution à la fois.

#### **II.1.16. Tampons pour les expériences utilisant les billes magnétiques à amylose**

On utilise les tampons pour colonnes utilisés pour les expériences des protéines fusionnées au MBP. Lors de l'étape d'élution, on utilise le même tampon que le tampon d'élution pour les protéines fusionnées au MBP.

#### **II.1.17. Tampons pour les expériences de *crosslinking* réversibles au DSP**

On prépare une solution de 25 mg/mL de DSP, pour Dithiobis (Succinimidyl Propionate), dilué dans du DMSO pour une concentration stock de 62 mM de DSP, qui doit être utilisé le jour même (la solution doit rester « fraîche »). Lors des expériences de *crosslinking* réversibles, on doit avoir 5 mM de DSP au final dans chaque mélange de protéines à « crosslinker » au DSP. Lors de la réversion du *crosslinking*, on ajoute du DTT et du  $\beta$ -mercaptoéthanol pour séparer les protéines liées par le DSP au moment de la dénaturation par chauffage à 95°C. Lors de ces expériences, les tampons suivants ont été utilisés après le *crosslinking* pour lier les protéines ArgR (toutes versions testées), et PepA. Tampon d'interaction, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM

d'imidazole, 0,005% Tween 20, 1 mM L-arginine, pH= 8,00. Tampon de lavage, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM d'imidazole, 0,005% Tween 20, 1 mM L-arginine, pH= 8,00. Tampon d'élution 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM d'imidazole, 0,005% Tween 20, 1 mM L-arginine, pH= 8,00. On remarquera que le tampon d'interaction a une composition identique à celle du tampon de lavage.

#### **II.1.18. Tampons pour purifier les protéines fusionnées à une queue 6-HIS avec des billes magnétiques Ni-NTA**

Après les expériences au DSP, on utilise un tampon d'interaction pour laver les billes magnétiques et conserver les protéines liées par le DSP, dont l'une est fusionnée avec 6-HIS. La composition de ce dernier est, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl et 20 mM imidazole, pH 8, avec 0,005% de tween 80 et si on travaille avec ArgR, on rajoute 1 mM de L-arginine. Le tampon d'élution est quant à lui composé de 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl et 250 mM imidazole, pH 8, 0,005% de tween, et sert à séparer les protéines des billes magnétiques. La plupart des expériences utilisant le DSP se font avec ArgR. En général, on ajoute de la L-arginine pour qu'ArgR soit active.

#### **II.1.19. Tampons de recombinaison *in vitro***

Le tampon de recombinaison *in vitro* possède une concentration finale de 50 mM Tris-HCl pH 8,0, 50 mM NaCl, 1,25 mM EDTA, 5 mM spermidine, 1 mM L-arginine, 10 % glycérol et 25 microgrammes/mL BSA. On fabrique un tampon 5X pour pouvoir diluer cette solution durant les expériences jusqu'à 1X.

### II.1.20.Solutions de Southern-blot

Il y a tout d'abord une solution de dépurination composée de 10,75 mL d'HCl à 11,6 M et de 489,25 mL d'eau distillée pour une concentration finale de 0,25 M. Puis une solution de dénaturation, 43,82 g de NaCl, 10 g de NaOH, compléter à 500 mL avec de l'eau distillée, soit 1,5 M de NaCl et 0,5 M de NaOH. Dans une solution de neutralisation contenant 30,28 g de Tris (0,5 M), 43,83 g de NaCl (0,5 M), mettre à pH 7 et compléter à 500 mL. Il y a également préparation d'une solution de préhybridation (25 mL de 20XSSC, 10 mL de blocking reagent ROCHE tm à 10%, 1 mL de N-Lauryl-sarcosine à 10%, 200 microlitres de SDS à 10% et 63,8 mL d'eau stérile) que l'on préchauffera à 65°C. Après les hybridations, on utilise des solutions de lavages fabriquées à partir de 20XSSC (pour 1L, 175,30 g de NaCl (3 M), 88,2 de citrate de sodium (0,342 M), pH=7). On fabrique les solutions suivantes, 2XSSC-0,2% de SDS (50 mL de 20XSSC, 5 mL de SDS à 20%, ajusté à 500 mL avec H<sub>2</sub>O distillée), 1XSSC-0,1% de SDS (25 mL de 20XSSC, 5 mL de SDS à 20% ajusté à 500 mL avec H<sub>2</sub>O distillée). On utilise les tampons suivants pour visualiser les bandes d'ADN après la migration. Tampon 1 ; pour un volume total de 500 mL, on a 100 mM d'acide maléique, 150 mM de NaCl, pH 7,5, on autoclave et garde la solution à 4°C. Tampon 2 ; pour un volume total de 500 mL, on rajoute 5 g de blocking reagent (Roche Diagnostique, Inc). On autoclave et garde la solution à 4°C. Tampon 3 ; pour un volume total de 500 mL, on a 100 mM de Tris-HCl, 100 mM de NaCl et 50 mM de MgCl<sub>2</sub>, pH 9,5. On filtre et garde la solution à 4°C. Pour réagir avec les anticorps antidig, on utilise une solution composée de 1,4 mL d'H<sub>2</sub>O distillée, 100 microlitres de diluant, 5 µL de CDP-star substrate de NEB tm.

### II.1.21.Solution pour la Topoisomérase 1

C'est une solution contenant au final 0,5 mM de dithiothreitol (DTT), 0,1 mM EDTA, 1 mM L-arginine, 10 mM MgCl<sub>2</sub>, 50 mM KCl, 50 mM Tris-HCl, pH 7,5. On fabrique généralement des solutions concentrées à 5X.

### II.1.22. Protéines ArgR et PepA utilisées

Lors de mes expériences j'ai utilisé de nombreuses versions des protéines ArgR et PepA, les versions sauvages ArgRWt et PepAWt, mais également des versions mutantes générées de différentes façons. En général, les techniques employées pour générer ces versions mutantes d'ArgR et PepA sont de la mutagenèse par transposon Tn4430 et la mutagenèse dirigée par PCR.

Ces résultats sont présentés sous forme d'un tableau récapitulatif montrant les différentes modifications apportées aux protéines générées au cours des expériences.

Protéines	Transposon Tn4430	Mutagenèse par PCR	Fusion Nt
ArgRWt	-	-	6 HIS
ArgR149	-	+	6 HIS
ArgR5aa	+	-	6 HIS
ArgRE150A	-	+	-
ArgR150	-	+	-
ArgR151	-	+	-
ArgRL149A	-	+	6 HIS
ArgRL151A	-	+	6 HIS
PepANtMBP	-	+	MBP
PepACtMBP	-	+	MBP

**Tableau 11.** Présentation des différentes façons d'obtenir les différents mutants pour ArgR et PepA. Les éventuelles fusions en position Nt ont été transcrites ici. Il est important de signaler que les protéines ArgRWt, ArgR149, ArgR5aa ont d'abord été générées sous leurs formes natives sans fusion.

### **II.1.23. Tampon de vérification de l'activité aminopeptidase A**

Le mélange que l'on utilise est une solution contenant de la L-leucine p-nitroamilide hydrochloride (LPNA). Pour un volume total de 1 mL, on met 200  $\mu$ L de TM spécial de concentration stock de 1 M pour le NaCl, 100  $\mu$ L de  $MnCl_2$  à 0,01M final, 100  $\mu$ L de LPNA à 10 mM de concentration stock et on complète avec de l'eau distillée (600  $\mu$ L). On rajoute 15  $\mu$ g de PepAWt ou de PepANtMBP ou de PepACtMBP en solution.

### **II.1.24. Tampon de charge pour les gels d'ADN**

La composition du tampon est la suivante, 0,25 % de xylène-cyanol, 0,25 % de bleu de bromophénol, 30 % de glycérol. Il s'agit d'une concentration de 5 X.

### **II.1.25. Tampon de charge pour les gels de protéines**

La composition de ce tampon est la suivante, Tris 0,4 M, pH=6,8, DTT 7,7 %, SDS 10 %, glycérol 50 %. On met un peu de poudre de bleu de bromophénol pour rendre la migration visible.

### **II.1.26. Tampon Tris-borate-EDTA**

Pour un volume de 500 mL, on a 54 g Tris, 27,5 g d'acide borique, 20 mL d'EDTA 0,5 M, pH=8,0. Ce mélange est appelé tampon TBE 1X. Garder au froid à 4°C. Lors des expériences de retard sur gel, la concentration de ce tampon sera de 0,25 X pour les boîtes arginine et de 0,5 X pour les séquences *cer*.

## II.2. Les méthodes

### II.2.1. Transformation chimique des bactéries

Des cellules compétentes ont d'abord été fabriquées. Pour cela, 5 mL d'un milieu LB liquide ont été inoculés avec une colonie venant d'une gélose LB-Agar, laisser croître toute la nuit. 1,5 mL de la pré-culture précédente ont été mis dans 150 mL de milieu de culture LB liquide. La croissance des bactéries a été laissée jusqu'à une Densité Optique (DO) de 0,375 à 595 nm, puis laisser reposer 10 minutes sur glace. Centrifuger 7 minutes à 5000 rotations par minute (RPM), puis resuspendre dans 10 mL de tampon calcium-HEPES froid et laisser 10 minutes sur glace avant de centrifuger à nouveau 7 minutes à 5000 RPM. Resuspendre à nouveau dans 2 mL de tampon calcium froid puis aliquoter 100 microlitres par eppendorf de 1,5 mL et congeler à -70°C.

La transformation par choc thermique suit ce protocole qui rend les cellules d'*Escherichia coli* compétentes. Laisser d'abord décongeler un tube de cellules compétentes sur la glace puis ajouter jusqu'à 10 microlitres d'ADN (2 µL pour un vecteur entier, et 5 à 10 microlitres de produits de ligation). Laisser 10 minutes sur glace et mettre 35 secondes à 42°C avant de rajouter 1 mL de LB froid. On laisse ensuite incuber une heure à 37°C avec agitation avant d'étaler sur une boîte de pétri pourvue de l'antibiotique de sélection approprié. Incuber toute la nuit pour voir si des colonies transformées se forment.

### II.2.2. Expériences *in vivo* recombinaison pCS210

Les expériences de recombinaisons *in vivo* se font dans les bactéries *Escherichia coli* de la souche DS956 déficiente en ArgR.

Ces souches sont compétentes pour une transformation par choc thermique de 30 secondes, et déjà transformées par le plasmide pCS210, contenant le gène *lacZ* flanqué de deux séquences *cer*.

S'il y a recombinaison du plasmide pCS210, on a perte du gène *lacZ* et il ne reste plus qu'une séquence *cer*.

Lorsqu'il y a recombinaison, on passe d'un plasmide pCS210 de 7,6 Kbs à un plasmide pCS211 de 4,6 Kbs. Cette délétion est visualisée par la formation de colonies blanches sur milieu TAXI, ou par électrophorèse des plasmides extraits des cellules transformées (après digestion avec HindIII, enzyme de restriction qui coupe le plasmide une seule fois).

### II.2.3. Expériences *in vivo* répression *argA::lacZ*

Ces expériences ont lieu après transformation, par choc thermique de 30 secondes, de bactéries compétentes de la souche EC146λ. Ces bactéries comportent une construction chromosomique avec le gène *lacZ* contrôlé en amont par une boîte arginine. Ces expériences visent à vérifier si les versions sauvages et mutantes d'ArgR sont toujours capables de se lier à la séquence d'ADN nommée boîte arginine, dont la séquence consensus est :

5' TNTGAATTTAAATTCANATAATNTGAATTTAAATTCANA 3'.

Mais la séquence que l'on a utilisée, est celle de la boîte arginine du gène *argA*

5' ACAGAATAAAAATACACTAATTTCGAATAATCATGCAAA 3'. (Maas, 1994)

Après cette transformation, on étale les bactéries sur gélose LB+agar+X-gal ou sur gélose Mac Conkey, pour vérifier la coloration des colonies qui indique s'il y a répression ou non de la fusion *argA::lacZ*.

#### **II.2.4. Préparation d'un gel d'agarose 1 %**

Dissoudre 1g d'agarose dans 100 mL de TAE 1X (40 mM Tris-Acétate, 1 mM EDTA, pH = 8,00) par chauffage en portant le mélange à ébullition, dans un récipient en verre. Faire ensuite couler le gel dans un bloc généralement pourvu de peigne et laisser refroidir le tout, ce qui va entraîner la polymérisation du gel.

#### **II.2.5. Préparation d'un mini-gel SDS-PAGE**

Deux parties le composent, le Resolving gel de composition suivante : 3,7 mL d'une solution de polyacrylamide (40% acrylamide, 0,8% bis-acrylamide), 3,8 mL de Tris-HCl 1M pH 8,8, 100 microlitres de SDS 10%, 100 microlitres d'APS (0,1 g/mL), 2 mL d'eau distillée, 20 microlitres de TEMED. Ce qui fait 15% d'acrylamide au final dans le Resolving gel.

Au-dessus de cette partie, on a le Stacking gel : 1,9 mL d'eau distillée, 325 microlitres d'une solution de polyacrylamide (33% acrylamide, 0,8% bis-acrylamide), 312,5 microlitres de Tris-HCl 1M pH 6,8, 25 microlitres d'APS (0,1 g/mL), 25 microlitres SDS 10%, 10 microlitres de TEMED. Ce qui fait 12,8% d'acrylamide au final dans le Stacking gel.

#### **II.2.6. Préparation d'un gel pour une expérience de retard sur gel**

Les gels utilisés pour ces expériences sont des gels de polyacrylamide 6%, avec une concentration en L-arginine de 0,1 mM au minimum. La composition du gel est la suivante : 1 mL de TBE 10X, 2 mL d'une solution d'acrylamide/bis-acrylamide 30%/0,75%, 50 mL de L-arginine 0,1M, 7mL H<sub>2</sub>O. Ajouter le tampon de migration électrophorétique, le TBE, contenant la L-arginine. Généralement, la concentration du TBE est de 1X (89 mM de Tris, 89 mM d'acide borique, 2 mM d'EDTA), mais des



concentrations de 0,5X et de 0,25X ont été plus fréquemment utilisées. Dans tous les cas, 1 mM de L-arginine a été mis dans ces solutions.

### **II.2.7. Induction de protéines par IPTG**

Préparer une préculture *overnight* de 5 mL de nos bactéries contenant un plasmide capable de générer la protéine ArgR ou la protéine PepA. Après cette culture, on met 4 mL de cette préculture dans 250 mL de LB avec antibiotique de sélection et on attend d'atteindre une D.O de 0,6 à 600 nm. Prendre un échantillon de 1 mL (pour éventuellement tester les protéines dans cet échantillon avec un gel SDS-PAGE) avant d'ajouter de l'IPTG pour une concentration finale de 1 mM à partir d'un stock de 100 mM d'IPTG. Induire ensuite pendant 4 heures (pour les plasmides avec un promoteur *tac* fort et 16 heures pour les plasmides dérivés du pGS38). Répartir les 250 mL de cultures induites entre les deux bouteilles et centrifuger à 4°C pendant 20 min à 5000 rpm. On utilise ensuite les protocoles pour une purification partielle.

### **II.2.8. Purification partielle de la protéine ArgR**

Cette expérience nécessite d'abord la surexpression de la protéine par le biais de souches d'*E.coli* ayant été transformées par un vecteur de surexpression, Les tampons utilisés ensuite pour la lyse des cellules dépendent des types de purifications.

Resuspendre les cellules bactériennes centrifugées 20 minutes à 4°C, dans 5 mL d'un mélange de TM et de NaCl (concentré à 1M au final), on rajoute 0,5 mM de PMSF (dilué dans de l'éthanol à 95%), pour empêcher l'action des protéases. 3 à 5 cycles gel-dégel sont effectués pour fragiliser les membranes cellulaires, en mettant les cellules resuspendues à -70°C puis à 37°C. Une fois les cycles terminés, environ 20 µL d'une solution de lysozyme à une concentration stock de 1 mg/mL est ajouté et on laisse agir 10 minutes avant de faire la sonication (10 cycles avec une minute de repos entre les cycles). Ensuite une centrifugation à 10000 rotations par minute (RPM) à 4°C durant 20

minutes a été faite, on récupère le surnageant et on y ajoute du sulfate d'ammonium (0,291 g par mL) que l'on laisse pendant 2 heures dans un endroit à 4°C et sous agitation, pour précipiter les protéines. Centrifuger ensuite pendant 20 minutes à 10000 rpm à 4°C. Pendant ce temps, on stérilise les boudins de dialyse en les faisant bouillir 15 min dans 500 mL d'H<sub>2</sub>O distillée contenant 1 mL d'EDTA à 0,5M et pH 8,0. Une fois la centrifugation terminée, le surnageant a été conservé, mais surtout, le culot a été resuspendu dans 2 à 5 mL de TM. Il est également possible de dialyser le surnageant. Dialyser au moins 2 fois pendant 2 heures pour la première dialyse, et toute la nuit pour la seconde dialyse.

Si la purification d'ArgR est recherchée, la première centrifugation est la même, mais une étape de précipitation à la L-arginine de ArgR est rajouté dans le surnageant conservé. Laisser sur glace 15 à 30 min, ensuite, recentrifuger de la même manière le mélange. Récupérer le culot et resuspendre la protéine dans du TM à haute concentration en sel.

### **II.2.9. Purification partielle pour la protéine PepA**

Cette expérience nécessite d'abord la surexpression de PepA par le biais de souches d'*E.coli* ayant été transformées par un vecteur de surexpression, Les tampons utilisés ensuite pour la lyse des cellules dépendent des types de purifications. Resuspendre les cellules bactériennes centrifugées 20 minutes à 4°C, dans 5 mL d'un mélange de TM et de NaCl (concentré à 1M au final), on rajoute 0,5 mM de PMSF (dilué dans de l'éthanol à 95%), pour empêcher l'action des protéases. Faire ensuite 3 à 5 cycles gel-dégel pour fragiliser les membranes cellulaires, en mettant les cellules resuspendues à -70°C puis à 37°C. Une fois les cycles terminés, rajouter du lysozyme à une concentration de 1 mg/mL et laisser agir 10 minutes avant de faire la sonication (10 cycles avec une minute de repos entre les cycles). Effectuer ensuite une centrifugation à 10000 rotations par minute (RPM) à 4°C durant 20 minutes, récupérer le surnageant et y ajouter du sulfate d'ammonium (0,291 g par mL) et laisser pendant 2 heures dans un

endroit à 4°C et sous agitation, pour précipiter les protéines. Centrifuger ensuite pendant 20 minutes à 10000 rpm à 4°C. Pendant ce temps, stériliser les boudins de dialyse en les faisant bouillir 15 min dans 500 mL d'H<sub>2</sub>O distillée contenant 1 mL d'EDTA à 0,5M et pH 8,0. Une fois la centrifugation terminée, conserver le surnageant, mais surtout resuspendre le culot dans 2 à 5 mL de TM. Il est également possible de dialyser le surnageant. Dialyser au moins 2 fois pendant 2 heures pour la première dialyse, et toute la nuit pour la seconde dialyse, contre un tampon de dialyse (tampon TM de 10 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, pH = 7,5). Centrifuger à 4°C pendant 20 min à 10000 rpm les fractions de culots dialysés et resuspendre le culot ainsi obtenu dans du TM spécial.

### **II.2.10. Autre purification partielle pour PepA**

Le but de cette expérience est de générer des purifications de PepA avec moins d'activité de nucléase par rapport au protocole précédent.

Laver les cellules DS957 (qui sont normalement *pepA*<sup>-</sup>, mais qui ont été transformées avec un plasmide permettant son expression) avec le tampon suivant, 10 mM Tris-HCl pH 7,5, 10 mM MgCl<sub>2</sub>, 10 mM β-mercaptoethanol. Les cellules ont été resuspendues dans le tampon suivant, 50 mM Tris-HCl pH 8,2, 1 M NaCl, 10 mM MgCl<sub>2</sub>, 0,1 mM EDTA, 0,1 mM DTT et faire une lyse par sonication, après centrifugation à 10000 rotations par minute (RPM), 4°C durant 20 minutes, le surnageant a été récupéré et les protéines ont été chauffées à 70°C pendant 10 minutes, détruisant l'activité nucléase. Après cela, les protéines ont été précipitées avec du sulfate d'ammonium (0,291 g par mL) à 4°C pendant 1 à 2 heures. Centrifuger ensuite à 10000 rotations par minute (RPM), à 4°C durant 20 minutes. Récupérer le culot et le resuspendre dans le tampon suivant, 20 mM Tris-HCl à pH 8,2, 200 mM KCl, 1 mM MgOAc, 0,1 mM EDTA. Dialyser ensuite deux fois à 4°C dans un tampon contenant 20 mM Tris-HCl à pH 8,2, 10 mM KCl, 1 mM MgOAc, 0,1 mM EDTA, pendant 3 heures pour la première dialyse, et toute la nuit pour la seconde. Après centrifugation

10000 rotations par minute (RPM), 4°C durant 20 minutes de la solution de dialyse, resuspendre le culot dans un tampon 20 mM Tris-HCl à pH 8,2, 200 mM KCl, 1 mM MgOAc, 0,1 mM EDTA, 50% glycérol.

### **II.2.11. Purification des protéines fusionnées à une queue N-terminale 6-HIS**

Les différentes versions d'ArgR et la protéine XerD fusionnées ainsi sont purifiées via un kit Qiagen grâce à de l'imidazole qui décroche la protéine fusionnée de la colonne de Ni-NTA. Après avoir préalablement fait une culture des bactéries exprimant la protéine fusionnée avec 6-HIS, culoter les bactéries avant de les resuspendre dans du tampon de lyse, avec 2 à 5 mL de tampon par gramme de cellules. Rajouter 20 µL de lysozyme 1 mg/mL (soit une concentration finale de 20 µg/mL) et laisser incubé pendant 30 minutes avant de faire la sonication. Rajouter 1 mL de NaCl à 5 M pour avoir une concentration finale en NaCl de 1 M, avant sonication, pour solubiliser les protéines le plus possible. Après sonication, on centrifuge 20 minutes à 4°C, à 10000 RPM. Une fois que le surnageant a été récupéré, ajouter 1 mL de résine Ni-NTA pour 4 mL de surnageant. Laisser agiter 60 minutes à 4°C avant de placer tout le mélange dans une colonne. Puis récupérer le *flow-through*, avant de rajouter deux fois 4 mL de tampon de lavage. Après cette étape, on élue les protéines. Normalement, l'élution se fait quatre fois avec un tampon d'élution aux concentrations standardisées. Mais pour les protéines ArgR, on utilise trois concentrations en imidazole différentes pour avoir une meilleure pureté, c'est-à-dire, 100 mM, 150 mM et 250 mM d'imidazole. On vérifie les échantillons par migration gel SDS-PAGE.

### **II.2.12. Purification des protéines fusionnées au MBP**

Les protéines fusionnées au MBP (PepACT-MBP, PepANT-MBP et XerC) ont été purifiées sur une colonne contenant une résine amylose *high-flow* (NEB tm) préalablement lavée par 5 mL de tampon pour colonne. Pour ce faire, passer dans la

colonne le surnageant préalablement récupéré après la sonication. Après, faire un lavage de cette colonne grâce à 1 mL de tampon de lavage n°1. Ensuite, laver avec 8 mL de tampon pour colonne. Enfin, éluer la protéine fusionnée au MBP par 5 mL de tampon d'élution, et récupérer 1 mL d'élution à la fois dans un eppendorf. Ces éluions sont sensées contenir les protéines fusionnées au MBP, à des concentrations différentes selon les eppendorfs.

Les colonnes de résine peuvent être régénérées de la manière suivante. Pour cela, faire une première étape avec passage de 8 mL d'eau distillée suivie de 10 mL d'une solution contenant de l'eau distillé avec 0,1 % de SDS. Lors de la deuxième étape, faire passer dans la colonne 8 mL d'eau distillée puis 8 mL de tampon pour la colonne.

### **II.2.13. Protocole de récupération des protéines par billes magnétiques avec amylose**

Ce protocole est adapté d'un protocole standardisé de chez NEB.

Tout d'abord, vortexer et resuspendre les billes magnétiques, on va ensuite utiliser 100  $\mu$ L de billes magnétiques pour 200  $\mu$ L de volume total de solution avec les protéines. Mettre le volume approprié de billes magnétiques, puis laver deux fois les billes avec 500  $\mu$ L de tampon MBP pour colonne (on laisse une minute au séparateur magnétique entre chaque lavage). Mettre les billes dans le mix avec les protéines (dans nos expériences, ArgR, PepANt-MBP et PepACT-MBP). On incube ensuite à 4°C pendant une heure.

Séparer ensuite les billes du surnageant grâce au séparateur magnétique. Faire ensuite 3 lavages avec 500  $\mu$ L du tampon MBP pour colonne.

Une élution a été pratiquée en mettant 30  $\mu$ L de tampon d'élution et laisser ensuite 1 minute avant d'utiliser le séparateur magnétique (lui aussi pendant 1 minute) et récupérer le surnageant. Répéter une seconde fois cette élution selon les mêmes termes.

#### **II.2.14. Étude d'interaction protéine-ADN *in vitro***

Lors de ces expériences, la boîte arginine marquée avec la digoxygénine soit avec HEX, a été mise en présence de trois versions d'ArgR, ArgRWt, ArgR149, ArgR5aa).

Les retards sur gel ont été réalisés sur des gels de polyacrylamide 6%. La composition du gel est la suivante : 1 mL de TBE 10X, 2 mL d'une solution d'acrylamide/bis-acrylamide 30%/0,75%, 50 mL de L-arginine 0,1M, 7 mL H<sub>2</sub>O. Le TBE 10X est fabriqué de la manière suivante pour un volume total de 500 mL, 54 g de Tris (890 mM), 27,5g d'acide borique (890 mM), 20 mL d'une solution d'EDTA de 0,5 M à pH=8 (2 mM au final, soit 7, 445 g d'EDTA dans 500 mL). Avant de mettre les mélanges ADN/protéines, le gel a été polarisé par une pré-électrophorèse d'une heure à 100 V à température pièce.

Les mélanges sont réalisés comme suit : pour un volume total de 20 microlitres, 10 microlitres de tampon de liaison dilué à 2X, 1 microlitre de poly(dI-dC), jusqu'à 5 microlitres d'ADN (environs 2 ng par puits), 1 à 6 microlitres d'ArgR de concentration comprise entre 9 nM à 41400 nM (quelle que soit la version) et 1 mM de L-arginine. L'incubation dure 30 minutes à 37°C, puis les mélanges ont été laissés 20 minutes sur la glace.

C'est alors que 2 microlitres d'un tampon de charge de bleu de bromophénol 5X ont été rajoutés. Une fois les puits chargés, ils ont été laissés migrer à 100 Volts, à 4°C. Les formes suivantes des protéines ont été utilisées : ArgRWt, ArgR149 et ArgR5aa. Il

en important de noter que ces conditions de migration sont valables pour la migration des séquences boîtes arginine. Pour la séquence *cer*, la migration a été faite à 65 Volts et à température ambiante.

Le tampon de migration pour les retards sur gels est composé de 0,25 X TBE pour la migration à 100 Volts à 4°C pour les boîtes arginine. Le tampon de migration pour les retards sur gels est composé de 0,5 X TBE pour la migration à 100 Volts à 4°C pour les retards sur gel des complexes ArgR-PepA-ADN. Le tampon de migration pour les retards sur gels est composé de 0,5X TBE pour la migration à 65 Volts à température ambiante.

Les bandes d'ADN ont été détectées grâce au Typhoon (laser 532 nm, filtre 555BP20) pour les ADN marqués par HEX.

Pour les ADN marqués à la dioxygénine, le gel a été transféré sur une membrane de nylon chargée positivement (Hybond-N d'Amersham Biosciences). La détection finale a eu lieu après une minute de lavage avec 15 mL de tampon 1, puis 20 min de lavage avec 30 mL de tampon 2. Il faut ensuite laver 30 minutes avec 30 mL de tampon 2 contenant 6 microlitres d'anticorps anti-DIG. Dans un nouveau bac, laver deux fois de suite 15 minutes la membrane dans 30 mL de tampon 1. Un dernier lavage a été fait avec le tampon 3 pendant 2 minutes. La solution réagissant avec les Anticorps et le marquage a été préparé comme suit, 1,4 mL d'H<sub>2</sub>O distillée, 100 microlitres de diluant, 5 microlitres de CDP-star substrate de NEB. La membrane a été mise entre 2 acétates avant de mettre la solution ci-dessus. Après 5 minutes, retirer l'excédent et sceller les acétates. Placer ensuite la membrane sur un film et laisser exposer le film photo Fujifilm SuperRX X-ray pendant 5 à 20 minutes dans le noir avant de développer.

### II.2.15. *Crosslinking* au glutaraldéhyde

Cette expérience permet de savoir si les protéines ArgR de types sauvages et mutantes sont toujours capables de se multimériser, sachant que le glutaraldéhyde permet de former et de visualiser les multimères d'ArgR en créant des liaisons fortes, qui ne sont pas dénaturées par un gel SDS-PAGE.

Les protéines sont mélangées volume à volume avec le tampon triéthanolamine-glutaraldéhyde (composition dans la partie II.1.13.). On incube les mélanges à température de la pièce pendant 2 heures. Puis on fait migrer les échantillons sur le gel SDS-PAGE.

Le tampon de migration est composé de Tris-glycine-SDS 1X.

### II.2.16. *Crosslinking* au DSP

Ces expériences utilisent un *crosslinker* réversible, le DSP, pour Dithiobis (Succinimidyl Propionate), qui lie des protéines qui interagissent ensemble en temps normal, permettant ainsi une meilleure récupération d'un complexe, si l'une des protéines « *crosslinkée* » est fusionnée à une queue 6-HIS et récupérable par les billes magnétiques Ni-NTA. Lors de ces expériences la concentration finale de DSP (dilué dans le DMSO à 25 mg/mL dans sa solution stock) est de 62 mM, et la quantité de protéines ArgR est de 10 µg (588,24 picomoles) et celle de PepA est de 30 µg (542,5 picomoles). Durant chaque expérience, le volume final des mélanges est de 100 µL. le tampon d'interaction utilisé lors de ces expériences est celui utilisé pour les interactions avec les billes magnétiques, avec ajout de L-arginine (la composition du tampon est la suivante, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 20 mM d'imidazole, 0,005% Tween 20, 1 mM L-arginine, pH= 8,00). En général, on ajoute au mélange 4 µL de BSA à 0,8 µg/µL, en gardant à l'esprit que le volume total de solution du *crosslinking* est de 100 µL. Ensuite laisser le mélange de *crosslinking* 30 min à température ambiante. Puis la



réaction a été stoppée en ajoutant 5  $\mu\text{L}$  de Tris-HCl 1M (concentration finale de 50 mM) et on laisse incuber pendant 15 minutes. Ensuite, on passe à la purification par billes magnétiques Ni-NTA des protéines.

### **II.2.17.Purification par billes magnétiques Ni-NTA des protéines**

Après le crosslinking au DSP, un ajout de Tris-HCl a été pratiqué pour stopper la réaction. C'est alors que les complexes ont été récupérés grâce aux billes magnétiques (Qiagen) en mettant au moins 10  $\mu\text{L}$  de ces billes (saturation à 2 microgrammes de protéines/microlitre de billes magnétiques Qiagen tm). Après ajout des billes magnétiques, les mélanges ont été agités pendant une heure à 4°C. C'est alors que les tubes contenant les mélanges ont été placés sur un séparateur magnétique (un aimant) et que le surnageant a été retiré. Deux lavages ont été effectués avec 500  $\mu\text{L}$  de tampon d'interaction, en séparant les billes du surnageant 1 minute avec le séparateur, et ce à chaque lavage. Après les lavages, l'élution des protéines a été faite en mettant 30  $\mu\text{L}$  (dans le cas de mes expériences) de tampon d'élution. Les complexes billes-protéines ont été laissés une minute dans le tampon, puis laissés encore une autre minute sur le séparateur avant de récupérer les surnageants contenant les protéines.

Pour renverser les *crosslink* et séparer les protéines liées, le DTT ou le  $\beta$ -Mercaptoéthanol sont utilisables, Le DTT est présent dans le tampon de charge à un taux de 7,7 %.

Il est à noter que le crosslinking au DSP a été fait dans des tampons sans Tris, substance qui empêche la réaction de liaison de se faire. Pour ce faire, un changement de tampon a été pratiqué par ultrafiltration sur membrane MILLIPORE, en mettant 50 mM  $\text{NaH}_2\text{PO}_4$ , 300 mM NaCl, pH 8, pour les extraits de PepA et d'ArgR si nécessaire. On tâche de retrouver le même volume de solution, après centrifugation à 10000 rpm pendant 10 à 20 minutes. La centrifugation se fait à 4°C.

## II. 2.18. Construction, Surexpression et purification de PepANtMBP et de PepACtMBP

Lors de ces expériences, la partie Nt et Ct de PepA a été fusionnée au MBP. La première contient la région Nt de PepA et la majeure partie de la région centrale. La seconde contient la région Ct de PepA et la fin de la région centrale. Le gène *pepA* a été amplifié par PCR, à partir du plasmide pCS126. Mais, au lieu d'amplifier le gène en entier, on amplifie seulement soit la région Nt, soit la région Ct. Pour cela, on utilise les amorces suivantes à une concentration de 100 pmol/μL :

- PepA Ct HindIII stop: 5`GCTGCTAAGCTTTTACTCTTCGCCGTT3`
- PepA Ct PstI start : 5`GCGCGCGCTGCAGCACGGTC3`
- PepA Nt HindIII stop: 5`CACGTTGAAAAGCTTCTAACGCAGCGG3`
- PepA Nt PstI : 5`GCACTGCAGATGGAGTTTAGTGTA AAAAGC3`

Après amplification par PCR, une double digestion par PstI et HindIII a été effectuée, à la fois sur les fragments amplifiés et sur les plasmides pMALC2 (qui permettent de fusionner les protéines tronquées au MBP, dans chaque cas, la partie Nt de la protéine tronquée est fusionnée au MBP, la partie Ct étant composée soit de la région centrale pour PepANt, soit de la partie Ct pour PepACt).

Les cycles de PCR pour générer les fragments sont les suivants :

- 1) 95°C, 3 minutes
- 2) 95°C, 30 secondes
- 3) 55°C, 30 secondes
- 4) 72°C, 45 secondes
- 5) 72°C, 10 minutes
- 6) 4°C, pause

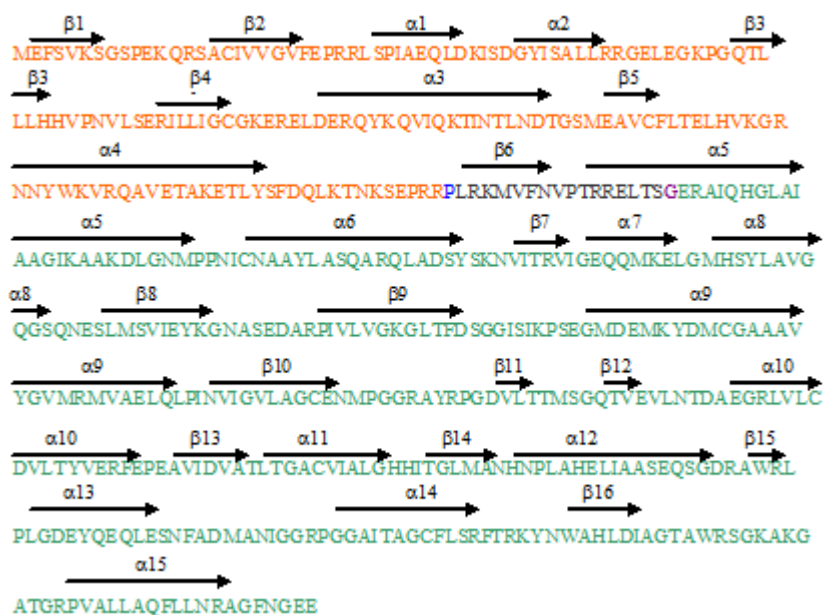
Les étapes 2 à 4, sont répétées pendant 30 cycles.

Les produits de PCR ont ensuite été purifiés par le QIA quick PCR purification kit de Qiagen Inc. Ensuite à une ligation a été faite toute la nuit à 16°C entre les

fragments et les plasmides traités par restriction enzymatique, grâce à la T4 DNA ligase.

Deux types de plasmides ont été ainsi obtenus:

- pMALc2pepaCt d'une taille de 7631 pbs contenant la partie Ct les acides-aminés 172 à 503, soit 331 acides-aminées sans le MBP.
- pMALc2pepaNt d'une taille de 7112 pbs contenant la partie Nt, contenant les acides-aminées 1 à 155, donc 155 acides-aminés sans le MBP (voir figure 51). Des cellules de la souche d'*Escherichia coli* DS955 ont été transformées par choc thermique, par l'un ou l'autre des deux plasmides. Après sélection des cellules transformées, les protéines sont induites par un milieu auto-inductible. La production des protéines a été induite toute la nuit sous agitation au bain-marie à 37°C. Les protéines font les poids moléculaires suivants, PepANtMBP = 60 kDas et PepACtMBP = 80 kDas.



**Figure 51.** Représentation de la protéine PepA d'*E.coli*. Sur cette figure on peut voir les hélices  $\alpha$  et les feuillets  $\beta$  indiqués par une flèche. La partie Nt que l'on retrouve dans la protéine PepANtMBP est indiquée en orange le dernier acide-aminé de cette partie (avant le MBP) étant indiqué en bleu. La partie en noir correspond à la partie déléetée. La partie Ct correspondant à PepACtMBP a été indiquée en bleu-vert et le premier acide-aminé de cette protéine est indiqué en violet.

## II. 2.19. Interaction des protéines PepANtMBP et PepACtMBP avec les dérivés d'ArgR

Lors des expériences où l'on veut savoir si ces protéines tronquées peuvent interagir avec les différentes versions d'ArgR, des *pulldowns* ont été pratiqués. Ensuite, des billes magnétiques à amylose ont été utilisées pour récolter les protéines fusionnées au MBP et éventuellement les protéines ArgR qui se seraient liées à PepA.

Pour le protocole de ces billes magnétiques, les étapes suivantes ont été effectuées:

- 1) Vortexer et resuspendre les billes magnétiques
- 2) Laver 2 fois les billes avec 500 microlitres de tampon pour colonne, on sépare les billes de la solution 30 secondes à une minute par un aimant.
- 3) Mettre 100 microlitres de billes magnétiques pour 400 microlitres de volume total (tampon d'interaction et extraits protéiques).
- 4) Incuber à 4°C pendant une heure.
- 5) Séparer les billes du surnageant avec l'aimant.
- 6) Faire trois lavages avec 500 microlitres de tampon pour colonne.
- 7) Mettre 30-50 microlitres de tampon colonne avec 10 mM de maltose.
- 8) Utiliser l'aimant et répéter.
- 9) Mettre de côté les éluions et garder.

Faire migrer les éluions sur gel SDS-PAGE pour vérifier s'il y a eu interaction entre les protéines PepA tronquées et les différentes versions d'ArgR.

## II.2.20. Recombinaison *in vitro* de pCS210

Ces expériences ont été réalisées avec le plasmide pCS210 (300 à 400 ng) contenant deux séquences *cer*, en présence des protéines XerC (300 nM monomère), XerD (300 nM monomère) venant de *Proteus mirabilis*, PepA (1800 nM de monomères) et des différentes versions d'ArgR (ArgRWt, ArgR149, ArgR5aa) chacune à 180 nM de monomères. Les différentes versions d'ArgR et XerD ont été purifiées par colonne Ni-NTA et la protéine XerC étant fusionnée au MBP à été purifiée par colonne

MBP. L'incubation des protéines avec l'ADN a été effectuée dans un volume total de 20 microlitres dans un tampon de recombinaison.

Après incubation des protéines avec l'ADN plasmidique pendant une heure à 37°C, une extraction des ADN par phénol, phénol-chloroforme et chloroforme a été effectuée pour retirer les protéines et récupérer l'ADN.

L'ADN a ensuite été précipité à l'éthanol 95%, à -20°C pendant 30 minutes, avant de culoter l'ADN et de resuspendre les plasmides dans du TE. C'est alors que l'on a digéré les plasmides extraits par EcoRV ou HindIII, avant de faire migrer les plasmides dans un gel d'agarose 0,75% dans un tampon TAE.

Une approche par Southern-blot d'un gel d'agarose 0,75 % a été choisie pour détecter les jonctions Holliday, un produit de la recombinaison. Cette visualisation est possible grâce aux sondes d'ADN, produites par PCR à partir du plasmide pCD34, contenant une séquence *cer* et s'hybridant avec les plasmides pCS210 lors du Southern blot.

Pour voir cela, le gel a été trempé dans une solution de dépurination pendant 15 minutes, avant de remplacer par 250 mL de solution de dénaturation et de laisser tremper 15 minutes. Cette dénaturation a été répétée une seconde fois puis la solution de dénaturation a été remplacée par 250 mL de solution de neutralisation pendant 15 minutes. Cette étape de neutralisation a été faite une deuxième fois. C'est alors que les ADN du gel ont été transférés sur une membrane hybond N+ en mettant les puits du gel vers le bas.

Dès lors, transférer toute la nuit. Après ce transfert, récupérer la membrane et tremper dans du tampon SSC 6X. Fixer ensuite les ADN aux UV avant d'effectuer la préhybridation. Lors de la préhybridation, la membrane a été placée dans un sac en plastique scellé avec 10 mL de solution de préhybridation préchauffée à 65°C. Agiter le tout à 65°C pendant 3 à 4 heures.

Lors de l'hybridation qui suit, on dénature les sondes à 100°C pendant 10 minutes, puis on met immédiatement sur glace. Pendant ce temps, on a mis à préchauffer à 65°C 20 mL de solution de préhybridation neuve. On en mettra un peu avec les sondes dénaturées conservées sur la glace avant de mettre les sondes dans 20 mL de solution et de laisser le tout s'hybrider toute la nuit à 65°C sous agitation dans un sac en plastique.

Après cette nuit d'hybridation, récupérer les sondes et les conserver à -20°C, puis laver la membrane 15 minutes dans du SSC2X 0.2%, puis dans du SSC1X 0,1% pour 15 minutes également. Pour voir les sondes marquées, les étapes suivantes ont été pratiquées. Laver d'abord une minute avec 15 mL de tampon 1, puis 20 min avec 30 mL de tampon 2. Laver ensuite 30 minutes avec 30 mL de tampon 2 contenant 6 microlitres d'anticorps anti-DIG. C'est alors que dans un nouveau bac, la membrane a été lavée deux fois de suite 15 minutes dans 30 mL de tampon 1. Un dernier lavage avec le tampon 3 a été fait pendant 2 minutes. Dès lors un Kit de détection, le phototope-Star Detection kit de NEB (tm) a été utilisé pour révéler les ADN hybridés. Pour cela, préparer la solution qui va réagir avec les Anticorps et le marquage (1,4 mL d'H<sub>2</sub>O distillée, 100 microlitres de diluant, 5 microlitres de CDP-star substrate de NEB), puis mettre la membrane entre 2 acétates avant de mettre la solution ci-dessus. Après 5 minutes, retirer l'excédent et sceller les acétates, puis placer la membrane sur un film et laisser exposer le film pendant 5 à 20 minutes dans le noir avant de développer.

### **II.2.21. Expériences sur la topologie de l'ADN**

Les expériences ont été réalisées avec l'utilisation des plasmides pNCK1, possédant une taille d'environ 5 kbs avec une séquence *cer* et une séquence *nmr* susceptibles de se lier à un complexe ArgR-PepA.

Lors de ces expériences, la protéine PepA a été purifiée différemment par rapport au protocole de purification habituel, pour éviter une dégradation de l'ADN. Le protocole décrit en II.2.10 a été utilisé lors de cette expérience.

La Topoisomérase I de la compagnie Promega a été utilisée pour relâcher les plasmides surenroulés après incubation à 37°C des 500 ng de plasmides avec les protéines ArgR (120 nM) et PepA (18 nM) pendant 15 minutes dans 50 microlitres d'une solution conçue spécialement pour l'expérience avec la Topoisomérase I.

Ensuite, une unité de Topoisomérase I a été rajoutée pour pouvoir relâcher l'ADN plasmidique, et laisser incuber le mélange à 37°C pendant 30 minutes avant d'arrêter la réaction par addition d'un tampon de charge contenant 0,2 mg/mL de protéinase K et 0,5% de SDS.

Ensuite, les échantillons ont été fait migrer pendant 19 à 24 heures sur un gel d'agarose *low* EEO à 0,8%, avec (ou sans) 0,5 microgramme/ml de chloroquine et avec du tampon 0,5X de Tris-borate-EDTA (TBE).

Une chasse de la chloroquine a ensuite été réalisée, qui est un agent intercalant sur l'ADN, grâce à deux bains de MgCl<sub>2</sub> à 1 mM, pendant une heure chacun. Ajouter ensuite un lavage du gel par 1L d'eau distillée, pendant 30 minutes. Colorer alors le gel au bromure d'éthidium (BET), 1 microgramme par mL, avant de laver 30 minutes à l'eau distillée. Puis analyser le gel via le programme Imager.

### **II.2.22.Imagerie par Typhoon**

La machine utilisée est le Typhoon trio, variable mode imager, compagnie Amersham Biosciences. Les programmes utilisés sont, pour la capture de l'image ; Typhoon Scanner Control 1 v5.0, et pour l'analyse et la retouche de l'image, ImageQuantTL v2005.

Cette machine a été utilisée pour voir les bandes d'ADN retardées marquées par HEX, avec détection de la fluorescence par laser 532 nm, filtre 555bp20.

### II.2.23. Mutagenèse dirigée sur *argR*

Lors des expériences de mutagenèses dirigées, les mutations ont été insérées dans les plasmides pGS38 en faisant des PCR avec des amorces contenant des mutations pour générer les codons stop et substitutions prévues sur la séquence *argR*. Toutes les mutations sont indiquées en caractères gras. Les concentrations pour les amorces sont environ de 100 pmol/μL.

Pour ArgR149 (la mutation stop est insérée dans le codon numéro 150), les amorces utilisées sont les suivantes :

F\_argR\_150 GTC AAA GAC CTG TAC GAA GCG ATT **TTA** CTG TTC GAC CAG  
GAGC

R\_argR\_150 GCT CCT GGT CGA ACA **GTT** ATA AAA TCG CTT CGT ACA GGT  
CTT TGA C

Après digestion du plasmide matrice par Dpn1, on transforme des cellules compétentes avec le produit de PCR, si la transformation est réussie, on a obtenu le plasmide pHSArgR149. La réussite de ces PCR et de ces transformations ont été vérifiées par séquençage.

Pour ArgRL149A, mutation de substitution de la leucine en position 149 par l'alanine les séquences sont les suivantes :

F\_argR\_L149A GAA GCG ATT **GCA** GAG CTG TTC GAC CAG

R\_argR\_L149A CTG GTC GAA CAG CTC **TGC** AAT CGC TTC

La façon de procéder à la transformation pour l'obtention du plasmide pQE31ArgRL149A identique à celle utilisée pour pHSArgR149.



Pour ArgRL151A, mutation de substitution de la leucine en position 151 par l'alanine les séquences sont les suivantes :

F\_argR\_L151A GAA GCG ATT TTA GAG **GCG** TTC GAC CAG

R\_argR\_L151A CTG GTC GAA **CGC** CTC TAA AAT CGC TTC

Pour ArgR150 (la mutation stop est insérée dans le codon numéro 151), les amorces utilisées sont les suivantes :

F\_argR\_151-JL4897/B9804 GTC AAA GAC CCT GTA CGA AGC GAT TTT AGA  
**GTA GTT** CGA GCA GGA GC

R\_argR\_151-JL4898/B9804 GCT CCT GGT CGA **ACT ACT** CTA AAA TCG CTT  
CGT ACA GGT CTT TGA C

Pour ArgR151 (la mutation stop est insérée dans le codon numéro 152), les amorces utilisées sont les suivantes :

F\_argR\_152-JL4899/B9804 GTC AAA GAC CTG TAC GAA GCG ATT TTA GAG  
CTC **TAA** GAC CAG GAG C

R\_argR\_152-JL4900/B9804 GCT CCT GGT **CTT** AGA GCT CTA AAA TCG CTT  
CGT ACA GGT CTT TGA C

Dans cette suite d'amorces présentée plus haut, les codons stop sont mis en évidence en gras.

Pour ces amorces F\_argR\_150 , R\_argR\_150, , les conditions de PCR sont les suivantes.

- 1) 95°C 1 min
- 2) 95°C 30 s
- 3) 55°C 1 min
- 4) 72°C 5 min
- 5) 72°C 10 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées pendant 19 cycles. La polymérase utilisée est la Vent DNA polymérase de chez NEB.

Pour la PCR impliquant les amorces F\_argR\_L149A, R\_argR\_L149A, F\_argR\_L151A et R\_argR\_L151A, les conditions sont les suivantes.

- 1) 98°C 1 min
- 2) 98°C 30 sec
- 3) 55°C 30 sec
- 4) 72°C 3 min
- 5) 72°C 10 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées pendant 30 cycles. La polymérase utilisée est la Phusion de chez NEB.

Pour les amorces F\_argR\_151-JL4897/B9804, R\_argR\_151-JL4898/B9804, F\_argR\_152-JL4899/B9804 et R\_argR\_152-JL4900/B9804 les conditions de PCR sont les suivantes :

- 1) 95°C 1 min
- 2) 95°C 30 s
- 3) 60°C 69 sec
- 4) 72°C 5 min
- 5) 72°C 10 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées pendant 25 cycles. La polymérase utilisée est la Vent DNA polymérase de chez NEB.

#### **II.2.24. Expérience de retard sur gel avec ArgBOX marquées par HEX**

Lors de cette expérience, on utilise l'amorce ArgboxD (dont la séquence est 5'CTT GCG GAT CCG AGC TTC G 3'), et l'amorce M13R-48HEX5' marquée à l'extrémité 5' par le marqueur fluorescent HEX (dont la séquence est 5'AGC GGA TAA CAA TTT CAC ACA GGA 3'). Cette séquence ArgBOX est amplifiée par PCR,

soit à partir du plasmide pGS38, soit à partir de la boîte arginine déjà purifiée. On obtient des fragments d'environ 115 pb.

Les concentrations des amorces sont de 100 pmol/ $\mu$ L. Les réactions de PCR sont réalisées dans des volumes totaux de 50  $\mu$ L et le cycle de PCR est le suivant :

- 1) 95°C 1 min
- 2) 95°C 30 s
- 3) 47°C 30 s
- 4) 72°C 15 s
- 5) 72°C 10 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées pendant 30 cycles. La polymérase utilisée est la Taq DNA polymerase de chez NEB

Une fois les produits de PCR purifiés, on les utilise pour les retards sur gels lors des études d'interaction protéine-ADN in vitro qui seront ensuite analysées par imagerie Typhoon.

#### **II.2.25. Expérience de retard sur gel avec ArgBOX marquées par la digoxygénine (dig)**

Lors de ces expériences, un fragment de 106 pb a été utilisé contenant la boîte arginine, qui sera marquée par la digoxygénine lors de la PCR. Les amorces utilisées, sont les suivantes, ArgboxD (dont la séquence est 5'CTT GCG GAT CCG AGC TTC G 3') et ArgboxG (dont la séquence est 5'TTT CAG CCG AAT TCA GGG CTG 3').

Les conditions de PCR sont les suivantes :

- 1) 95°C 1 min
- 2) 95°C 30 s
- 3) 55°C 30 s

- 4) 72°C 30 s
- 5) 72°C 1 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées durant 30 cycles. On utilise la Taq DNA polymérase de chez NEB. Les concentrations des amorces sont de 100 pmol/μL.

### **II.2.26. Expérience de retard sur gel avec les séquences *cer* marquées par la digoxygénine (dig)**

Lors de ces expériences, un fragment de 600 pb contenant deux demi-boîtes arginine a été utilisé, qui sera marqué par la digoxygénine lors de la PCR. Les amorces utilisées, sont les suivantes, M13-F-40-GH6222 (dont la séquence est 5'CGC CAG GGT TTT CCC AGT CAC GAC 3') et M13-R-48-GH6223 (dont la séquence est 5'AGC GGA TAA CAA TTT CAC ACA GGA 3'). Les concentrations en amorces sont de 493,6 pmol/μL pour M13-F-40 et de 482,4 pmol/μL pour M13-R-48. Les conditions de PCR sont les suivantes :

- 1) 95°C 2 min
- 2) 95°C 15 s
- 3) 42°C 30 s
- 4) 72°C 1 min
- 5) 72°C 5 min
- 6) 4°C pour stopper la réaction

Les étapes 2 à 4 sont répétées durant 30 cycles. On utilise la Taq DNA polymérase de chez NEB.

### **II.2.27. Imagerie par Imager**

L'appareil utilisé est le suivant, Alpha innotech chemi imager 5500, Innotech corporation, modèle de série 981651. Le programme utilisé est l'AlphaImager.

Cet appareil a été utilisé pour détecter les ADN colorés au BET, par réaction avec les rayons UV émis par la machine.

### II.2.28. Dosage des protéines par Bradford Assay (Kit Biorad TM)

Lors des dosages Bradford, diluer d'abord un réactif au 1/5 avec de l'eau déionisée stérile. Puis préparer une gamme d'étalonnage du BSA (de concentration 1mg/mL).

Le tableau suivant montre la gamme d'étalonnage.

<b>Concentration de BSA (mg/mL)</b>	<b>0</b>	<b>0,2</b>	<b>0,4</b>	<b>0,6</b>	<b>0,8</b>	<b>0,9</b>
<b>Volume de BSA (µL)</b>	<b>-</b>	<b>200</b>	<b>400</b>	<b>600</b>	<b>800</b>	<b>900</b>
<b>Volume H<sub>2</sub>O</b>	<b>1000</b>	<b>800</b>	<b>600</b>	<b>400</b>	<b>200</b>	<b>100</b>

**Tableau 12.** Gamme d'étalonnage pour le Bradford *assay*.

Lors de la réaction, on mélange 20 µL de chacun de ces tubes d'étalonnage avec 1 mL de réactif déjà dilué. Laisser incuber 5 min à température ambiante, puis mesurer la DO au spectrophotomètre à 595 nm.

On peut désormais faire un graphique dit de courbe d'étalonnage et en définir l'équation de la droite.

Dès lors, faire le dosage des protéines extraites des bactéries, avec des dilutions au 1/5 et au 1/10 dans un volume total de 50 µL pour chaque dilution. Soit 10 µL de protéine et 40 µL d'eau déionisée pour la dilution au 1/5, et 5 µL de protéines et 45 µL d'eau déionisée.

La réaction se fait avec 20  $\mu$ L de chaque dilution et 1 mL de réactif dilué. Mesurer ensuite les DO avec le spectrophotomètre à 595 nm. Les mesures ont été ensuite reportées sur un graphique, puis effectuer ensuite les calculs de concentrations. Le spectrophotomètre utilisé est un Biorad SmartSpec 3000.

### **II.2.29. Vérification de l'activité Aminopeptidase A**

Cette expérience consiste à vérifier si les différentes versions de PepA, soit PepAWt, PepANtMBP et PepACtMBP, sont toujours capables de couper des peptides au niveau de leur partie Nt. Pour cela, de la L-leucine p-nitroamilide hydrochloride (LPNA) a été mise en solution. Puis incuber à 37°C les solutions de LPNA avec PepA ajouté à l'intérieur pendant au moins 30 minutes. Si la solution prend une teinte jaune, la protéine PepA fonctionne et est active.

## **CHAPITRE III**

### **RÉSULTATS**

## CHAPITRE III.RÉSULTATS

### III.1.Introduction

Dans la partie « introduction » de cette thèse, on a vu que lors de la réplication du plasmide ColE1, des accidents de réplication pouvaient dimériser ce dernier. On a également décrit que la séquence *cer*, située sur ce plasmide, permettait de résoudre ces accidents en recrutant les recombinases XerC et XerD ainsi que les protéines accessoires ArgR (156 acides-aminés) et PepA (503 acides-aminés), ce qui génèrera deux monomères identiques.

Mon travail de thèse a consisté à étudier les interactions entre les deux protéines ArgR et PepA. Plus particulièrement la partie Ct de la protéine ArgR et cela grâce à différentes mutations générées dans notre laboratoire. Ces mutations empêchant la recombinaison du plasmide pCS210 ont été générées dans la partie Ct de la protéine ArgR soit par insertion du transposon Tn4430, soit par mutagenèse dirigée. Ces mutations n'affectent pas la capacité d'ArgR à se lier à l'ADN.

On avait vu précédemment dans l'introduction que la protéine AhrC était capable de compléter des *E.coli* déficientes en ArgR, malgré les différences de séquences. On a notamment remarqué dans de précédentes études, des similitudes intéressantes entre ArgR et AhrC dans la région C-terminale, au sein de la sixième hélice  $\alpha$  de la protéine ArgR et de la protéine AhrC. Cette région Ct avait été ciblée par de précédentes expériences de notre laboratoire car la partie Ct (d'après les modèles de Sträter et Reijns) entre en contact avec PepA et permet la recombinaison. On avait déjà mis en avant dans l'introduction que la partie Nt d'ArgR entre en contact avec l'ADN. On avait déjà obtenu les protéines mutantes ArgR149 et ArgR5aa qui ne permettent pas la recombinaison en *cer*, mais permettent toujours la liaison à l'ADN. Puisque l'on sait que c'est dans la région Ct que ces mutations empêchant la recombinaison sont situées, j'ai tout d'abord cherché à démontrer l'association entre ArgR et PepA au niveau de



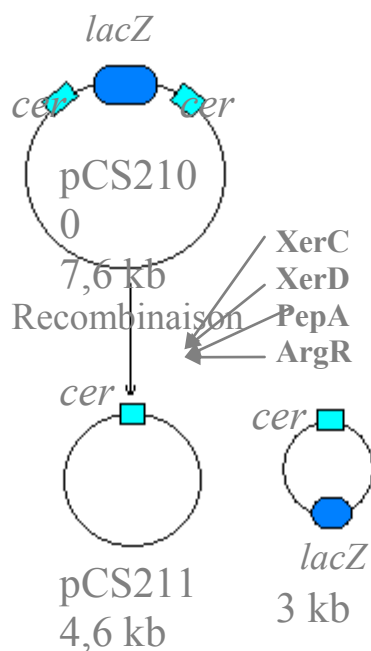
cette région C-terminale d'ArgR, puis à démontrer que des mutants d'ArgR en C-terminal ne pouvaient pas interagir correctement avec PepA. J'ai également cherché quelles régions de PepA interagissent avec ArgR.

Lors de cette thèse, de très nombreuses expériences ont été réalisées dans le but d'éclaircir ces interrogations.

### **III. 2. Caractérisation des mutants d'ArgR et PepA**

#### **III.2.1. Recombinaison *in vivo* du plasmide pCS210 dans les cellules DS956 avec différents mutants d'ArgR en C-terminal.**

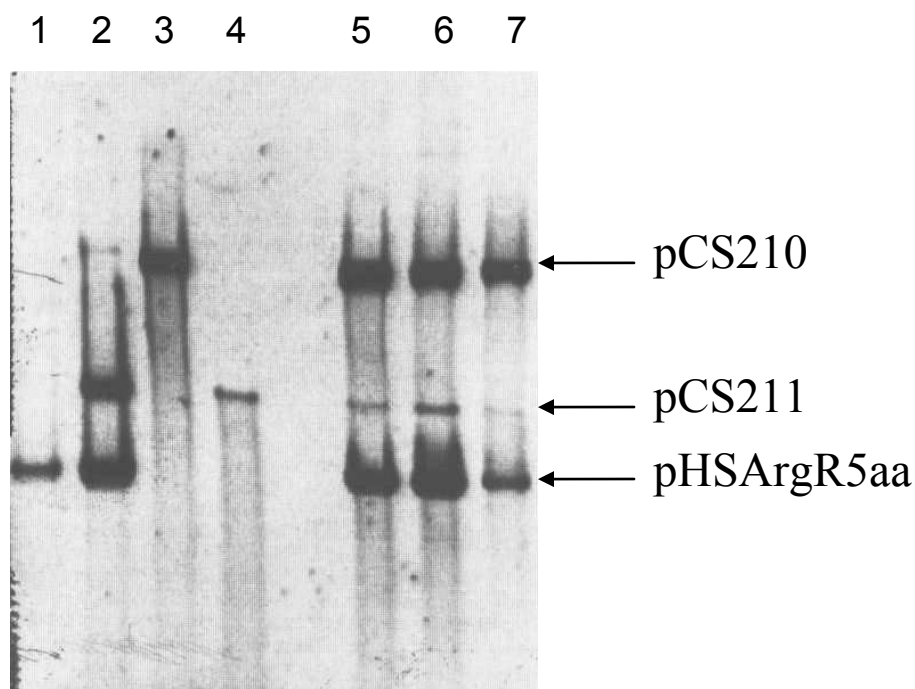
Lors de ces expériences, des cellules compétentes de la souche DS956 ont été transformées avec des plasmides contenant le gène des versions sauvages ou mutées de la protéine ArgR. Pour la version sauvage d'ArgR (soit ArgRWt), les plasmides pDB169, pGS38, pQE31ArgRWt ont été utilisés. Sur une gélose LB contenant de l'X-gal, de l'IPTG, de la tétracycline et ampicilline (surnommée gélose TAXI), les colonies sont devenues blanches (surtout en périphérie), indiquant une recombinaison au niveau des deux sites *cer* du plasmide pCS210 avec lequel on a déjà transformé les cellules DS956 avant de les rendre à nouveau compétentes. Cette recombinaison explique la coloration blanche des colonies, qui ont perdu le gène *lacZ* après recombinaison du plasmide pCS210, ce gène étant compris entre les deux séquences *cer* de ce plasmide. Sinon, s'il n'y a pas de recombinaison, elles apparaissent bleues, à cause de la présence du gène *lacZ* (figure 52).



**Figure 52.** Recombinaison pCS210 en pCS211.

Le plasmide pCS210 de 7,6 kb possède le gène *lacZ* flanqué de deux séquences *cer*. Lorsque toutes les protéines capables de réaliser la recombinaison *cer* spécifique sont présentes et fonctionnelles dans la cellule, il y a recombinaison intramoléculaire et perte de 3 kb avec conservation d'un seul site *cer*, et formation du plasmide pCS211. Les 3 kb perdus se retrouvent dans un plasmide de 3 kb qui sera rapidement perdu par les cellules.

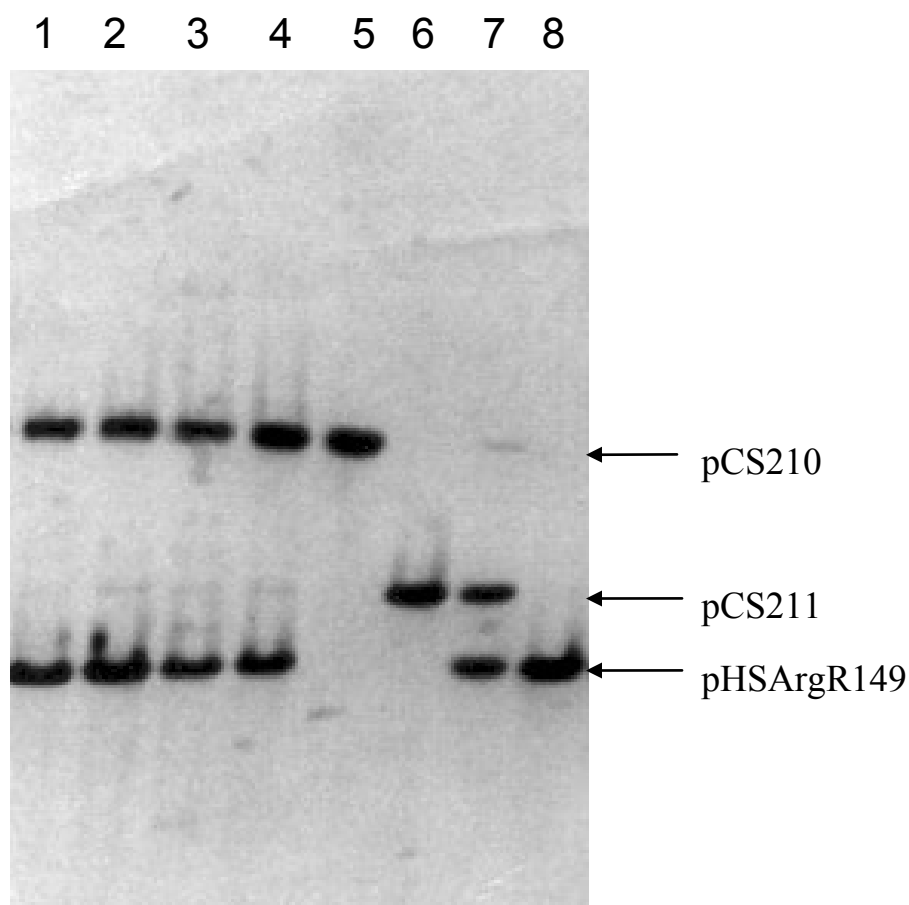
Ce qui est confirmé par extraction des plasmides contenus dans ces cellules, où après migration sur gel d'agarose 1%, il a été vu que le plasmide pCS210 a perdu 3 kb et est devenu le plasmide renommé pCS211 (figure 53, puits n°2).



**Figure 53.** Image en négatif d'une migration électrophorétique sur gel d'agarose 0,8% des plasmides pCS210, pCS211, pGS38 et pHSArgR5aa. Les plasmides pGS38 et pHSArgR5aa ont été extraits de cellules DS956 (*argR*<sup>-</sup>) transformées et digérés par HindIII. Le puits numéro 1 montre le plasmide pGS38 seul, le second puits montre la recombinaison donnant pCS211 en présence de pGS38. Le troisième puits montre pCS210. Dans le quatrième puits, on montre pCS211 seul. Dans les puits 5, 6 et 7, trois clones indépendants contenant les plasmides pHSArgR5aa ont été visualisés (avec autorisation, Sénéchal *et al.*, 2010).

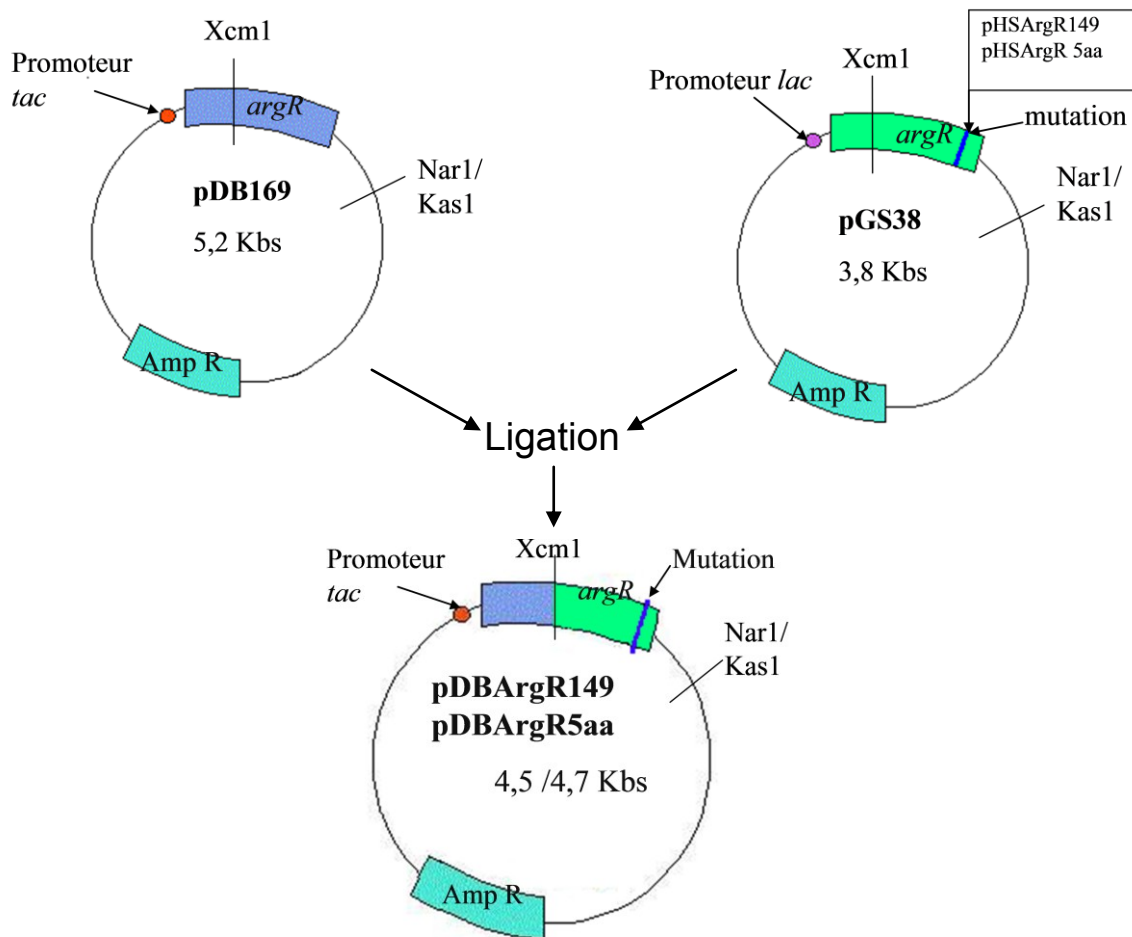
Pour la version mutante ArgR149 (générée par les plasmides pHSArgR149, pDBArgR149, pQE31ArgR149), sur le même type de gélose TAXI, on ne voit pas de perte de la coloration bleue des colonies de DS956. Ce qui indique que la recombinaison n'était pas détectable. Pour la version mutante ArgR5aa (générée par les plasmides pHSArgR5aa, pDBArgR5aa, pQE31ArgR5aa), sur le même type de gélose TAXI, la perte de la coloration bleue des colonies de DS956 n'a pas été visualisée en présence de ce mutant. Ce qui indique également que la recombinaison n'a pas eu lieu avec ArgR5aa. Ce qui signifie que les deux mutants ArgR149 et ArgR5aa empêchent la recombinaison au niveau des deux séquences *cer*. Des migrations sur gel d'agarose des plasmides pCS210, pCS211, pGS38 et pHSArgR5aa ont été précédemment réalisées après extraction des plasmides de cellules DS956. Ce qui permet de vérifier les

recombinaisons suspectées *in vivo*. Il a été remarqué qu'en présence de pHSArgR5aa, une très faible recombinaison de pCS210 a été observée, contrairement à ce qui a été observé avec pGS38. Des migrations sur gel d'agarose des plasmides pCS210, pCS211, pGS38 et pHSArgR149 ont été aussi réalisées après extraction des plasmides de cellules DS956. Ce, dans le but de vérifier ou d'infirmer les recombinaisons *in vivo*. Il a été remarqué qu'en présence de pHSArgR149, une très faible recombinaison de pCS210 a été observée comme avec pHSArgR5aa (figure 54).



**Figure 54.** Migration sur gel d'agarose 0,8% des plasmides pCS210, pCS211, pGS38 et pHSArgR149. Les puits 1 à 4 montrent 4 clones indépendants contenant les plasmides pCS210 en présence de pHSArgR149. Cette migration a été observée après extraction des plasmides de cellules DS956 (*argR*). Le puits numéro 5, montre le plasmide pCS210 seul. Dans le puits numéro 6, c'est le plasmide pCS211 qui est seul. Le puits numéro 7 montre la recombinaison de pCS210 en présence de pGS38. Le puits numéro 8 montre le plasmide pHSArgR149 seul (avec autorisation, Sénéchal *et al*, 2010).

Pour générer une meilleure surproduction des protéines ArgR149 et ArgR5aa, les plasmides pDB169 comportant le promoteur *tac* ont été recombines avec la partie de pHSArgR149 ou pHSArgR5aa comportant la mutation après coupure enzymatique. Les enzymes utilisées sont XcmI et NarI ou KasI (figure 55).

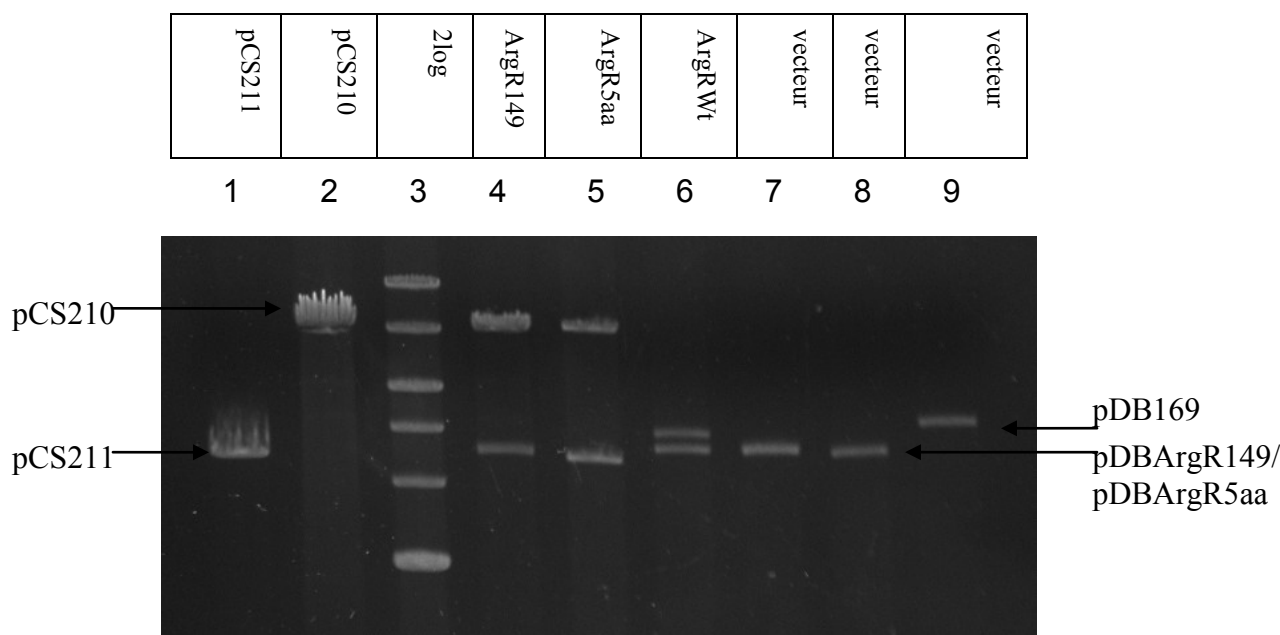


**Figure 55.** Représentation des plasmides pDB169 et pGS38 utilisés pour générer les plasmides pDBArgR149 et pDBArgR5aa. Les différences de promoteur, *tac* pour pDB169 qui génère ArgRWt et *lac* pour les dérivés de pGS38 (pHSArgR149 et pHSArgR5aa) qui comportent les mutations donnant les protéines ArgR149 et ArgR5aa ont été visualisé. Le promoteur *tac* devrait entraîner une meilleure surproduction des protéines après ligation. Ces constructions ont été vérifiées par séquençage.

Dans ces expériences, toute la partie de pDB169 comportant le promoteur *tac* (soit un fragment de 3,9 kbs) et la partie des pHSArgR149 et pHSArgR5aa comportant les mutations (soit un fragment d'environ 769 pbs que l'on insère dans celui de 3,9 kbs)

ont été récupérées. Si le plasmide pDB169 possède une taille de 5,2 kbs, le plasmide pDBArgR149 possède quand à lui une taille de 4,7 kbs et pDBArgR5aa fait une taille de 4,5 kbs.

Nous avons vérifié si le plasmide pCS210 inséré dans des cellules DS956 a subi ou non la recombinaison donnant le plasmide pCS211 après transformation par pDBArgR149, pDBArgR5aa ou pDB169 de ces cellules DS956 (figure 56).

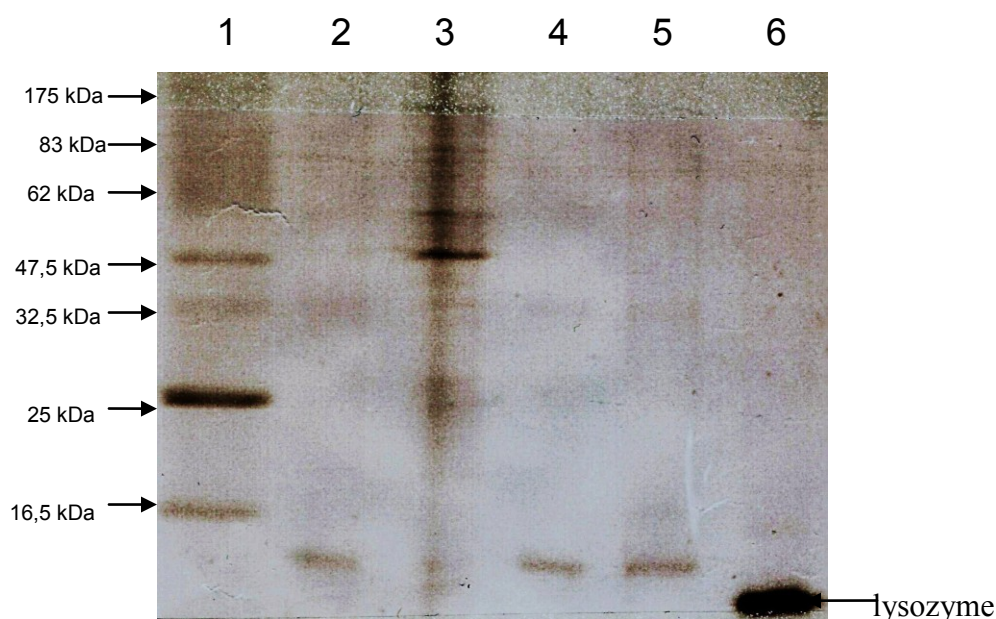


**Figure 56.** Migration sur gel d'agarose 0,8 % des plasmides pDBArgR149, pDBArgR5aa et pDB169 et leur influence sur la recombinaison du plasmide pCS210. Tous les plasmides présentés ici ont été linéarisés par HindIII. Le premier puits montre un témoin pCS211 (produit de recombinaison de pCS210). Le second puits montre un témoin pCS210. Le troisième montre le marqueur 2 Log. Les puits 4 à 6 montrent des extractions de plasmides de DS956. Le quatrième puits montre les plasmides pCS210 et pDBArgR149. Le cinquième puits montre les plasmides pCS210 et pDBArgR5aa. Le sixième puits montre les plasmides pCS211 et pDB169. Les puits 7, 8 et 9 montrent respectivement les plasmides pDBArgR149, pDBArgR5aa et pDB169 seuls et en temps que témoins.

Les résultats présentés avec les plasmides pGS38, pHSArgR149 et pHSArgR5aa avec ces plasmides dérivés du pDB169 ont pu donc être vérifiés. C'est alors que la purification partielle des protéines ArgRWt, ArgR5aa et ArgR149 par

précipitation à la L-Arginine a été tentée (figure 57, 58 et 59). Les résultats pour la purification d'ArgRWt et ArgR5aa montrent que l'on retrouve les protéines avec une pureté satisfaisante. Par contre, pour ArgR149, il a fallu faire une dialyse du surnageant juste après la sonication par rapport au protocole de purification partielle, ce qui correspond à une modification de ce protocole standardisé.

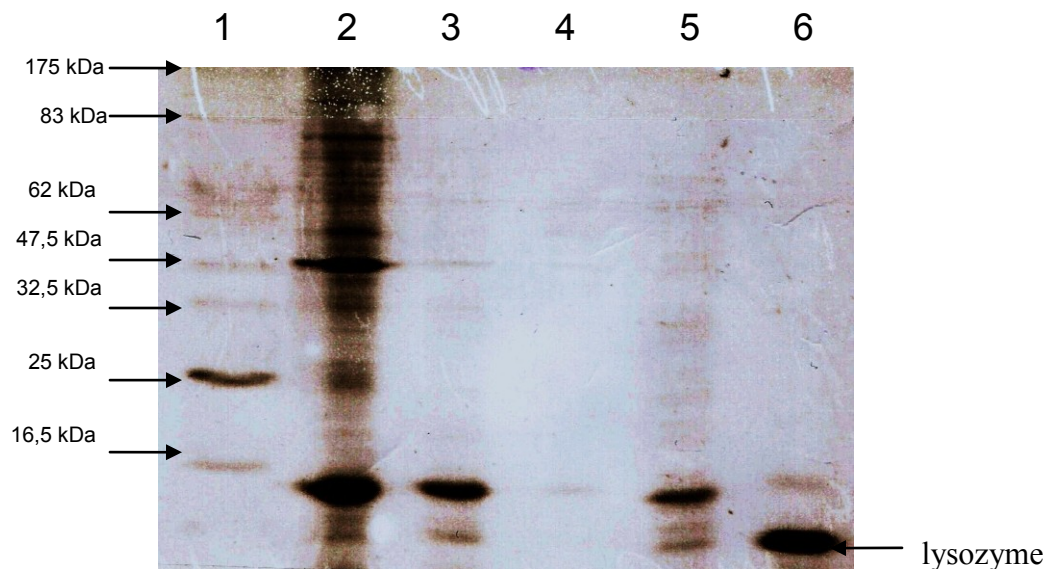
La photo qui suit montre une des purifications de la protéine ArgRWt.



**Figure 57.** Purification d'ArgRWt native. Le premier puits montre le marqueur. Le second puits montre le surnageant issu de la première précipitation au sulfate d'ammonium (29,1 %) et on peut y voir la protéine ArgRWt. Le troisième puits montre le culot après cette première précipitation. Le quatrième puits montre le culot après dialyse contre le tampon TM. Le cinquième puits montre le surnageant après dialyse et après précipitation par L-Arginine à 1 mM. Le sixième puits montre une protéine témoin, le lysozyme. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.

Pour ArgRWt, la purification a été jugée efficace.

La photo suivante montre une des purifications de la protéine ArgR5aa.

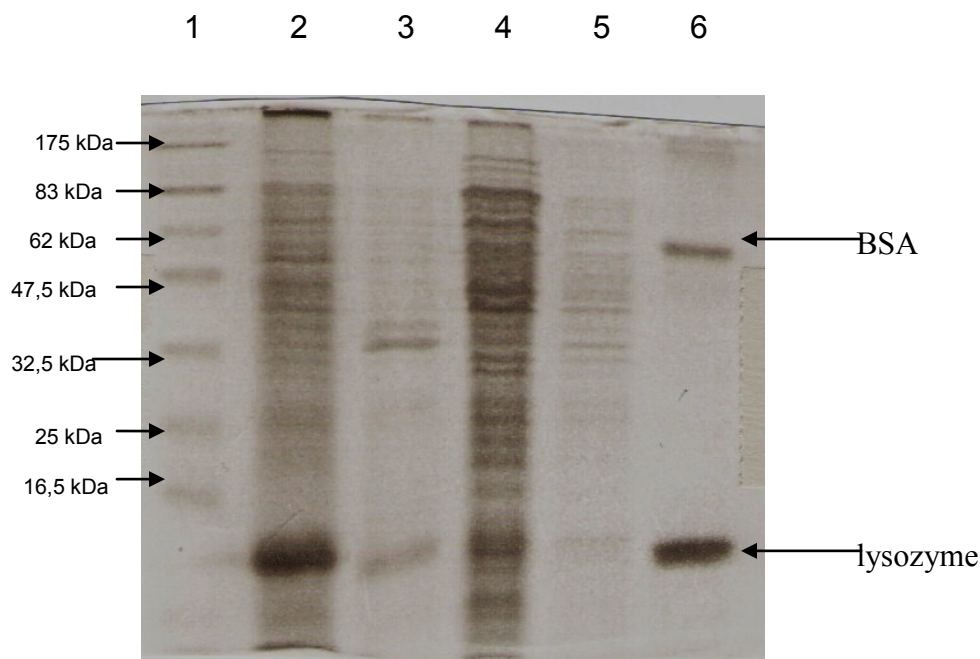


**Figure 58.** Purification d'ArgR5aa native. Le premier puits montre le marqueur. Le second puits montre le culot issu de la première précipitation au sulfate d'ammonium (29,1 %) et la protéine ArgR5aa y a été visualisée. Le troisième puits montre le surnageant après cette première précipitation. Le quatrième puits montre le culot après dialyse. Le cinquième puits montre le surnageant après dialyse contre le tampon TM et après précipitation par L-Arginine à 1 mM. Le sixième puits montre une protéine témoin, le lysozyme. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.

Pour ArgR5aa, la purification a été jugée efficace, même si des reconcentrations des protéines sont à envisager d'après cette photo, comme pour ArgRWt. La plupart des protéines ont également été précipitées par le sulfate d'ammonium 29,1 %.

La photo qui suit montre la tentative de purification de la protéine ArgR149. Les modifications au protocole standard de purification partielle incluent une première dialyse après sonication et une seconde après utilisation du sulfate d'ammonium.



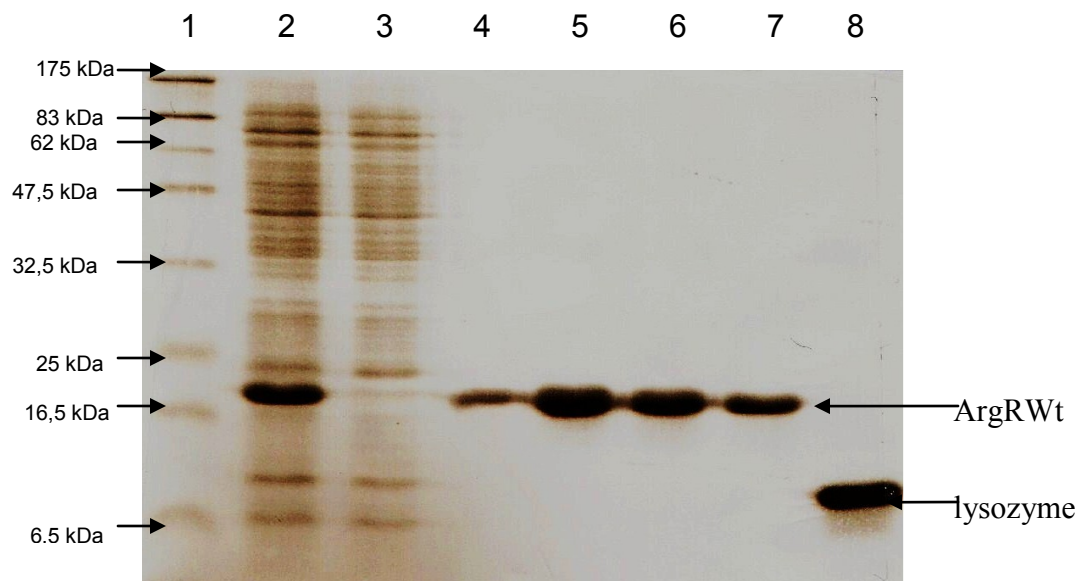


**Figure 59.** Purification de la protéine ArgR149. Le premier puits montre le marqueur. Le second puits montre le surnageant après dialyse de la sonication contre le tampon TM. Le troisième puits montre le surnageant issu de la resuspension après la seconde dialyse. Le quatrième puits montre le culot après dialyse du surnageant de la sonication contre le tampon TM. Le cinquième puits montre le précipité (culot) après dialyse contre le tampon TM. Le sixième puits montre un témoin BSA et lysozyme. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.

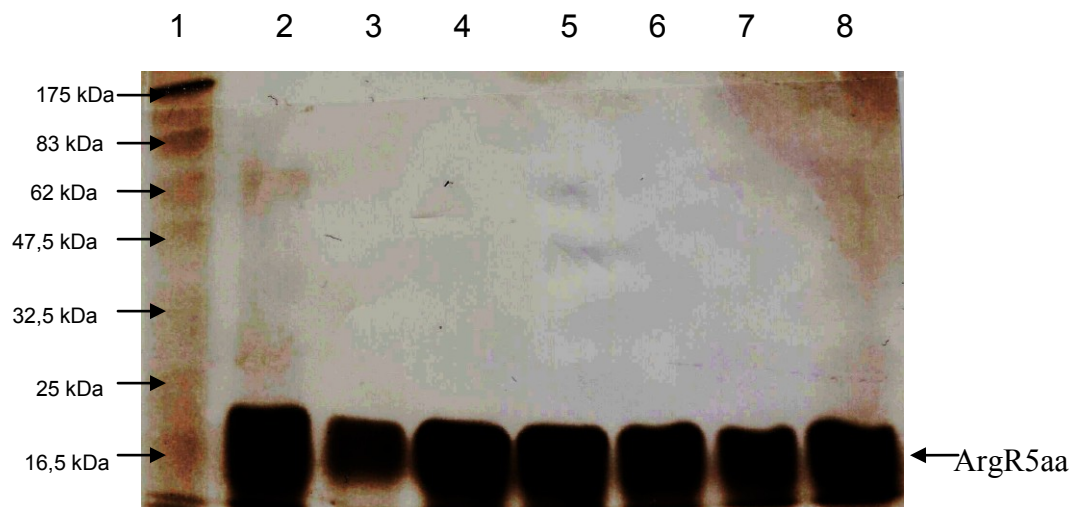
La purification d'ArgR149 montre que la protéine mutante a précipité très facilement dès la sonication et qu'il a trop de contaminants même après dialyse pour que les échantillons soient utilisables pour les expériences.

Ce qui m'amena à tenter la purification de protéines ArgRWt, ArgR5aa et ArgR149 après surexpression de ces protéines grâce à un vecteur pQE31 où les protéines seront fusionnées à 6-HIS (figure 60, figure 61, figure 62), ce qui permet de récupérer les protéines qui ont été captées par colonne Ni-NTA après élution avec imidazole. On obtient donc les plasmides pQE31ArgRWt, pQE31ArgR149 et pQE31ArgR5aa. Ces constructions ont été vérifiées par séquençage.

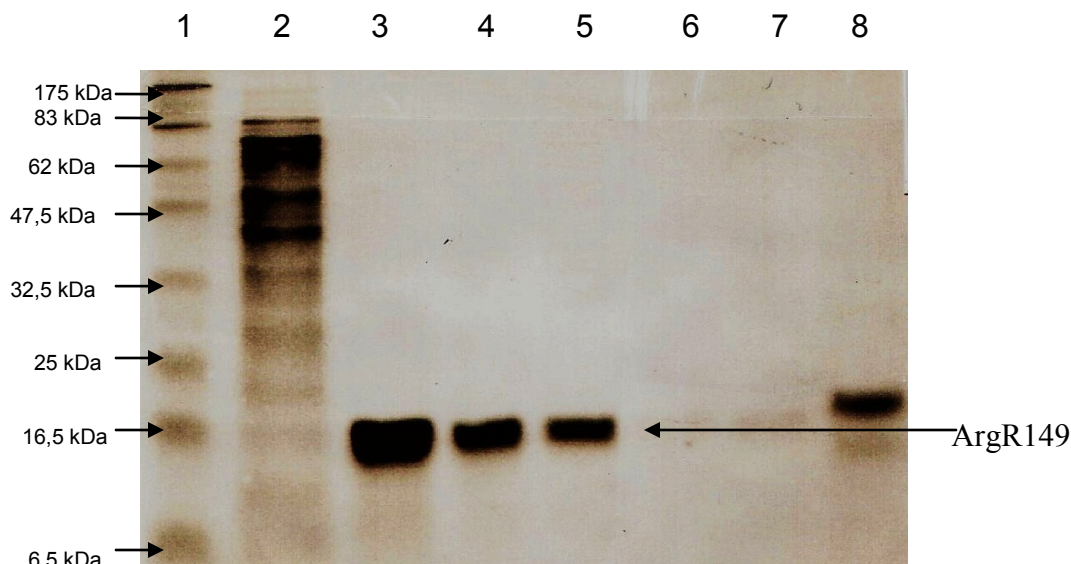
Les photos suivantes montrent les purifications des protéines fusionnées.



**Figure 60.** Purification de la protéine ArgRWt fusionnée avec 6-HIS. Le premier puits montre le marqueur. Le second puits montre le surnageant avant passage dans la colonne de résine Ni-NTA. Le troisième puits montre le lavage de la résine. Les puits 4 à 7 montrent respectivement les éluions 1 à 4 de la protéine ArgRWt à 150 mM d'imidazole. Le puits numéro 8 montre un témoin lysozyme. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.



**Figure 61.** Purification de la protéine ArgR5aa fusionnée avec 6-HIS. Le premier puits montre le marqueur. Le second puits montre l'éluion de la protéine à 100 mM d'imidazole. Les puits 3 à 6 montrent respectivement les éluions 1 à 4 de la protéine ArgR5aa à 150 mM d'imidazole. Les puits 7 et 8 montrent les éluions de la protéine ArgR5aa à 250 mM d'imidazole. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.



**Figure 62.** Purification des protéines ArgR149 fusionnées avec 6-HIS. Le premier puits montre le marqueur. Le second puits montre le lavage de la résine sur laquelle se sont fixées les protéines ArgR149. Les puits 3 et 4 montrent respectivement les éluions 3 et 4 de la protéine ArgR149 à 150 mM d'imidazole. Les puits 5 à 7 montrent respectivement les éluions 1 à 3 de la protéine ArgR149 à 250 mM d'imidazole. Le puits numéro huit montre un témoin ArgRWt fusionné à 6-HIS. Le gel SDS-PAGE contient 15 % d'acrylamide et a été coloré au bleu de Coomassie.

### III.2.2. Vérification de la répression de la fusion *argA::lacZ* dans les cellules EC146λ ( $\lambda$ AZ-7) par les différentes versions d'ArgR

Le but principal de ces expériences, est de vérifier la répression de la fusion *argA::lacZ* par ArgR. En effet, les mutations portent sur la partie Ct de la protéine, or c'est la partie Nt de la protéine qui se lie à l'ADN (Eckhardt 1980).

Durant ces expériences, les plasmides utilisés portent les gènes des versions mutantes ArgR5aa, ArgR149 et de la version sauvage ArgRWt. Ces versions ont été utilisées pour vérifier la répression de la fusion *argA::lacZ* portée par le chromosome de la bactérie. Il est important de savoir que ces cellules EC146λ ( $\lambda$ AZ-7) possèdent la construction chromosomique suivante, le promoteur de *argA* est situé en amont du gène rapporteur *lacZ* et cette région promotrice comporte une partie de la séquence codante

du gène *argA*. Donc ArgR peut se fixer sur cette région et réprimer l'expression du gène *lacZ*, et dans ce cas précis, la colonie sera blanche et donnera un phénotype *lac<sup>-</sup>* en présence d'ArgR.

Pour la transformation les cellules EC146 $\lambda$  ( $\lambda$ AZ-7), les plasmides pDB169, pGS38, pQE31ArgRWt ont été utilisés pour exprimer la protéine ArgRWt dans ces cellules. En effet, le plasmide pDB169 contient un promoteur *tac* plus fort que le promoteur *lac* utilisé pour les plasmides pGS38 et ses dérivés (pHSArgR149, pHSArgR5aa), tandis que les plasmides dérivés du vecteur pQE31 contiennent un promoteur T5 lui aussi plus fort que le promoteur *lac* utilisé pour les plasmides pGS38 et ses dérivés. Après culture sur gélose LB plus ampicilline, X-gal et IPTG, ou sur gélose Mac Conkey, les colonies étaient blanches, ce qui indique la répression de la fusion *argA :: lacZ*. Les plasmides pHSArgR5aa, pDBArgR5aa, pQE31ArgR5aa ont été utilisés pour produire ArgR5aa.

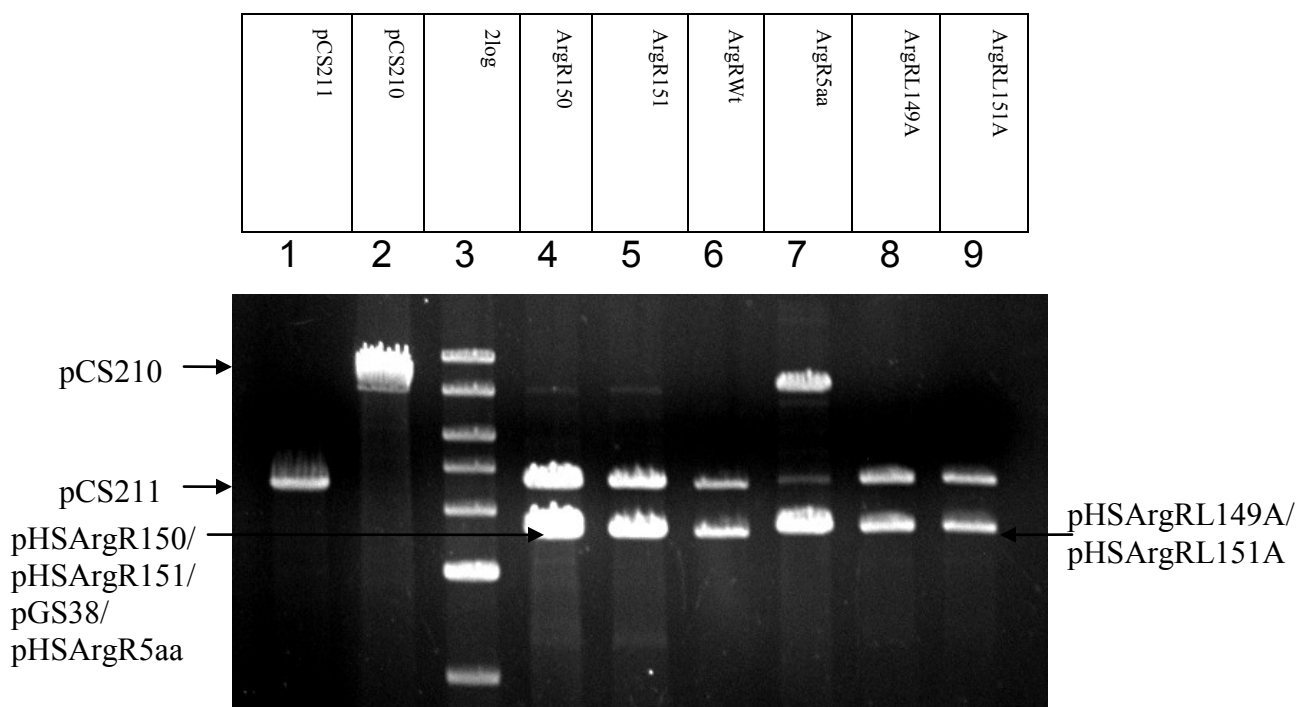
Après culture sur gélose MacConkey plus ampicilline (on n'a pas besoin de rajouter l'X-gal et l'IPTG), les colonies transformées par tous ces plasmides sont blanches (les colonies non transformées sont rouges). Ce qui indique également une répression de la fusion *argA :: lacZ* par la fixation d'ArgRWt sur le promoteur en amont du gène. Dans nos expériences, les colonies seraient rouges si le gène *lacZ* était actif. Les plasmides pHSArgR149, pDBArgR149, pQE31ArgR149 ont également été utilisés pour produire ArgR149. Sur un milieu MacConkey avec ampicilline, les colonies sont devenues blanches, ArgR149 est donc capable de réprimer le promoteur *argA*, comme ArgRWt.

Sur un milieu MacConkey pourvu d'ampicilline, les colonies transformées avec pHSArgR5aa, pDBArgR5aa, pQE31ArgR5aa sont devenues blanches, ArgR5aa est donc capable de réprimer le promoteur *argA*, comme ArgRWt. Toutes les versions étudiées dans ces expériences décrites ci-dessus permettent la répression de la fusion *argA :: lacZ*.

### III.2.3. Construction des ArgR tronquées

Lors de ces expériences, des mutagenèses dirigées ont été utilisées pour construire des versions du gène plus courtes que la version sauvage ArgRWt, les mutants ArgR150 et ArgR151. Les gènes portant ces mutations ont été clonés dans les vecteurs suivants : pHSArgR150 pour le mutant ArgR150 et pHSArgR151 pour le mutant ArgR151, ceci à partir du pGS38 dans lequel on a inséré les fragments d'ADN générés par PCR.

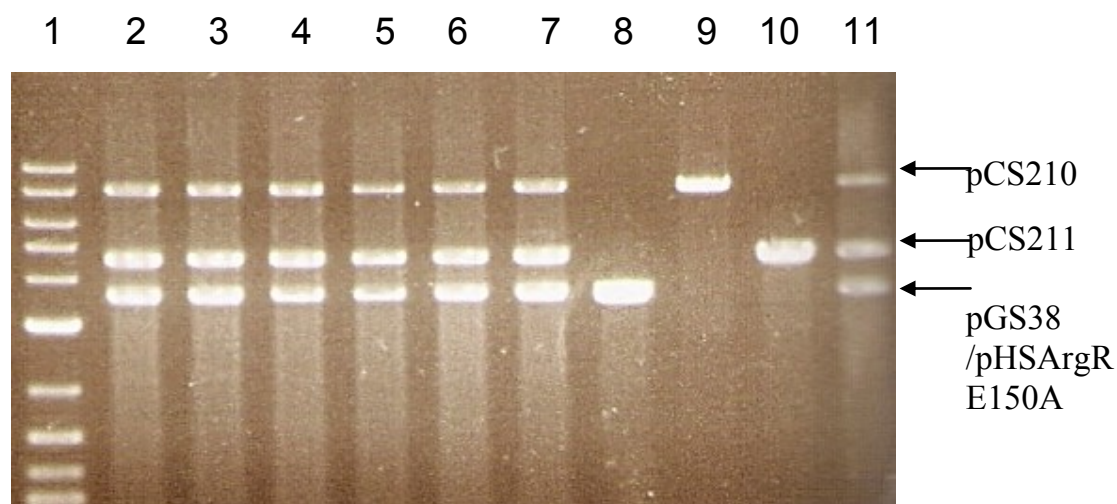
Comme des cellules DS956 possédant le plasmide pCS210 ont été transformées par l'un ou l'autre de ces plasmides, et que les colonies ont été incubées sur des géloses TAXI, il a été visualisé que les colonies devenaient blanches. Les plasmides de ces souches transformées par pHSArgR151 ou pHSArgR150 ont été extraits, et après migration sur gel d'agarose à 1%, il a été observé que les plasmides pCS211, signe de la recombinaison intra-plasmidique, se sont formés (figure 63). Ces constructions pHSArgR151 et pHSArgR150 ont été vérifiées par séquençage.



**Figure 63.** migration sur gel d'agarose 0,8 % des plasmides pHSArgR150, pHSArgR151, pGS38, pHSArgR5aa, pHSArgRL149A, pHSArgRL151A et leur influence sur la recombinaison du plasmide pCS210. Tous les plasmides présentés ici ont été linéarisés par HindIII. Le premier puits montre un témoin pCS211 (produit de recombinaison de pCS210). Le second puits montre un témoin pCS210. Le troisième montre le marqueur 2 Log. Les puits 4 à 9 montrent des extractions de plasmides de DS956. Le quatrième puits montre les plasmides pCS211 et pHSArgR150. Le cinquième puits montre les plasmides pCS211 et pHSArgR151. Le sixième puits montre pCS211 et pGS38. Le septième puits montre pCS210, pCS211 et pHSArgR5aa. Le huitième puits montre pCS211 et pHSArgRL149A. Le neuvième puits montre pCS211 et pHSArgRL151A.

Il en va de même pour les constructions pHSArgRL149A et pHSArgRL151A, qui ont aussi été séquencées et qui surexprime les protéines ArgRL149A et ArgRL151A. On voit encore des plasmides pCS210 dans les puits 4 et 5 de la figure 62 car il y a beaucoup de pCS210 par rapport au témoin dans le puits 6, malgré les repiquages, tous n'ont pas été recombinaison. La recombinaison de pCS210 a été visualisée avec ces versions d'ArgR, malgré le fait qu'elles ne possèdent qu'un acide-aminé de différence avec ArgR149 pour ArgR150, et deux acides-aminés pour ArgR151. De plus ArgRL149A possède la substitution de la leucine par l'alanine en

position 149, et ArgRL151A possède la substitution de la leucine par l'alanine en position 151. Ces deux nouvelles versions sont donc fonctionnelles, ce qui semblerait montrer que les résidus jusqu'à la position 150 (l'acide glutamique) auraient un rôle important dans la recombinaison. Mais il est important de signaler que la substitution ArgRE150A a toujours permis de générer des recombinaisons (figure 64). Si on voit encore des plasmides pCS210 non recombinés c'est parce que les clones bactériens n'ont qu'une génération après leurs transformations.



**Figure 64.** Migration sur gel d'agarose 0,8% de nombreux clones contenant les plasmides pCS210, pCS211, pGS38 et pHSArgRE150A. Les puits 2 à 7 contiennent chacun six clones indépendants des DS956 porteurs de pCS210 transformées par pHSArgRE150A. Le puits numéro 8 montre pGS38. Le puits numéro 9 montre le plasmide pCS210 seul. Le puits numéro 10 montre pCS211 seul. Le puits numéro 11 montre les plasmides pCS210, pCS211 et pGS38. Tous les plasmides ont été linéarisés par HindIII.

#### III.2.4. Mutagenèse site dirigée en positions 149 et 151

Ces substitutions par l'alanine (A) ont été générées par mutations dirigées grâce à la technique de PCR et cloné dans des plasmides de type pHS, comme expliqué dans la partie III.2.3, puis les fragments comportant les mutations ont été sous-clonés dans pQE31, pour donner les plasmides pQE31ArgRL149A et pQE31ArgRL151A, exprimant respectivement les protéines ArgRL149A et ArgRL151A. Ces plasmides recombinants ont été transformés dans des cellules DS956 comprenant déjà le plasmide

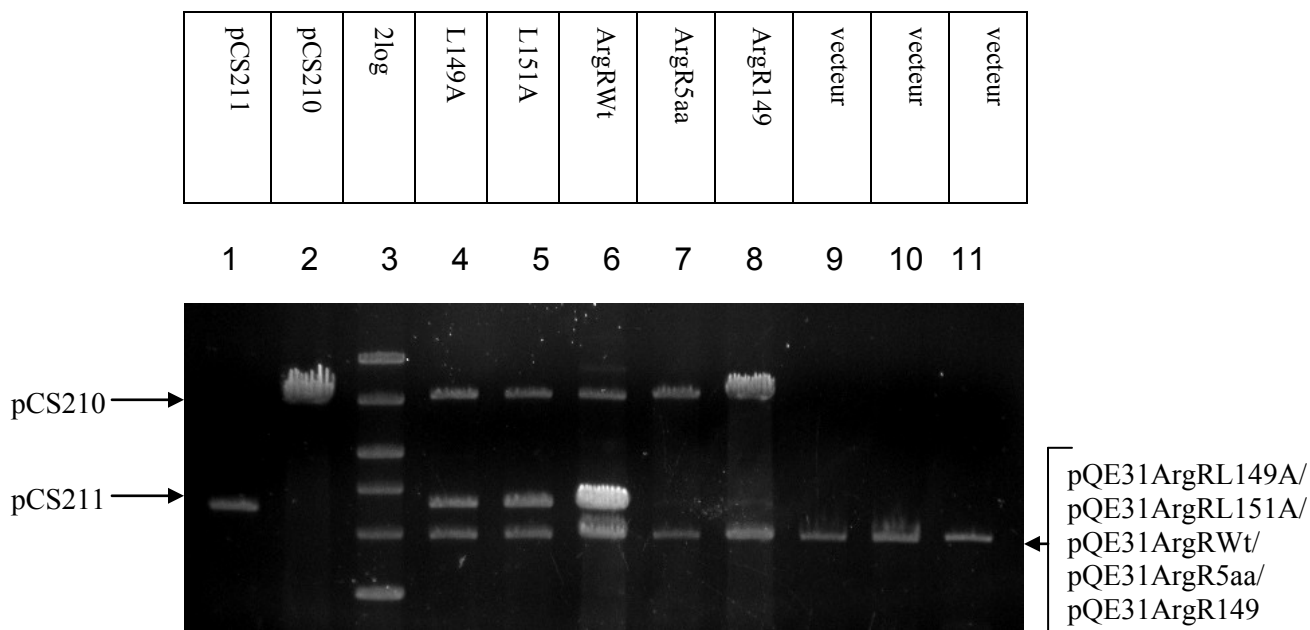
pCS210. Comme pour les précédentes transformations, une recombinaison intramoléculaire générant le plasmide pCS211 a été recherchée.

Lorsque les colonies après transformation par l'un ou l'autre des plasmides ont été observées, une tendance à devenir blanches de la part de ces colonies a été aperçue, au moins en périphérie desdites colonies sur le milieu TAXI. Ce qui indique une perte du gène *lacZ* au fur et à mesure des générations.

Les plasmides ont été extraits et une recombinaison de pCS210 a été observée dans les deux cas (c'est-à-dire avec pQE31ArgRL149A et pQE31ArgRL151A), donc la formation de pCS211. Si un résidu Leucine (en position 149 ou en position 151) et un résidu Acide glutamique (en position 150) ont été conservés dans cette région de la sixième hélice  $\alpha$  de la protéine ArgR, une recombinaison de pCS210 a toujours été observée (figure 65).

Mais comme il a été précédemment observé, s'il reste les deux résidus leucine et que l'acide glutamique est conservé dans cette hélice, il y a toujours recombinaison pour le plasmide pCS210, car il y a apparition des plasmides pCS211.





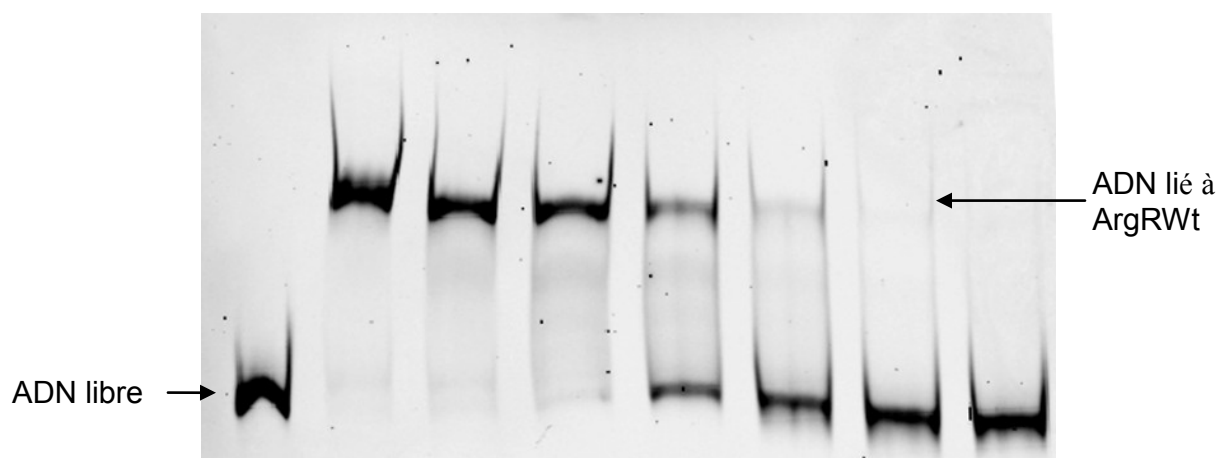
**Figure 65.** Migration sur gel d'agarose 0,8 % des plasmides pQE31ArgRWt, pQE31ArgR149, pQE31ArgR5aa, pQE31ArgRL149A, pQE31ArgRL151A et leur influence sur la recombinaison du plasmide PCS210. Tous les plasmides présentés ici ont été linéarisés par HindIII. Le premier puits montre un témoin pCS211 (produit de recombinaison de pCS210). Le second puits montre un témoin pCS210. Le troisième montre le marqueur 2 Log. Les puits 4 à 8 montrent des plasmides issus extraient de DS956. Le quatrième puits montre les plasmides pQE31ArgRL149A, pCS210 et pCS211. Le cinquième puits montre les plasmides pQE31ArgRL151A, pCS210 et pCS211. Le sixième puits montre les plasmides pQE31ArgRWt, pCS210 et pCS211. Le septième puits montre les plasmides pQE31ArgR5aa, pCS210 et pCS211. Le huitième puits montre les plasmides pQE31ArgR149, pCS210 et pCS211. Les puits 9 à 11 montrent respectivement les plasmides pQE31ArgRL149A, pQE31ArgRL151A et pQE31ArgRWt.

### III. 3. Activité de liaison des mutants avec l'ADN *in vitro*

#### III.3.1. Caractérisation des activités de liaison d'ArgRWt, ArgR149, ArgR5aa sur la séquence boîte arginine

Les expériences de retard sur gel d'acrylamide des séquences d'ADN ArgBOX (boîte arginine) marquées par HEX ont été réalisées pour caractériser la capacité des formes d'ArgR à se lier à l'ADN. Les différentes migrations ont été faites avec les

formes purifiées d'ArgR suivantes : ArgRWt, ArgR149, ArgR5aa. Dans chaque puits et pour chaque expérience de retard sur gel, 2 ng d'ADN (soit 1,3 nM) et les versions fusionnées avec 6-HIS de ces protéines ont été utilisées en présence de 1 µg de poly-dIdC. Ces migrations ont eu lieu à 4°C et avec du tampon TBE 0,25 X (comme dans Chen *et al*, 1997). Dans ces conditions, les protéines ArgRWt, ArgR5aa et ArgR149 sont capables de se multimériser, et les concentrations données ci-dessous ont été calculées grâce à biomath calculator, programme informatique accessible sur le site de Promega. Toutes es concentrations sont présentées en ordre décroissant.

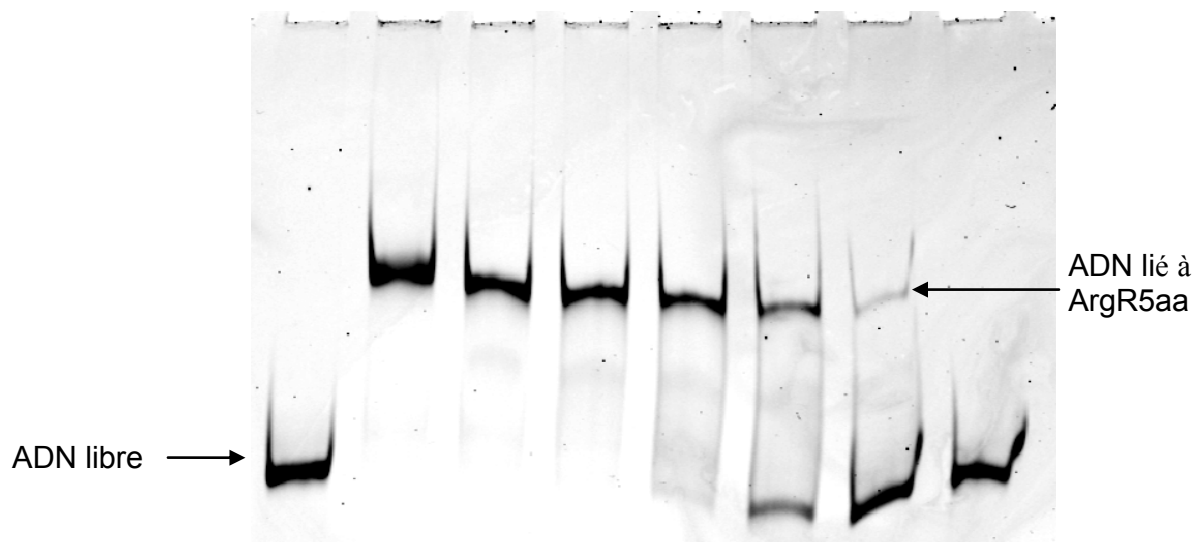


Numéro du puits	1	2	3	4	5	6	7	8
[ArgRWt]	0 nM	576 nM	288 nM	144 nM	72 nM	36 nM	18 nM	9 nM
µg	0	0,2	0,1	0,05	0,025	0,0125	0,00625	0,00313

**Figure 66.** Retard sur gel d'ArgRWt avec la boîte arginine. Les concentrations de protéines sont données en nanomolaires de monomères sur la séquence d'ADN. Dans chaque puits, 2 ng d'ADN (soit 1,3 nM) et 1 µg de poly-dIdC ont été utilisés. Ces migrations ont eu lieu à 4°C et avec du tampon TBE 0,25 X dans un gel d'acrylamide 5 %.

Pour la protéine ArgRWt, un retard net de la séquence ArgBOX en présence de la protéine de type sauvage a été bien vu, correspondant à la liaison d'un hexamère d'ArgRWt sur l'ADN (figure 66). Moins il y a de protéine, moins il y a d'ADN retardé, d'ailleurs des brins libres et des brins retardés pour des concentrations de 72 et 36 nM (très légèrement pour une concentration 91,2 nM) ont été visualisés. À des

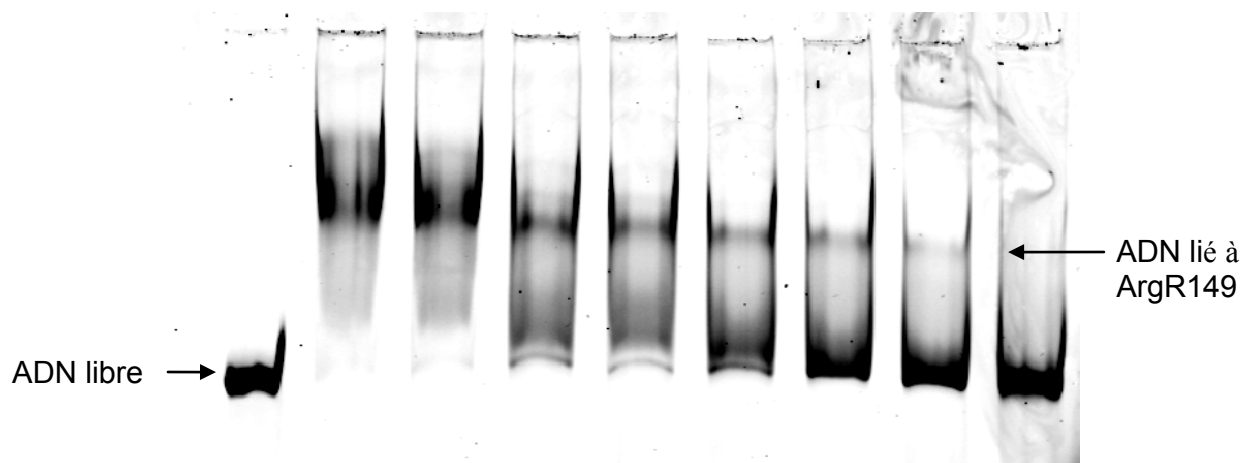
concentrations de 96 nM et 48 nM des brins d'ADN ont été totalement retardés. Il a été visualisé que pour une concentration de 72 nM de protéine, la moitié des ADN sont retardés.



Numéro du puits	1	2	3	4	5	6	7	8
[ArgR5aa]	0,0	2910	1455	727,5	363,6	181,8	91,2	45,6
	nM	nM	nM	nM	nM	nM	nM	nM
µg	0	1	0,5	0,25	0,125	0,0625	0,0313	0,0156

**Figure 67.** Retard sur gel d'ArgR5aa avec la boîte arginine. Les concentrations de protéines sont données en nanomolaires de monomères sur la séquence d'ADN. Dans chaque puits, on a 2 ng d'ADN (soit 1,3 nM) 1 µg de poly-dIdC. Ces migrations ont eu lieu à 4°C et avec du tampon TBE 0,25 X dans un gel d'acrylamide 5 %.

Pour la protéine ArgR5aa, le même type de retard que pour ArgRWt a été vu. Il a été observé que le retard est d'environ de 50 % pour 30,3 nM d'ArgR5aa, soit deux fois moins que pour ArgRWt (figure 67). Un retard total de l'ADN a été observé pour des concentrations de 2910 nM à 363,6 nM, alors qu'un mélange brins d'ADN retardés et libres a été observé pour des concentrations de 181,8 nM à 91,2 nM.



Numéro du puits	1	2	3	4	5	6	7	8	9
[ArgR149]	0	4911	3675	1837,8	1227,6	918	613,8	459,6	229,8
	nM	nM	nM	nM	nM	nM	nM	nM	nM
µg		1,67	1,25	0,63	0,42	0,31	0,21	1,56	0,78

**Figure 68.** Retard sur gel d'ArgR149 avec la boîte arginine. Les concentrations de protéines monomériques sont données en nanomolaires sur la séquence d'ADN. Dans chaque puits, 2 ng d'ADN (soit 1,3 nM) et 1 µg de poly-dIdC ont été utilisés. Ces migrations ont eu lieu à 4°C et avec du tampon TBE 0,25 X dans un gel d'acrylamide 5 %.

Pour la protéine ArgR149, un retard de la séquence ArgBOX a été vu, notamment pour des concentrations de 4911 nM et 3675 nM. Mais lorsque les concentrations en ArgR149 ont été diminuées, l'apparition d'une traînée entre l'ADN libre et l'ADN retardé a été observée, surtout pour des concentrations de 1837,8 nM à 459,6 nM. Il a été visualisé que le retard commençait à s'observer à partir de 456,6 nM de protéine (figure 68).

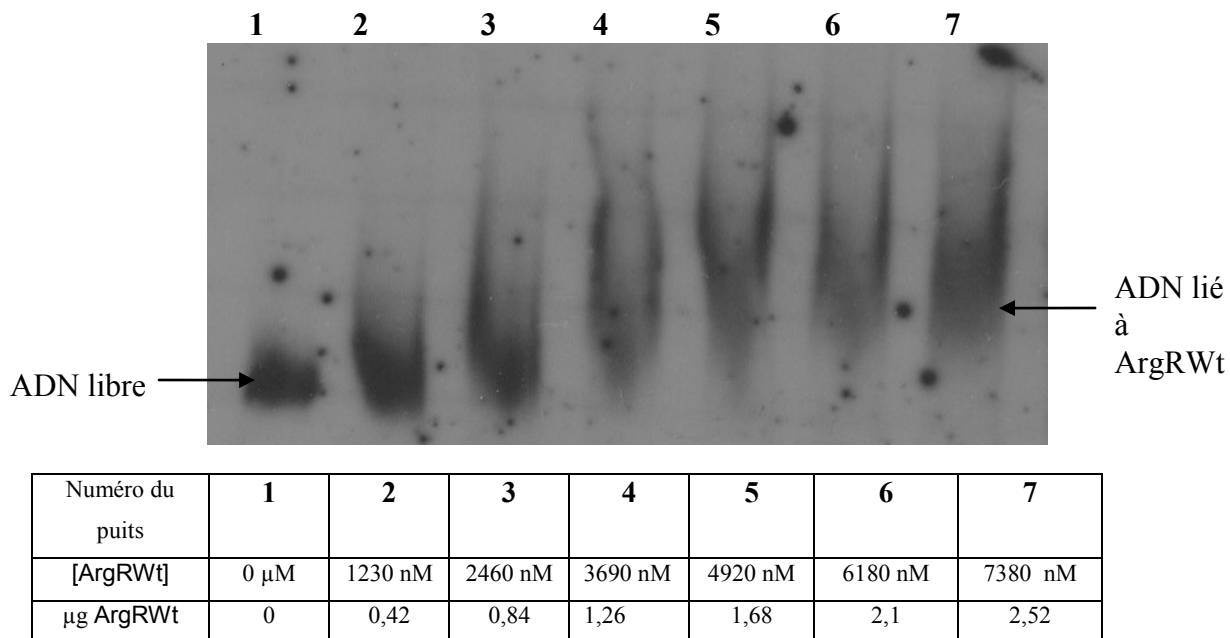
Des retards sur gels ont été également effectués avec les formes natives des protéines ArgRWt, ArgR149 et ArgR5aa avec deux séquences *cer* séparées de 600 pb dans la partie **III.3.2.**

### III.3.2. Caractérisation des activités de liaison d'ArgRWt, ArgR149, ArgR5aa sur la séquence *cer*

Des expériences sur gel d'acrylamide ont été réalisées comme précédemment mais avec deux séquences *cer* directement répétées et séparées de 600 pb, pour permettre une liaison de l'ADN autour de l'hexamère d'ArgR, la séquence *cer* ne comportant qu'une demie-boîte arginine. *In vivo*, cette construction est capable de subir la recombinaison lorsqu'elle est clonée dans un plasmide. La migration a eu lieu en présence de L-arginine pour permettre la fixation des hexamères d'ArgR sur les deux séquences *cer* d'ADN. Les versions natives des protéines ArgRWt, ArgR149 et ArgR5aa ont été mises en présence des deux séquences *cer*, à différentes concentrations, pour vérifier si ces formes de la protéine ArgR étaient toujours capables de se lier à l'ADN de deux séquences *cer*. Dans ces expériences, les protéines ArgR et ses mutants ont été purifiés par purification partielle. Les migrations effectuées à température ambiante et avec du tampon TBE 0,5 X (comme dans Sénéchal *et al*, 2010) ont donné les profils suivants avec 2 ng d'ADN (soit 0,25 nM) et 1 µg de poly-dIdC. Ces techniques de retards sur gel sont inspirées de Tian *et al*, 1992 et de Stirling *et al*, 1988b.

Dans ces conditions, les protéines ArgRWt, ArgR5aa et ArgR149 sont capables de se multimériser, et les concentrations données ci-dessous ont été calculées grâce à biomath calculator, programme informatique accessible sur le site de Promega.

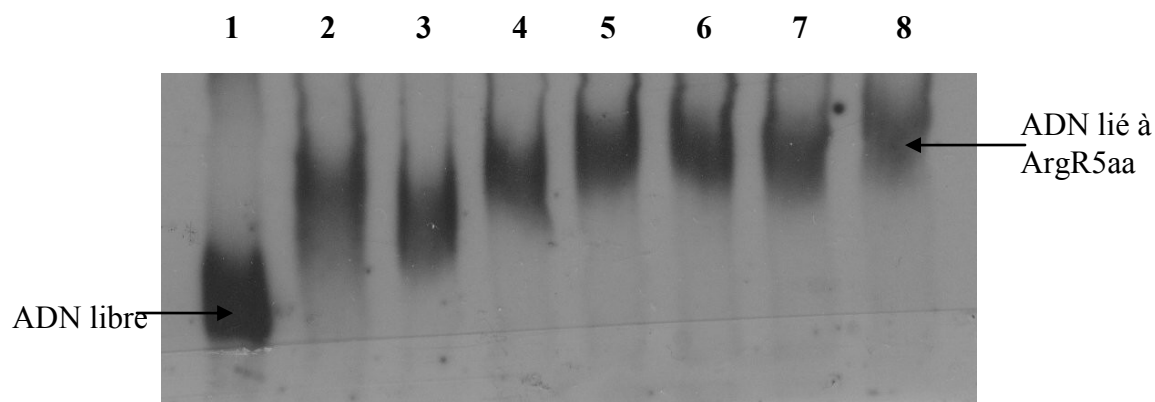
Voici Les résultats avec la forme sauvage.



**Figure 69.** Retard sur gel d'ArgRWt avec deux séquences *cer*. Les concentrations de protéines ArgRWt monomériques sont indiquées. Le pourcentage de pureté de la protéine est de 79 % et a été pris en compte dans les calculs de concentration. Dans chaque puits, 2 ng (0,25 nM) d'ADN et 1  $\mu$ g de poly-dIdC ont été utilisés. Les migrations ont été effectuées à température ambiante et avec du tampon TBE 0,5 X dans un gel d'acrylamide 5 %.

Ces retards montrent que la forme sauvage se lie à deux séquences *cer*, qui reforment une boîte arginine entière. Plus il y a de protéines, plus un retard a été visualisé, jusqu'à un certain seuil où les séquences ne migrent plus, traduisant une liaison de tous les ADN à la protéine ArgRWt (normalement sous forme hexamérique) et ce à partir de la concentration 3690 nM, un retard de l'ADN situé au même niveau ayant été visualisé pour une concentration de 6180 nM et 7380 nM (figure 69).

Les résultats avec la forme mutante ArgR5aa sont présentés sur la figure suivante.

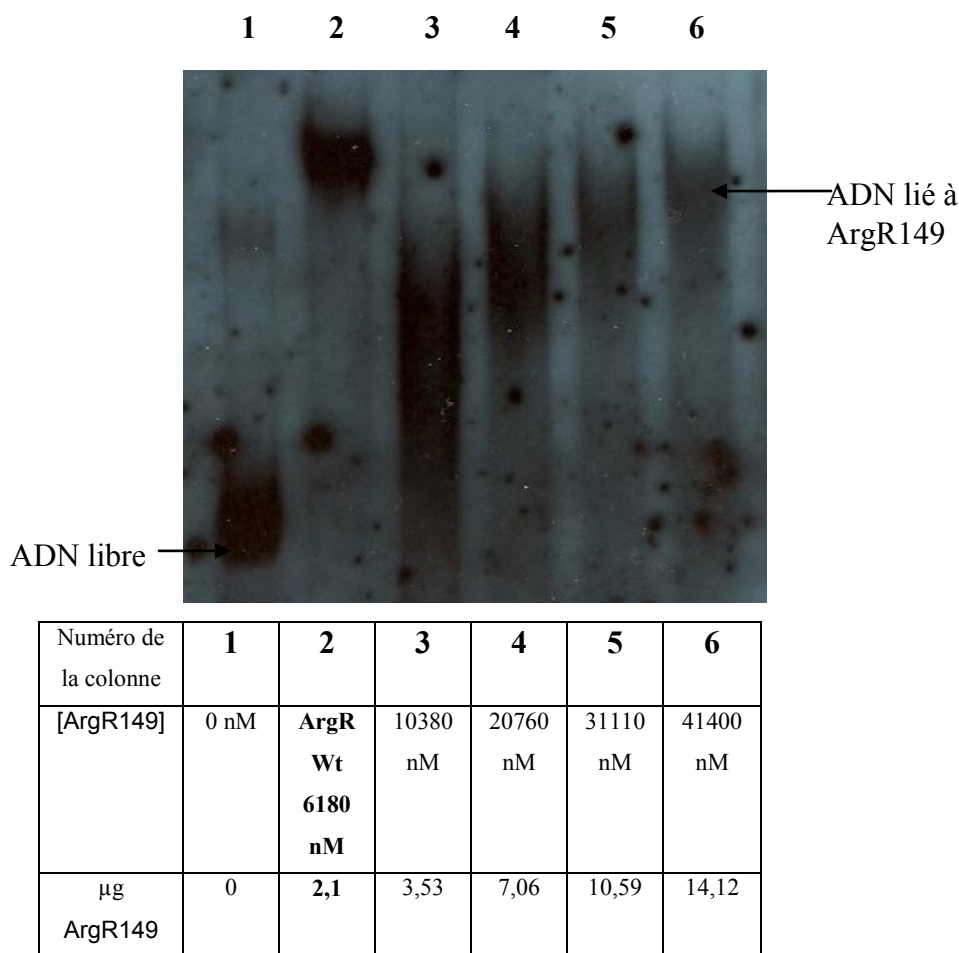


Numéro du puits	1	2	3	4	5	6	7	8
[ArgR5aa]	0 nM	<b>ArgR Wt 6180 nM</b>	2496 nM	4998 nM	7500 nM	10020 nM	12480 nM	15000 nM
µg ArgR5aa	0	2,1	0,85	1,7	2,55	3,4	4,25	5,1

**Figure 70.** Retard sur gel d'ArgR5aa avec deux séquences *cer*. Les concentrations des protéines ArgRWt et ArgR5aa monomériques sont indiquées. Le pourcentage de pureté de la protéine ArgR5aa est de 82 % et a été pris en compte dans les calculs de concentration. Dans chaque puits, 2 ng (0,25 nM) d'ADN et 1 µg de poly-dIdC ont été utilisés. Les migrations ont été effectuées à température ambiante et avec du tampon TBE 0,5 X dans un gel d'acrylamide 5 %.

Ces retards montrent que la forme ArgR5aa de la protéine permet aussi le retard de la séquence d'ADN contenant deux séquences *cer* utilisées lors des expériences. Le retard engendré par ArgR5aa a été comparé à celui engendré par ArgRWt, et il a été visualisé que la forme mutante ArgR5aa est également capable de générer un retard. Comme avec ArgRWt, il y a une concentration pour laquelle tous les ADN sont retardés (7500 nM), et cela à partir du puits 5, un retard de l'ADN situé au même niveau ayant été visualisé pour les puits 6 à 8 (figure 70).

Les résultats pour la protéine ArgR149 sont présentés sur la figure suivante.



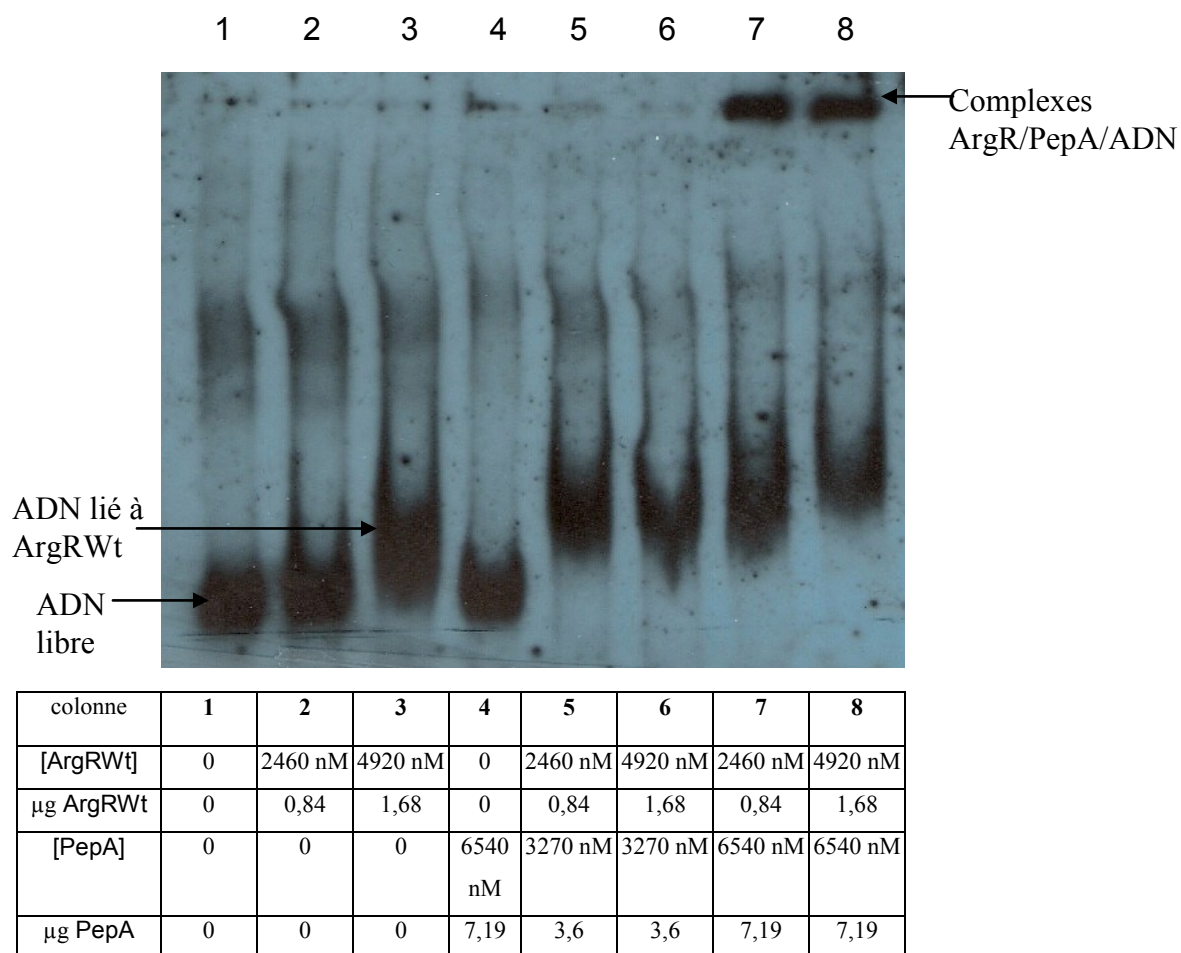
**Figure 71.** Retard sur gel d'ArgR149 avec deux séquences *cer*. Les concentrations des protéines ArgRWt et ArgR149 monomériques sont indiquées. Le pourcentage de pureté de la protéine ArgR149 est de 52 % et a été pris en compte dans les calculs de concentration. Dans chaque puits, 2 ng (0,25 nM) d'ADN et 1 µg de poly-dIdC ont été utilisés. Les migrations ont été effectuées à température ambiante et avec du tampon TBE 0,5 X dans un gel d'acrylamide 5 %.

Ces retards montrent que la forme ArgR149 de la protéine permet aussi le retard de la séquence d'ADN utilisée lors des expériences. Mais les retards qui ont été visualisés sont toujours plus diffus et moins intenses que ceux occasionnés par ArgRWt. Il a été visualisé que lorsque la concentration en ArgR149 augmentait, les bandes retardées étaient moins nettes (figure 71), voir plus étalées. Les profils de migration des deux séquences *cer* sont comparables à ce que l'on voit dans les publications de Tian *et al* en 1992 et 1994.



Il est important de signaler que des retards sur gel avec les protéines ArgRWt, PepA et les deux séquences *cer* avaient été tentés, mais lorsque les complexes se forment, ils ont une taille trop importante et restèrent au mieux dans le fond du puits (surtout pour les puits 7 et 8 de la figure 72), sans jamais rentrer dans le gel lors de migrations à 4°C.

De plus, les protéines PepA, ne semblent pas causer de retard sur gel des séquences d'ADN quand PepA n'est pas en présence d'ArgR.



**Figure 72.** Retard sur gel de complexes PepA-ArgRWt avec deux séquences *cer*. Les concentrations des protéines ArgRWt et PepA monomériques sont indiquées. Le pourcentage de pureté de la protéine ArgRWt est de 79 % et a été pris en compte dans les calculs de concentration, tout comme le pourcentage de pureté de PepA qui est de 60 %. Dans chaque puits, 2 ng d'ADN (0,25 nM) et 1 µg de poly-dIdC ont été utilisés et la migration c'est faite à 4°C. Les migrations ont été effectuées avec du tampon TBE 0,5 X dans un gel d'acrylamide 5 %.

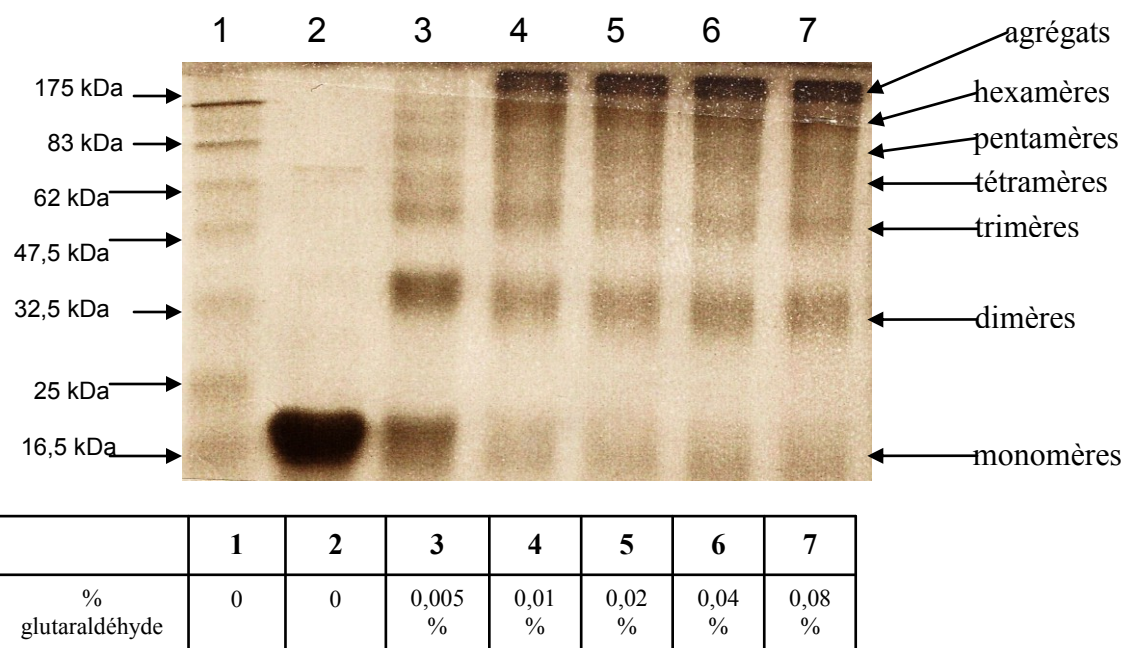
### III.4.Caractérisation des capacités d'interactions des protéines

#### III.4.1.*Crosslinkings* au glutaraldéhyde

##### III.4.1.1.*Crosslinkings* avec ArgR

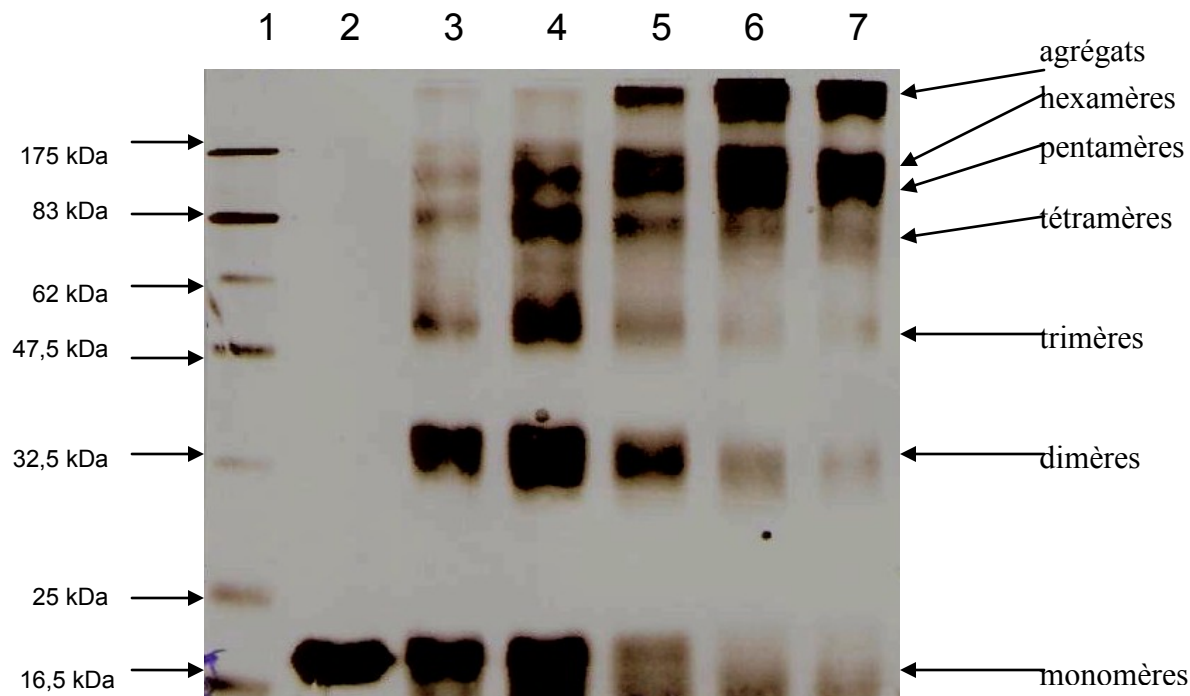
Selon la littérature, ArgRWt est capable de former des complexes hexamériques en présence de la L-arginine, il reste à savoir si la visualisation de ces complexes *in-vitro* est possible. Car si, comme on le suppose, l'hexamérisation des formes mutantes est difficile à obtenir, l'hexamère serait alors peu stable, ce qui pourrait expliquer les difficultés de recombinaison et liaison observées. Une technique utilisant le glutaraldéhyde à différents pourcentages, en solution permet de visualiser les multimères formés par ArgR et ses différentes formes. En effet, le glutaraldéhyde stabilise de manière covalente, par liaison chimique entre les groupements amines des protéines, les sous-unités qui interagissent les unes avec les autres. Ces expériences ont été faites avec les protéines ArgRWt, ArgR149, ArgR5aa (surexprimées à partir des vecteurs d'expression de type pQE31 donc possédant une partie 6-HIS) et ont été réalisées selon le protocole décrit dans matériel et méthodes et inspiré de Lu *et al*, 1992.

Les quantités de protéines ArgRWt, ArgR149, ArgR5aa sont de 5 µg dans chaque puits lors de ces expériences (soit 294,12 picomoles de monomères). Pour le *crosslinking* de la protéine ArgRWt, les profils suivants ont été obtenus. Les concentrations données ci-dessous ont été calculées grâce à biomath calculator, programme informatique accessible sur le site de Promega.



**Figure 73.** *Crosslinking* au glutaraldéhyde de la protéine ArgRWt. Dans le puits numéro 5  $\mu\text{g}$  de la protéine ArgRWt (soit 294,12 picomoles) a été observée en absence de glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la même protéine ArgRWt a été mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %.

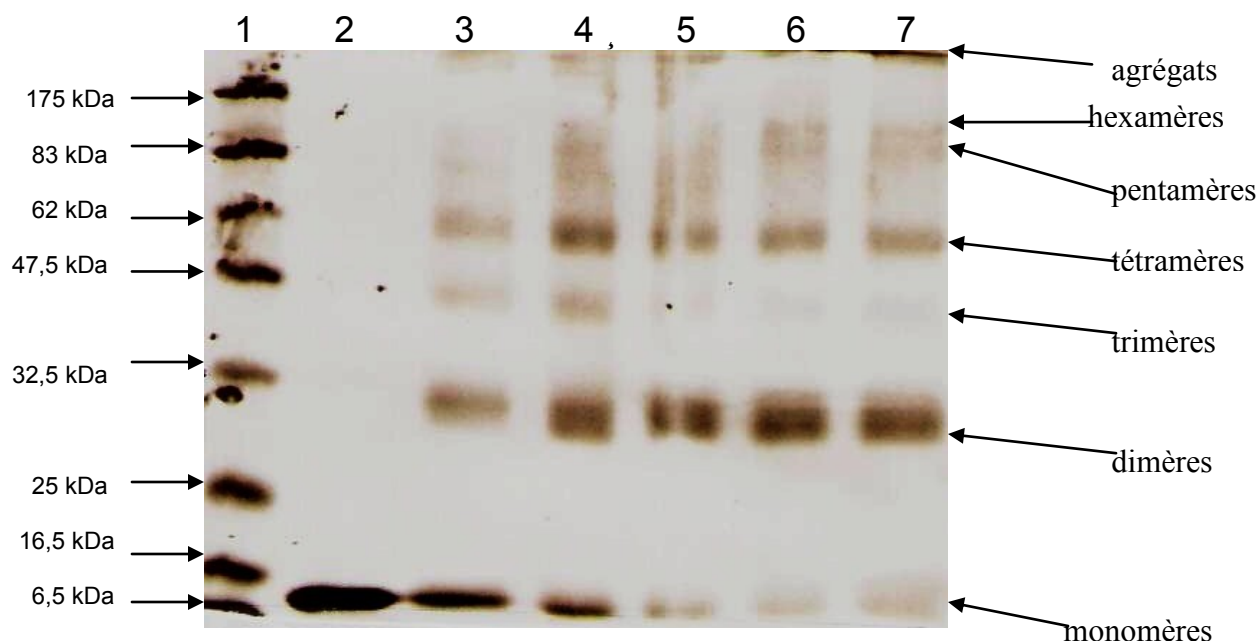
La protéine ArgRWt a été capable de se multimériser comme c'est le cas dans les cellules de la bactérie *E.coli*. ArgRWt forme des agrégats à partir de 0,01% de glutaraldéhyde. Les hexamères ont été plus fortement observés à partir d'une concentration de 0,01 % de glutaraldéhyde (figure 73).



	1	2	3	4	5	6	7
% glutaraldéhyde	0	0	0,005%	0,01%	0,02%	0,04%	0,08%

**Figure 74.** *Crosslinking* au glutaraldéhyde ArgR5aa. Dans le puits numéro 2, la protéine ArgR5aa (294,12 picomoles) a été observée en absence de glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la protéine ArgR5aa a été mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %. 294,12 picomoles de monomères d'ArgR ont été utilisés.

La protéine ArgR5aa a été capable de se multimériser elle aussi. La multimérisation d'ArgR5aa a été observée pour les concentrations en glutaraldéhyde de 0,005% à 0,08%, la multimérisation d'ArgR5aa ressemble assez à celle d'ArgRWt. D'ailleurs, les hexamères ont commencé à être aperçus à partir d'une concentration de 0,02 % de glutaraldéhyde. Toutefois, les agrégats ont été moins rapidement obtenus (figure 74).



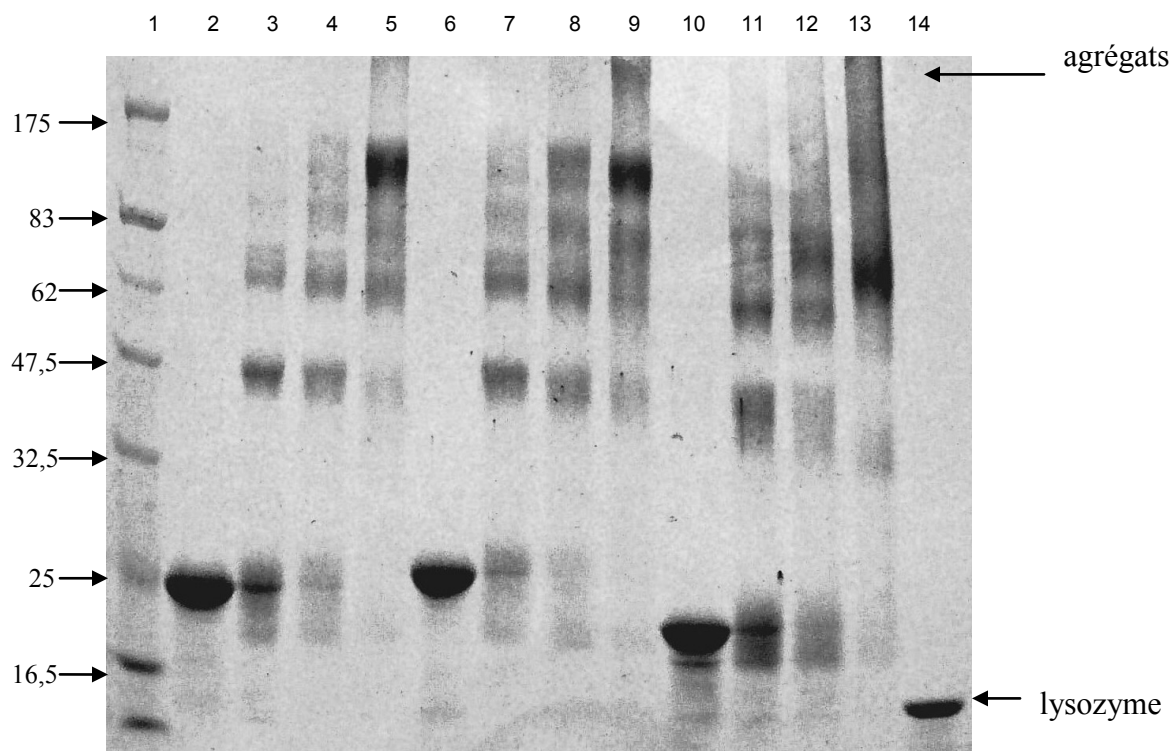
	1	2	3	4	5	6	7
% glutaraldéhyde	0	0	0,005%	0,01%	0,02%	0,04%	0,08%

**Figure 75.** *Crosslinking* au glutaraldéhyde d'ArgR149. La concentration en ArgR149 de 294,12 picomoles est égale entre les puits. Dans le puits numéro 2, la protéine ArgR149 a été observée en absence de glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la protéine ArgR149 a été mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %.

La protéine ArgR149 a été capable de se multimériser. En effet pour des concentrations en glutaraldéhyde de 0,005% à 0,08%, des dimères ont été souvent visualisés. Mais des hexamères ont été également vus seulement à partir du puits numéro 6, bien que plus difficilement. Ce sont surtout des dimères qui ont été observés (des puits 4 à 7). Les hexamères sont bien visibles à partir de 0,04 % de glutaraldéhyde. De plus, peu d'agrégats ont été remarqués (figure 75).

Lorsque l'on augmente le pourcentage de glutaraldéhyde de 0,005% à 0,08% on a toujours formation de multimères avec ces formes d'ArgR. Mais si les hexamères sont rapidement visualisés avec les formes ArgRWt et ArgR5aa, les hexamères sont

plus difficiles à faire apparaître pour la forme ArgR149, pour laquelle plus de dimères et de tétramères ont été visualisés, avec une prédominance de dimères.

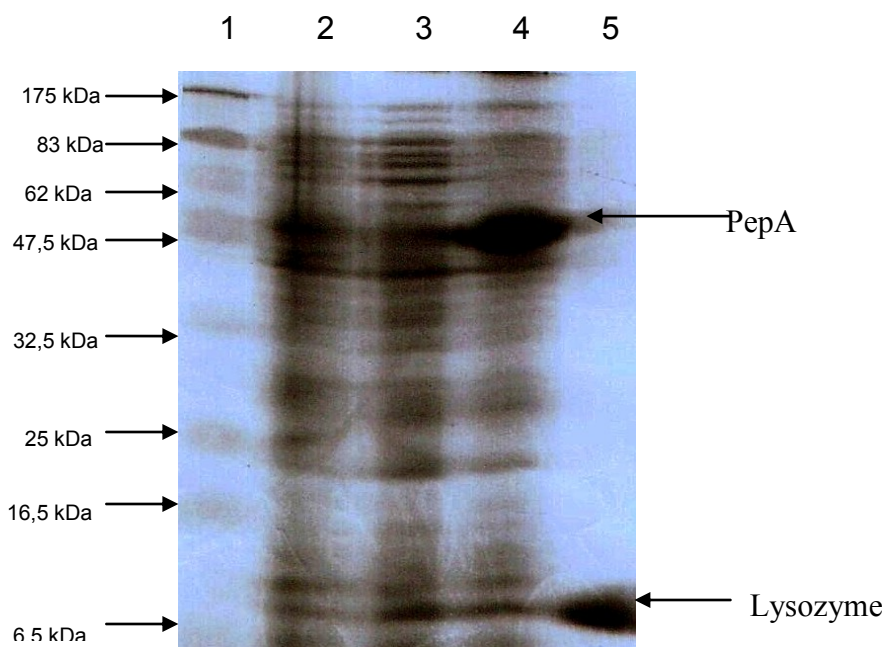


**Figure 76.** *Crosslinking* au glutaraldéhyde sur un gel SDS-PAGE à gradient des protéines ArgRWt, ArgR5aa et ArgR149. Les puits 2 à 5 montrent ArgRWt en présence de respectivement 0 %, 0,005 %, 0,01 % et 0,08 % de glutaraldéhyde. Les puits 6 à 9 montrent ArgR5aa en présence de respectivement 0 %, 0,005 %, 0,01 % et 0,08 % de glutaraldéhyde. Les puits 10 à 13 montrent ArgR149 en présence de respectivement 0%, 0,005 %, 0,01 % et 0,08 % de glutaraldéhyde. Le puits numéro 1 correspond au marqueur de poids moléculaire dont les poids sont indiqués en kDa. 294,12 picomoles de protéines ont été utilisées pour chaque version d'ArgR.

Le gel de la figure 76 montre les trois versions précédemment étudiées sur un même gel. On peut voir les légères différences de poids moléculaires entre ces versions et les multimérisations de 49,02 picomoles de chacune des versions d'ArgR étudiées, à savoir respectivement, ArgRWt, ArgR5aa et ArgR149. On peut également voir qu'il y a des différences de multimérisation entre chaque version, déjà visualisées au niveau des figures 73, 74 et 75.

### III.4.1.2. *Crosslinkings* avec PepA

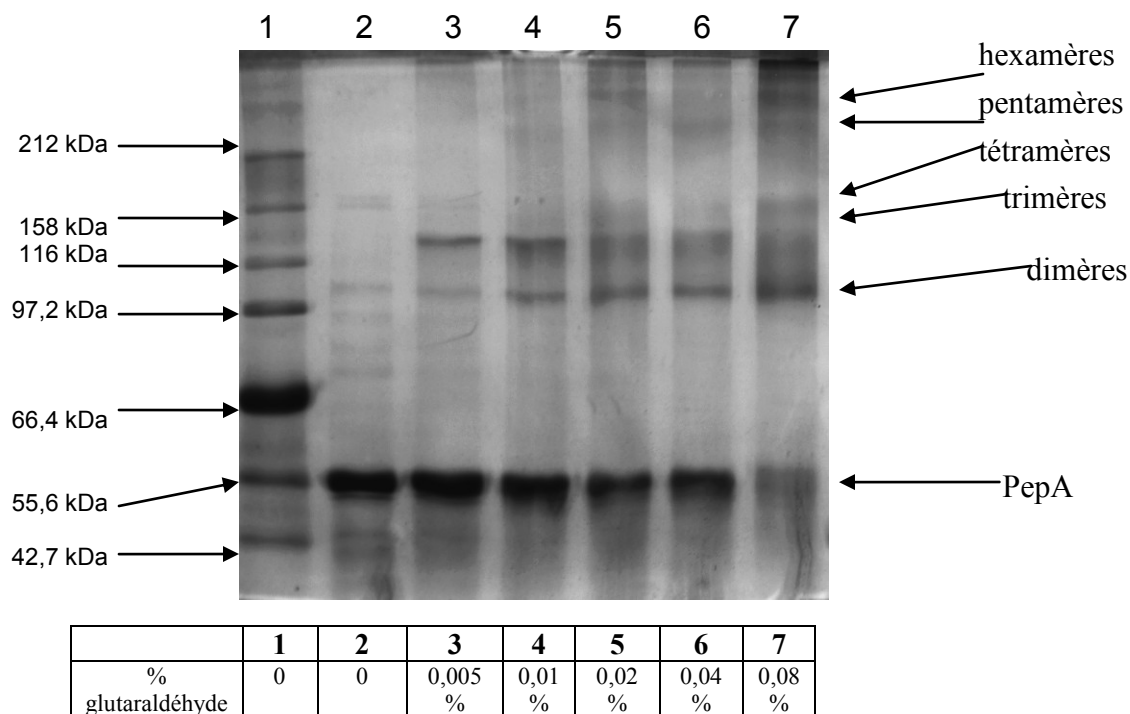
La purification partielle de PepA a été réalisée avant les expériences de *crosslinking* (figure 77).



**Figure 77.** Purification partielle de PepA. Le premier puits montre le marqueur prestained. Le deuxième puits montre le surnageant après sonication et précipitation au sulfate d'ammonium (29,1 %) avant dialyse contre le TM. Le troisième puits montre le surnageant après cette dialyse. Le quatrième puits montre le culot resuspendu après dialyse où on peut voir la protéine PepA. Le cinquième puits montre le témoin lysozyme.

Un *crosslinking* au glutaraldéhyde a également été réalisé pour PepA, en poursuivant la même idée que pour ArgR, à savoir, vérifier si la protéine PepA est capable de se multimériser *in-vitro*. 10,7  $\mu\text{g}$  de PepA ont été utilisés lors de cette expérience (soit 193,49 picomoles de monomères), cette concentration prend en compte le pourcentage de pureté de 60 %.

Les concentrations données ci-dessous pour PepA ont été calculées grâce à biomath calculator, programme informatique accessible sur le site de Promega.



**Figure 78.** *Crosslinking* au glutaraldéhyde de la protéine PepA. Le puits numéro 2, contient PepA sans glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la protéine PepA a été mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %. 193,49 picomoles de monomères ont été utilisés

Les formes multimériques de PepA ont été détectées en présence de concentration croissante de glutaraldéhyde, notamment des hexamères. Des trimères ont été visualisés à partir d'une concentration de 0,005% de glutaraldéhyde. Les hexamères ont été visualisés, quant à eux, à partir d'une concentration de 0,02% en glutaraldéhyde. Il a été remarqué que dans cette expérience, les trimères de PepA ont été très facilement obtenus. L'hexamérisation de PepA passerait d'abord par une trimérisation des sous-unités (figure 78). Il est à noter que peu d'agrégats de la protéine PepA version sauvage ont été visualisés durant cette expérience de *crosslinking*, même à des concentrations de glutaraldéhyde comprises entre 0,02% et 0,08%.



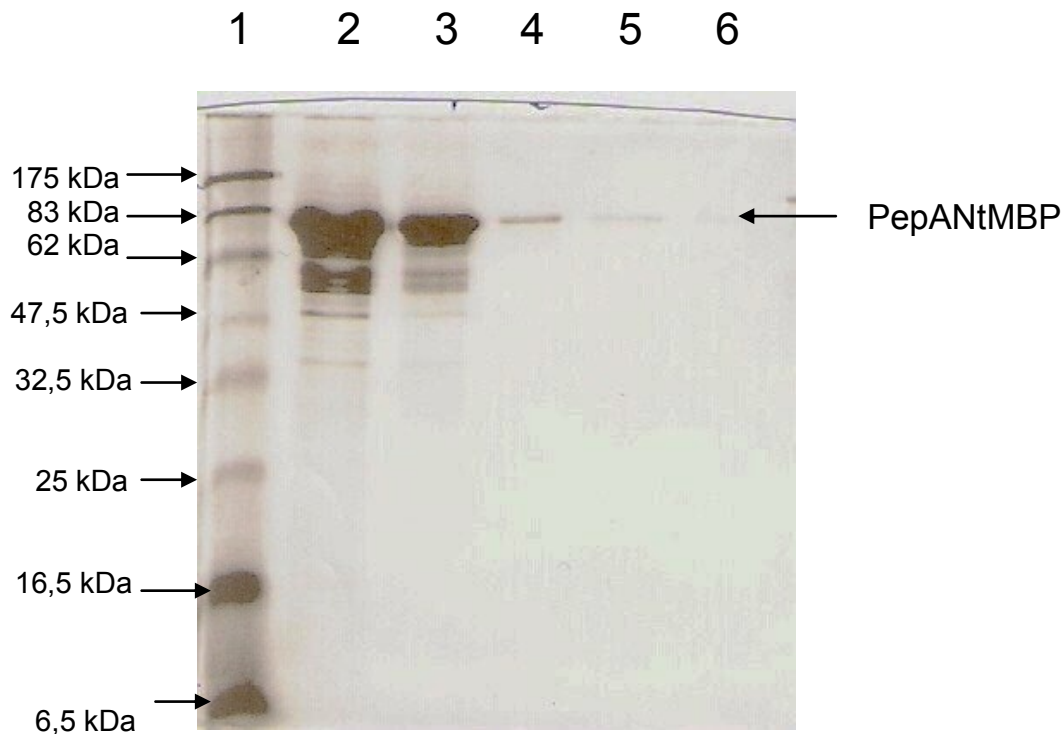
### III.4.1.3. *Crosslinkings* avec les versions tronquées de PepA

Pour les versions tronquées de la protéine PepA obtenues par fusion, soit de la partie Nt, contenant les résidus 1 à 155 de PepA fusionnés au MBP (*Maltose Binding Protein*) à sa partie Nt, soit de la partie Ct, contenant les résidus 156 à 288 de PepA fusionnés au MBP à sa partie Nt, à savoir les protéines PepANtMBP et PepACtMBP, des *crosslinking* au glutaraldéhyde ont aussi été réalisés, dans le but de voir si la protéine PepA tronquée (que ce soit en partie Nt ou Ct) était toujours capable de se multimériser sous l'effet du glutaraldéhyde.

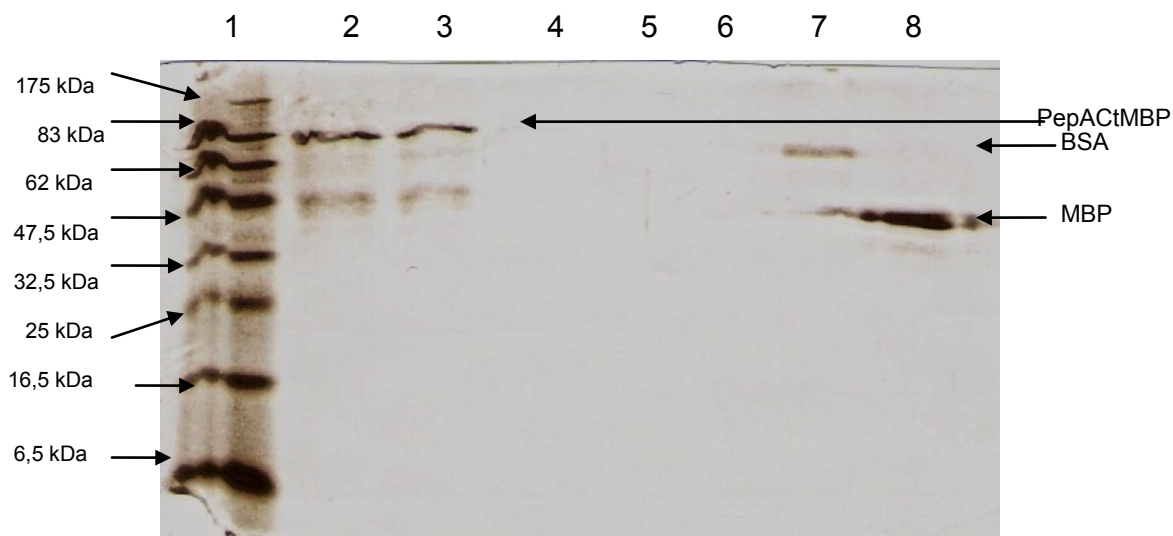
Cette multimérisation, comme avec ArgRWt, a été tentée avec l'utilisation du glutaraldéhyde quand son pourcentage en solution a été augmenté de 0,005% à 0,08% pour visualiser de plus en plus de multimères de haut poids moléculaire.

Pour les versions tronquées de PepA, PepANtMBP et PepACtMBP, 10,7 µg de protéines ont été utilisés durant ces expériences (soit respectivement 178,33 picomoles et 133,75 picomoles de monomères). Les concentrations données ont également été calculées grâce à biomath calculator, via le site de Promega.

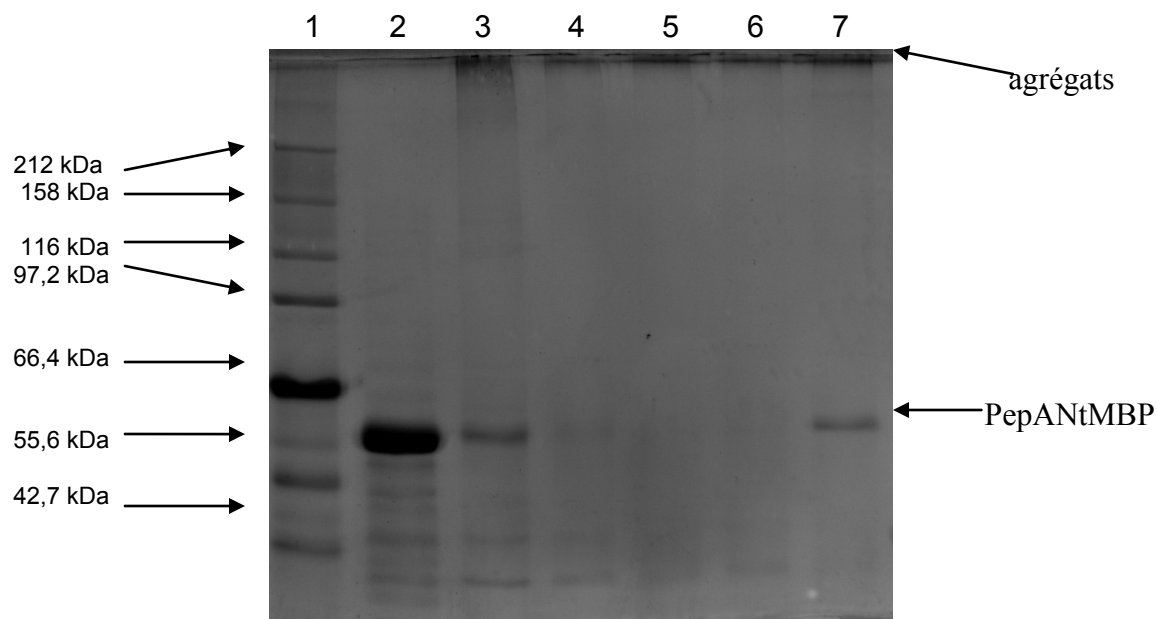
Il est à préciser que le poids moléculaire de PepANtMBP est 60 kDa et celui de PepACtMBP est de 80 kDa.



**Figure 79.** Purification de la protéine PepANtMBP. Le second puits montre la première élution de la protéine PepANtMBP. Les puits 3 à 6 montrent respectivement les élutions 2, 3, 4 et 5. La pureté de l'élution est de 70 % et la migration a été réalisée sur un gel SDS-PAGE à 15 % d'acrylamide.



**Figure 80.** Purification de la protéine PepActMBP. Les puits 2 à 6 montrent respectivement les élutions 1, 2, 3, 4 et 5. Les puits numéro 7 et 8 montrent respectivement un témoin BSA et un témoin MBP. La pureté de l'élution est de 40 % et la migration a été réalisée sur un gel SDS-PAGE à 15 % d'acrylamide.

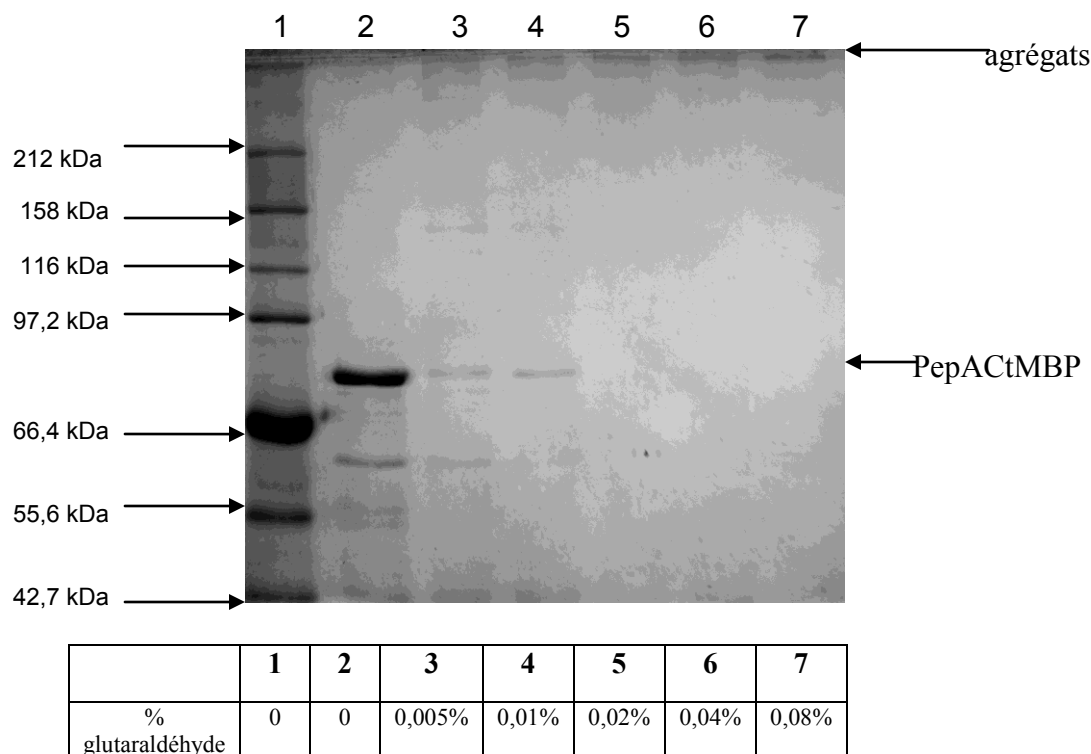


	1	2	3	4	5	6	7
% glutaraldéhyde	0	0	0,005%	0,01%	0,02%	0,04%	0,08%

**Figure 81.** *Crosslinking* au glutaraldéhyde de PepANtMBP. Le puits numéro 2 présente PepANtMBP sans glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la protéine PepANtMBP a été mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %. 178,33 picomoles de monomères ont été utilisés.

Des multimères de PepANtMBP n'ont visiblement pas été formés lors de ces expériences. Mais des agrégats ont été visualisés dès 0,005% de glutaraldéhyde (figure 79).

Les *crosslinking* au glutaraldéhyde pour PepACtMBP sont présentés sur la figure suivante.



**Figure 82.** *Crosslinking* au glutaraldéhyde de PepACtMBP. Le puits numéro 2, présente PepACtMBP sans glutaraldéhyde. Dans les puits 3, 4, 5, 6 et 7, la protéine PepACtMBP est mise en présence respectivement de 0,005%, 0,01%, 0,02%, 0,04% et 0,08% de glutaraldéhyde. La concentration du gel en acrylamide est de 15 %. 133,75 picomoles de monomères ont été utilisés.

Des multimères de PepACtMBP n'ont visiblement pas été formés lors de ces expériences. Mais des agrégats ont été visualisés dès 0,005% de glutaraldéhyde, voir un possible dimère, comme pour PepANtMBP (figure 80).

Une autre technique avait été testée pour PepA et les protéines tronquées PepANtMBP et PepACtMBP. Elle consistait à vérifier l'activité aminopeptidase de chacune des versions de PepA. En effet, lorsque l'on met la protéine PepA en présence de L-leucine p-nitroanilide hydrochloride et que l'on incube la solution à 37°C, on va obtenir une solution de couleur jaune par dégradation du substrat.

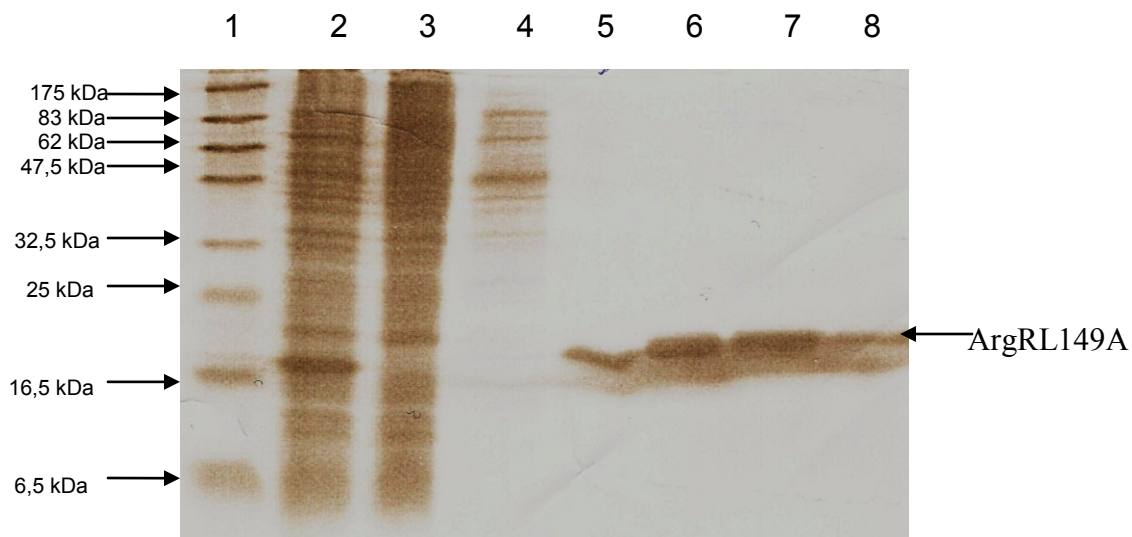
Seule la forme sauvage de PepA s'est avérée capable d'agir normalement sur le L-leucine p-nitroanilide hydrochloride, substrat de cette expérience. Les deux autres

versions ont été incapables de générer cette couleur jaune, même quand PepANtMBP et PepACtMBP ont été mélangées dans la même solution, montrant leur absence d'activité aminopeptidase.

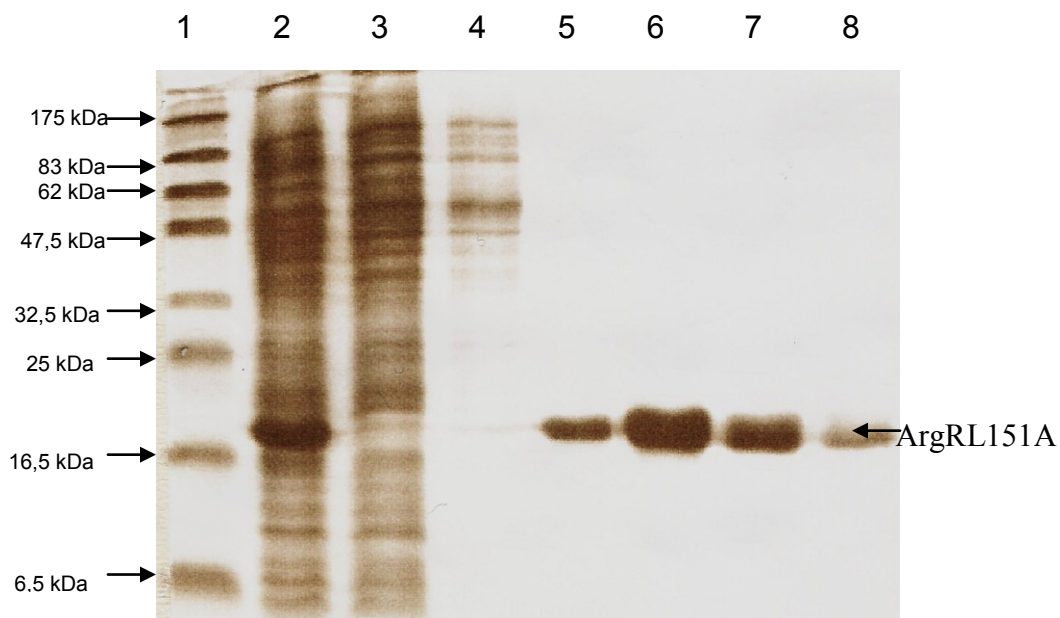
### III. 4.2. *Crosslinking* réversibles avec DSP

Les expériences de *crosslinkings* réversibles sont menées avec un « *crosslinker* » réversible nommée DSP qui stabilise les interactions protéine à protéine en formant des liaisons covalentes entre les dites protéines. On a récupéré les complexes formés en présence de DSP avec des billes magnétiques Ni-NTA reconnaissant la partie 6-HIS des protéines ArgR que l'on a utilisé dans ces expériences pour connaître les interactions entre les protéines PepA et ArgR. Après l'ajout du DTT ou du  $\beta$ -mercaptoéthanol, les liaisons croisées formées par le DSP se brisent libérant ArgR et les protéines liées à ce dernier. Les versions suivantes d'ArgR fusionnées à 6-HIS, ont été utilisées pour ces expériences : ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A, et ce respectivement dans les puits 2, 3, 4, 5 et 6 des deux figures 67 et 68. Pour toutes ces versions d'ArgR, on a utilisé 10  $\mu$ g de protéine (588,24 picomoles de monomères). Seules ces versions des protéines ArgR ont été utilisées pour interagir avec 30  $\mu$ g de PepA (542,5 picomoles de monomères) avec une pureté de 60 %.

Les protéines ArgRL149A et ArgRL151A ont été préalablement purifiées et éluées avec une solution à 150 mM d'imidazole (figure 81 et 82).

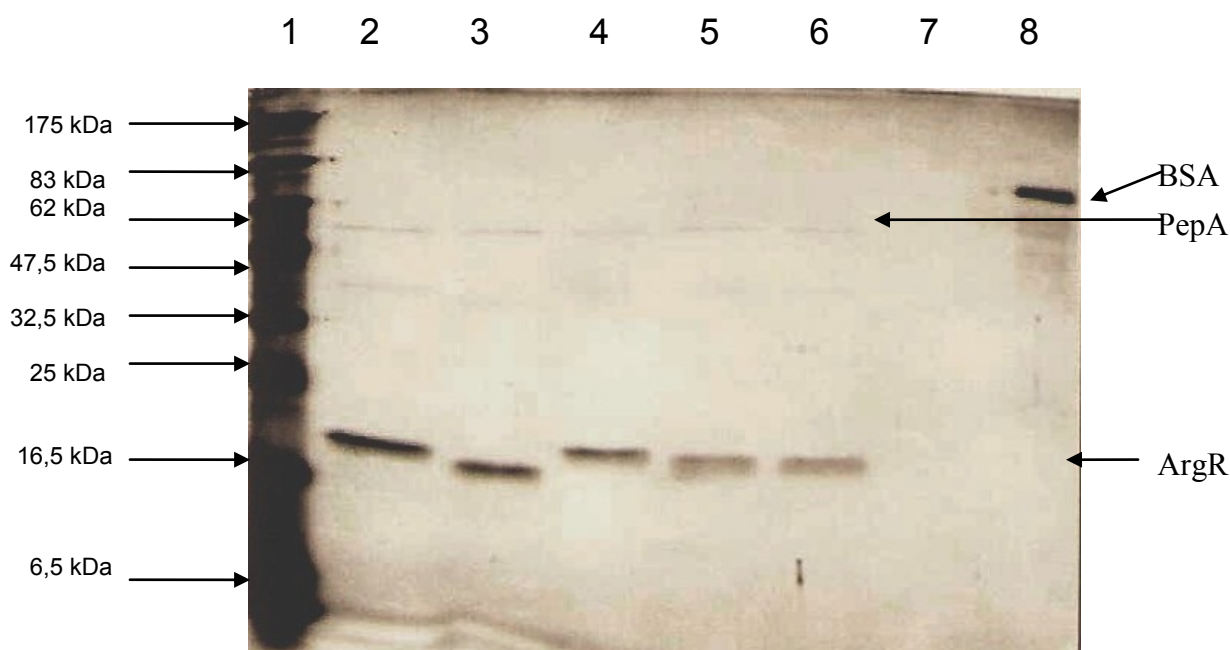


**Figure 83.** Purification de la protéine ArgRL149A. Le premier puits montre le marqueur. Le second montre un extrait du surnageant de sonication. Le troisième puits montre le *flow through*. Le quatrième puits montre le lavage de la résine Ni-NTA. Les puits 5 à 8 montrent respectivement les éluions 1 à 4 de la protéine ArgRL149A avec 150 mM d'imidazole. La concentration du gel en acrylamide est de 15 %.



**Figure 84.** Purification de la protéine ArgRL151A. Le premier puits montre le marqueur. Le second montre un extrait du surnageant de sonication. Le troisième puits montre le *flow through*. Le quatrième puits montre le lavage de la résine Ni-NTA. Les puits 5 à 8 montrent respectivement les éluions 1 à 4 de la protéine ArgRL151A avec 150 mM d'imidazole. La concentration du gel en acrylamide est de 15 %.

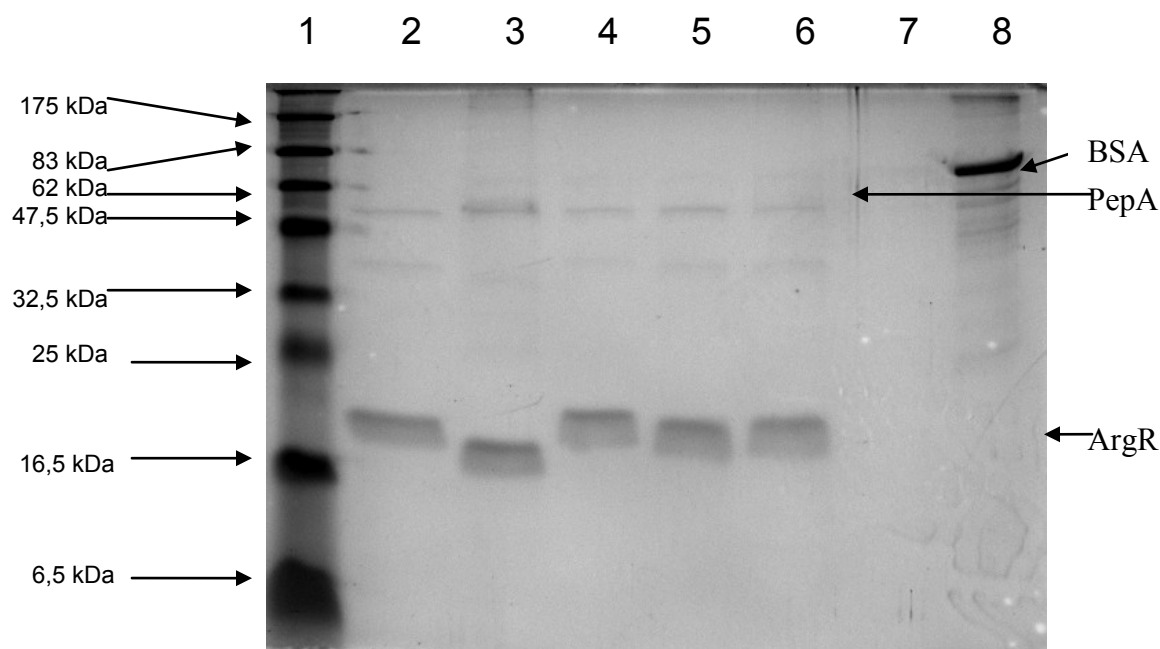
Des expériences en présence (figure 83) et en absence de L-arginine (figure 84) ont été réalisées, mais les profils sont semblables dans ces deux expériences. À noter que le puits numéro 7 de chaque figure est un contrôle où seul PepA a été mis en présence de DSP et des billes magnétiques Ni-NTA. Le puits numéro 8 contient un témoin BSA, car les mélanges contiennent tous du BSA, on cherche à voir la spécificité des billes magnétiques Ni-NTA vis-à-vis de ce dernier. Les concentrations données pour les différentes versions d'ArgR et la protéine PepA ont été calculées grâce à biomath calculator.



	1	2	3	4	5	6	7	8
PepA	-	+	+	+	+	+	+	-
ArgRWt	-	+	-	-	-	-	-	-
ArgR149	-	-	+	-	-	-	-	-
ArgR5aa	-	-	-	+	-	-	-	-
ArgRL149A	-	-	-	-	+	-	-	-
ArgRL151A	-	-	-	-	-	+	-	-
BSA	-	-	-	-	-	-	-	+

**Figure 85.** Migration sur gel SDS-PAGE 15 % des différentes formes d'ArgR avec L-arginine. Chaque forme d'ArgR testée a été mise en présence de PepA. Le premier puits montre le marqueur prestained. Les puits 2 à 6 montrent des mélanges de PepA avec respectivement ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A. Le puits numéro 7 montre un contrôle des billes magnétiques avec PepA. Le puits numéro 8 montre un contrôle BSA sans utilisation des billes magnétiques. 588,24 picomoles de monomères des différentes protéines ArgR et 542,5 picomoles de monomères de PepA ont été utilisés.

Cette photo est très similaire à celle qui va suivre où une expérience similaire sans L-arginine a été réalisée avec les mêmes protéines.



	1	2	3	4	5	6	7	8
PepA	-	+	+	+	+	+	+	-
ArgRWt	-	+	-	-	-	-	-	-
ArgR149	-	-	+	-	-	-	-	-
ArgR5aa	-	-	-	+	-	-	-	-
ArgRL149A	-	-	-	-	+	-	-	-
ArgRL151A	-	-	-	-	-	+	-	-
BSA	-	-	-	-	-	-	-	+

**Figure 86.** Migration sur gel SDS-PAGE 15 % des différentes formes d'ArgR sans L-arginine. Chaque forme d'ArgR testée a été mise en présence de PepA. Le premier puits montre le marqueur prestained. Les puits 2 à 6 montrent des mélanges de PepA avec respectivement ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A. Le puits numéro 7 montre un contrôle des billes magnétiques avec PepA. Le puits numéro 8 montre un contrôle BSA sans utilisation des billes magnétiques. 588,24 picomoles de monomères des différentes protéines ArgR et 542,5 picomoles de monomères de PepA ont été utilisés.

Toutes les versions d'ArgR ont été utilisées avec PepA sauvage lors de ces expériences, toutes les versions peuvent se lier à la protéine PepA sauvage. En effet PepA a été recueillie grâce à ArgR dans chaque cas, avec ou sans L-arginine. Les puits



numéro 7 de chaque figure montrent qu'il n'y a pas de liaison non-spécifique de PepA avec les billes magnétiques. La grande similarité entre les deux expériences malgré l'absence de la L-arginine dans la seconde a été remarquée. Il a été visualisé sur les deux figures précédentes que le contrôle avec PepA montre qu'il n'y a pas de problème de liaison non-spécifique avec les billes magnétiques Ni-NTA. En effet, PepA seule ne se lie pas aux billes magnétiques sans ArgR, il en est de même avec le BSA.

Il a été envisagé de tester les éventuelles capacités des protéines PepANtMBP et PepACtMBP à se lier aux formes suivantes d'ArgR, ArgRWt (fusionnée 6 HIS ou non), ArgR149, ArgR5aa, ArgRL149A, ArgRL151A, via la technique employant les billes d'amylose magnétiques.

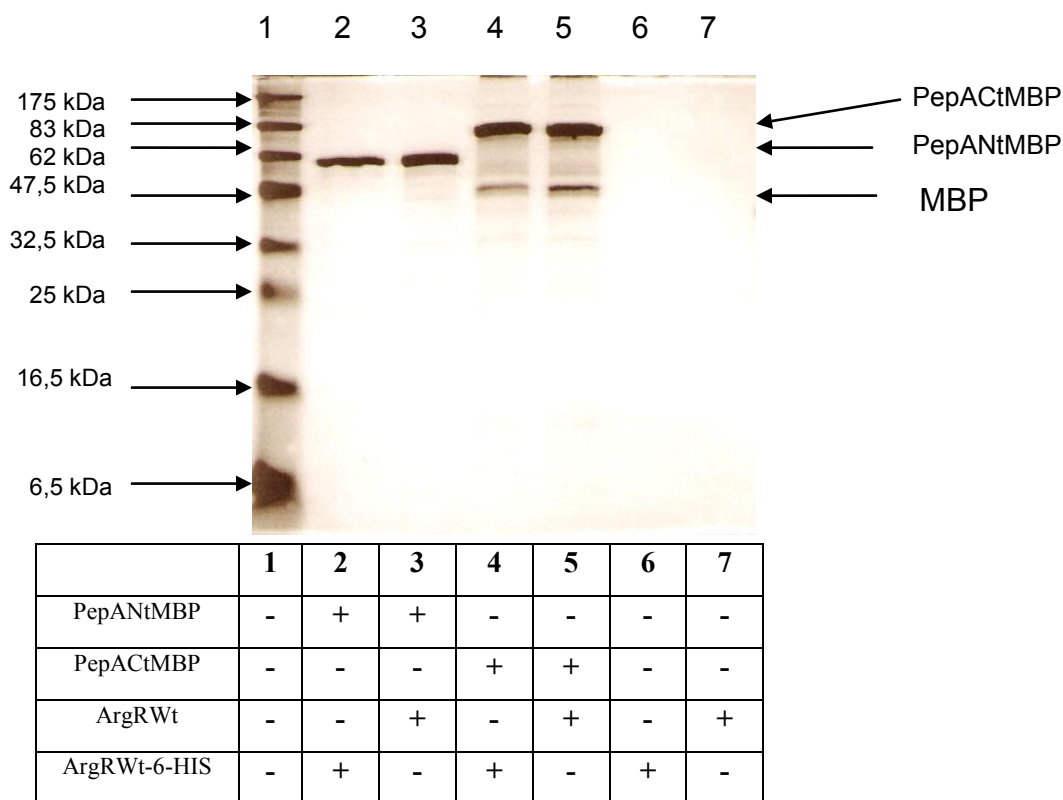
Durant les expériences d'interaction protéine à protéine, les formes tronquées PepANt et PepACt ont été mises en présence de la forme sauvage d'ArgR, et ce pour voir si en présence de cette forme d'ArgR, des complexes ArgR-PepANt ou des complexes ArgR-PepACt allaient se former, montrant le rôle d'une des régions dans la liaison ArgR-PepA et donc dans la recombinaison en *cer*.

Lors de ces expériences réalisées avec les protéines PepANtMBP, PepACtMBP et ArgRWt (avec 6-HIS ou non), où on utilise des billes magnétiques à l'amylose, qui sont capables de capter la partie MBP des protéines PepA tronquées. L'utilisation de ces billes magnétiques à l'amylose a été décidée car des protéines PepA fusionnées à 6-HIS fonctionnelles n'ont pas encore été obtenues, ni purifiées malgré le fait que d'après le séquençage, le marquage de la protéine a été réussi. De plus, le choix de travailler sur des versions tronquées de cette protéine est également dû au fait que la plupart des mutations ponctuelles déjà obtenues avec PepA, surtout en Nt et peu en Ct, n'interviennent que dans la liaison entre l'ADN et PepA, et non entre ArgR et PepA.

Les profils suivants ont pu être observés. Les quantités de protéines sont les suivantes, 15 µg de protéine PepANtMBP (250 picomoles de monomères) et 15 µg de PepACtMBP (187,5 picomoles de monomères), les différentes protéines ArgR ont une

quantité de 5  $\mu\text{g}$  (294,12 picomoles de monomères) dans les puits où elles sont présentes. Les résultats sont présents sur la figure suivante.

Les concentrations données pour les différentes versions de PepA et la protéine ArgRWt ont été calculées grâce à biomath calculator.



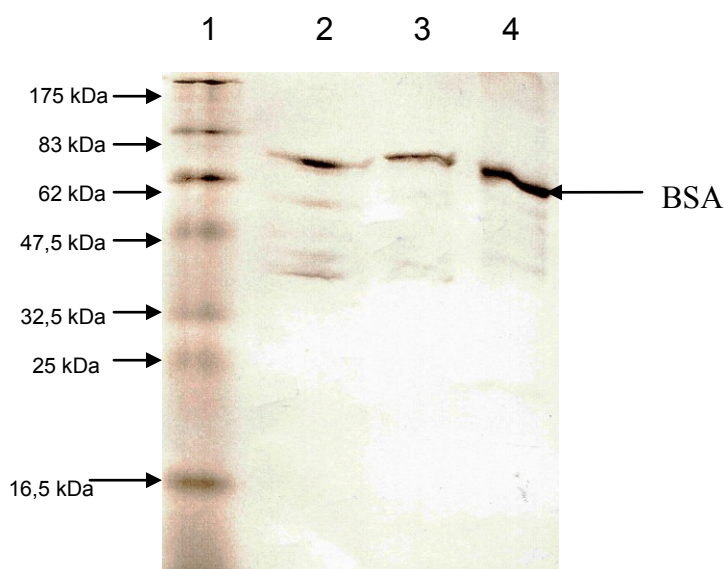
**Figure 87.** Migration des protéines PepA tronquées lors des interactions protéine à protéine entre ArgR et les formes tronquées de PepA. La concentration du gel en acrylamide est de 15 %. 250 picomoles de monomères de PepANtMBP, 187,5 picomoles de monomères de PepACtMBP et 294,12 picomoles de monomères d'ArgR natives ou fusionnées avec 6-HIS ont été utilisés.

Dans tous les cas de figure, aucune formation de complexe entre ArgRWt (ni avec d'autres formes, données non-montrées), que ce soit avec la forme fusionnée avec 6-HIS comme le montre le puits 2, ou la forme sauvage comme dans le puits 3, et les deux formes de PepA fusionnées au MBP, n'a été vue, il n'y a pas d'interaction. En effet, seules les versions tronquées de PepA (PepACtMBP et PepANtMBP) ont été captées (figure 85). C'est à la vue des résultats avec ArgRWt qu'il a été décidé de ne pas poursuivre les expériences avec les formes ArgR149, ArgR5aa, ArgRL149A et

ArgRL151A, car on s'attend à des résultats similaires à ceux des formes ArgRWt-6-HIS et ArgRWt.

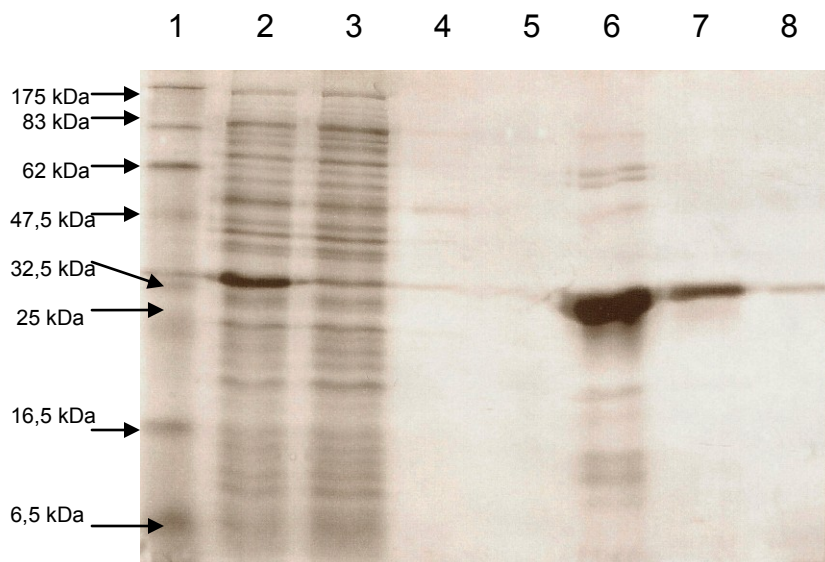
### III.4.3. Expériences de recombinaison *in vitro*

Ces expériences ont été réalisées avec le plasmide pCS210 contenant deux séquences *cer*, en présence des protéines XerC, XerD, PepA et des différentes versions d'ArgR (ArgRWt, ArgR149, ArgR5aa). Pour ces expériences, la purification de la protéine XerC et de la protéine XerD a été réalisée (figure 85).



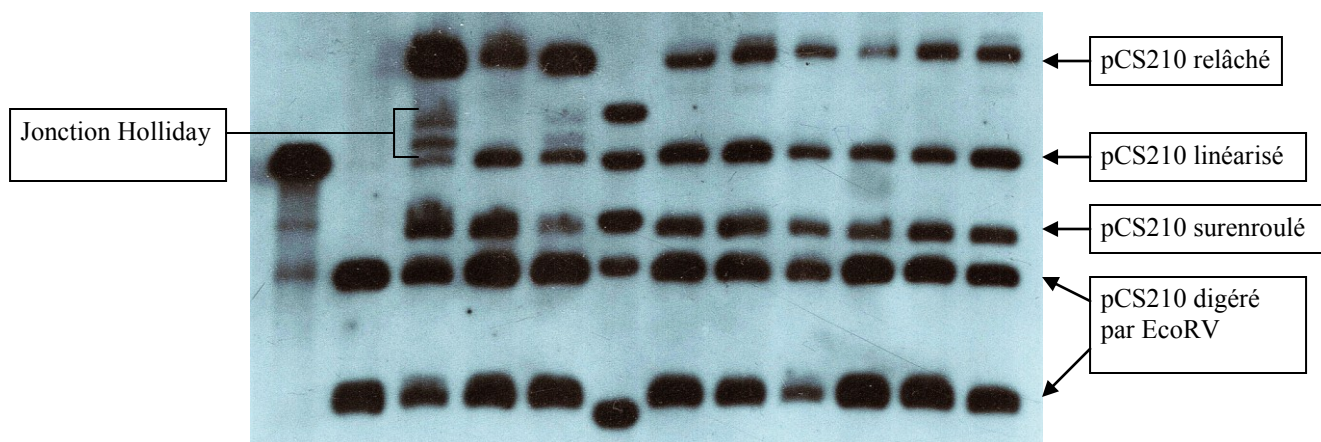
**Figure 88.** Purification de XerC. Le premier puits montre le marqueur. Le second puits montre la première élution de la protéine XerC. Le troisième puits montre la deuxième élution de la protéine XerC. Le quatrième puits montre un témoin BSA.

Cette purification de XerC (figure 86) a été permise par fusion avec le MBP (*Maltose Binding Proteine*). La purification suivante, celle de XerD (figure 87), a été permise par fusion avec 6-HIS.



**Figure 89.** Purification de XerD. Le premier puits montre le marqueur. Le second puits montre la migration d'un échantillon du surnageant avant passage dans la résine. Le troisième puits montre les protéines qui ne se sont pas liées à la résine après passage. Le puits numéro 4 montre le lavage de la colonne. Les colonnes 5 à 8 montrent les éluions respectivement 1 à 4 en utilisant 250 mM d'imidazole.

Après incubation des protéines avec 400 ng d'ADN plasmidique, une approche par Southern-blot d'un gel d'agarose 0,75 % a été choisie pour détecter les jonctions Holliday, un produit de la recombinaison. Cette visualisation est possible grâce aux sondes d'ADN marquées avec DIG et contenant une séquence *cer* s'hybridant avec les plasmides pCS210 lors du Southern blot (d'après les conditions utilisées dans M.Villion and G.Szatmari, 2003 et Colloms *et al*, 1996).



	1	2	3	4	5	6	7	8	9	10	11	12
pCS210	+	+	+	+	+	-	+	+	+	+	+	+
ArgRWt	-	-	+	-	-	-	+	-	-	-	-	-
ArgR149	-	-	-	+	-	-	-	+	-	-	-	-
ArgR5aa	-	-	-	-	+	-	-	-	+	-	-	-
PepA	-	-	+	+	+	-	-	-	-	+	-	-
XerC	-	-	+	+	+	-	-	-	-	-	+	-
XerD	-	-	+	+	+	-	-	-	-	-	-	+

**Figure 90.** Migration *Southern-blot* des différentes recombinaisons *in vitro*. Dans le premier puits, le plasmide pCS210 linéarisé a été visualisé, les puits 2 à 5 montrent des pCS210 digérés par EcoRV (deux coupures) et les puits 3 à 5 montrent différentes formes du plasmide pCS210 en présence de PepA, XerC et XerD et respectivement des formes d'ArgR suivantes, ArgRWt, ArgR149, ArgR5aa. Les puits 7 à 12 montrent des contrôles avec respectivement les protéines ArgRWt, ArgR149, ArgR5aa, PepA, XerC et XerD. Les concentrations sont les suivantes, [ArgRWt]=180 nM de monomères (0,06 µg), [ArgR149]=180 nM de monomères (0,06 µg), [ArgR5aa]=180 nM de monomères (0,06 µg), [PepA]=1800 nM de monomères (1,98 µg), [XerC]=300 nM monomère (0,48 µg) et [XerD]=300 nM monomère (0,24 µg).

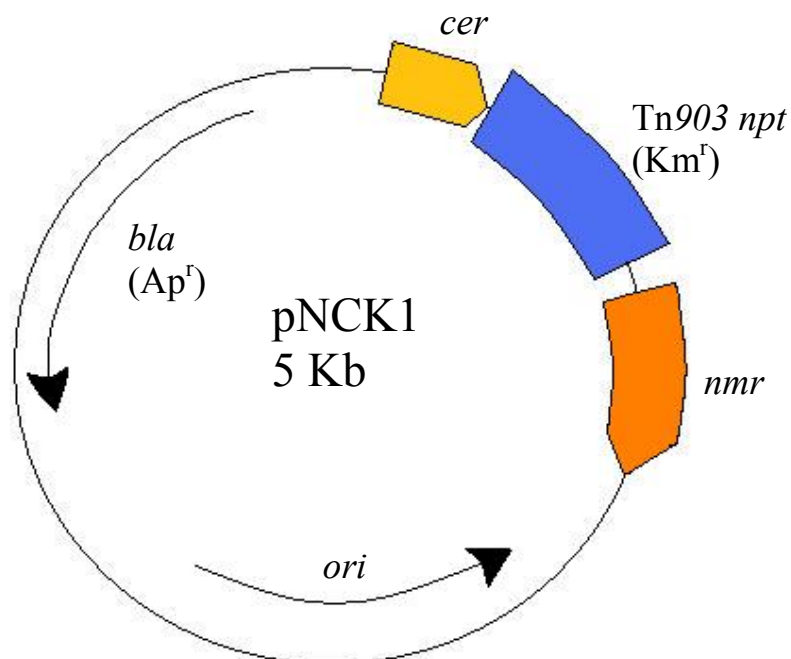
En présence de la protéine ArgRWt, les jonctions Holliday sont bien visibles (puits 3 de la figure 87), mais il n'y en a pas avec ArgR149 (puits 4 de la figure 88). L'apparition des jonctions d'Holliday avec ArgR5aa, a été visualisée d'une manière plus faible (puits 5 de la figure 87). Dans le puits 2, qui est un témoin avec seulement le plasmide pCS210, seules des bandes correspondant à pCS210 ont été visualisées après coupure par EcoRV. Fait important, les protéines XerC et XerD ont été ajoutées dans les mélanges qui ont été fait migrés dans les puits 3 à 5. Ce qui permet d'avoir toutes les protéines impliquées dans la recombinaison de présentes. Dans les puits 7 à 12 où

sont situés des contrôles avec seulement une seule protéine, il n'y a jamais de jonction d'Holliday qui se forment. Ces résultats suggèrent qu'effectivement, ArgR149 est incapable de former un complexe de recombinaison en présence PepA, XerC et XerD, et qu'ArgR5aa *in vitro* voit sa capacité à former ce complexe diminué par rapport au type sauvage.

#### **III.4.4. Expériences de topologie de l'ADN**

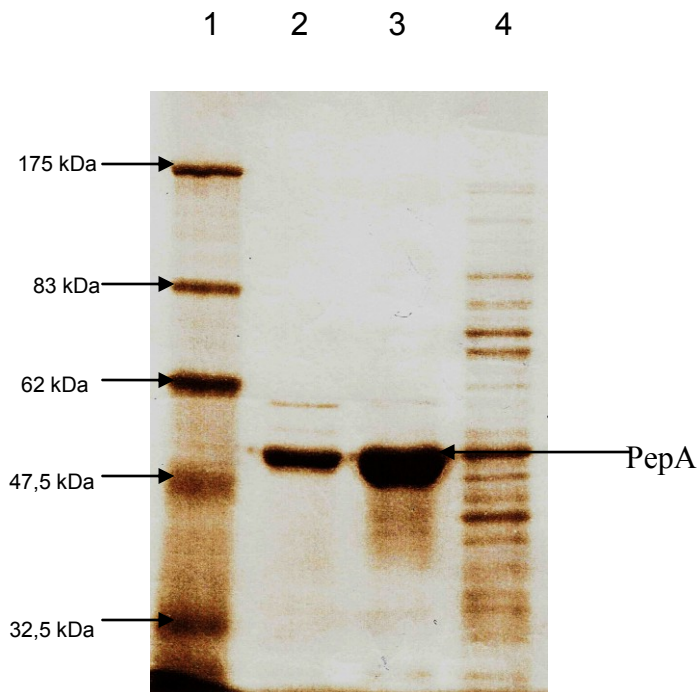
Ces expériences ont été réalisées avec l'utilisation des plasmides pNCK1 (Sirois and Szatmari, 1995; Zakova and Szatmari, 1995) possédant une taille d'environ 5 kbs et deux séquences de recombinaison (*cer* et *nmr*) susceptibles de se lier à un complexe ArgR-PepA et donc d'entraîner des différences de topologie.

Ce plasmide (figure 89) est capable d'être recombiné *in vivo* par une souche d'*E.coli* possédant les gènes *argR*, *pepA*, *xerC* et *xerD*.



**Figure 91.** Carte du plasmide pNCK1. Les sites *cer* et *nmr* qui sont capables de se lier aux protéines ArgR et PepA ont été identifiés, et de se recombiner si les protéines XerC et XerD sont présentes. Le gène *bla* comporte la résistance à l'ampicilline, et le gène *npt* du transposon Tn903 comporte celle à la kanamycine.

Si on utilise le plasmide pNCK1, c'est pour empêcher des événements de recombinaison *in vitro*, car si les séquences *cer* et *nmr* sont suffisamment proches pour permettre la liaison des protéines ArgR, PepA, XerC et XerD, elles ne le sont pas suffisamment pour permettre une recombinaison spontanée lors des expériences de topologie. Lors de ces expériences de topologie, une autre purification partielle utilisant un chauffage à 70°C (figure 90) pour inactiver les nucléases qui risquent de générer des trainées par dégradation de l'ADN (McCulloch, *et al*, 1994).

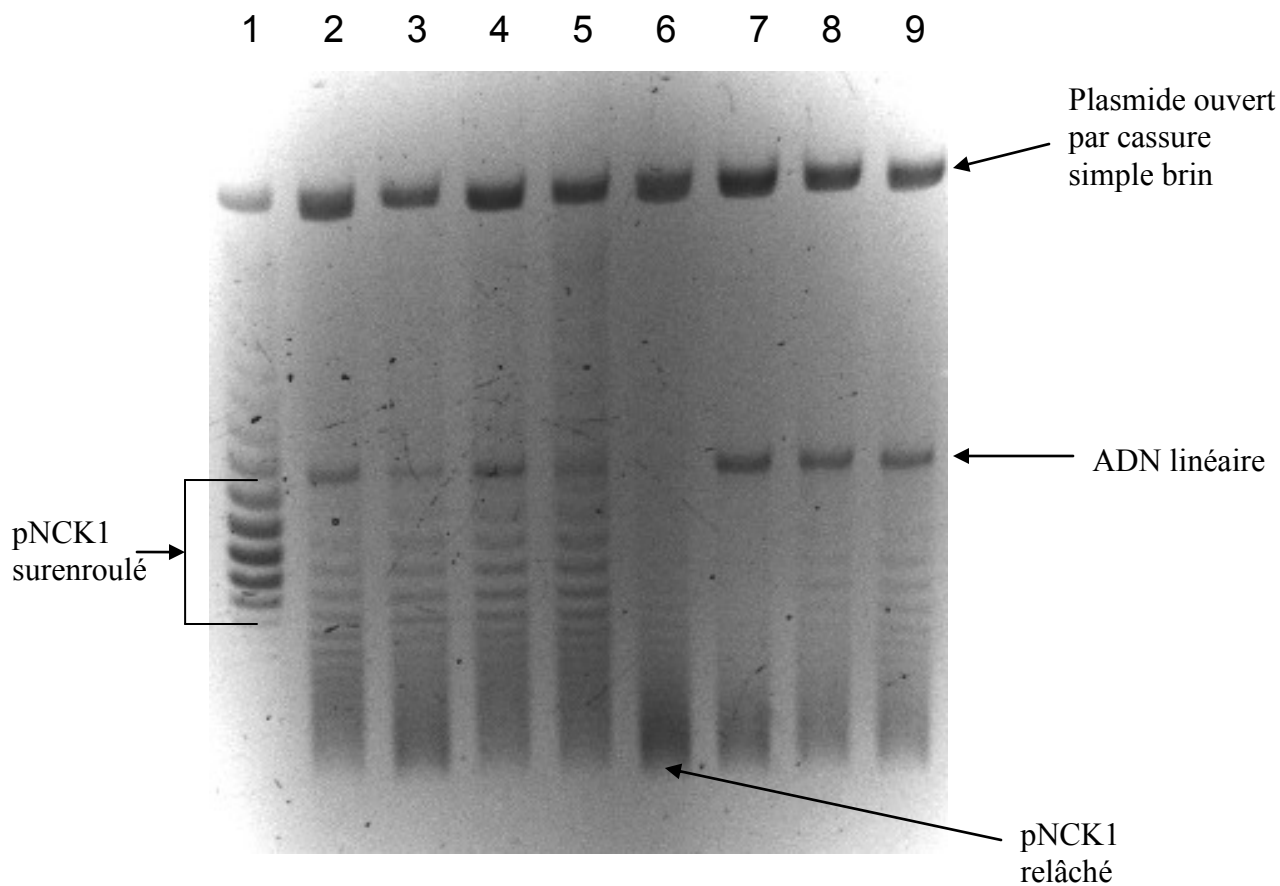


**Figure 92.** Purification de PepA après chauffage à 70°C. Le premier puits montre le marqueur. Le second puits montre la migration du surnageant dialysé contre du tampon TM, la protéine PepA y a été visualisée. Le troisième puits montre le surnageant du culot dialysé contre le tampon TM après précipitation au sulfate d'ammonium (29,1 %), PepA y a été visualisée en grande quantité. Le quatrième puits montre les protéines présentes dans le culot resuspendu. La concentration du gel en polyacrylamide est de 15 % et la coloration du gel a été faite au bleu de Coomassie.

Dans ces expériences, la Topoisomérase I (de Promega) venant du blé, a été utilisée pour relâcher les plasmides surenroulés après incubation à 37° C desdits plasmides avec ArgR et PepA pendant 15 minutes, afin de savoir si les deux protéines étudiées vont avoir un effet sur la topologie de pNCK1 (d'après Alén *et al*, 1997) dans des volumes de 50 µL. La fraction de PepA utilisée après purification est celle visible sur la figure dans le puits numéro 3. Cette fraction ne contient pas d'activité endonucléase et est particulièrement pure (80%). Les échantillons ont été fait migrés pendant 19 à 24 heures sur un gel avec (ou sans) chloroquine.

Avec de la chloroquine présente dans le gel, les types de gels suivants ont été obtenus :





	1	2	3	4	5	6	7	8	9
pNCK1	+	+	+	+	+	+	+	+	+
Topo1	-	+	+	+	+	+	+	+	+
ArgRwt	-	-	+	-	-	-	+	-	-
ArgR149	-	-	-	+	-	-	-	+	-
ArgR5aa	-	-	-	-	+	-	-	-	+
PepA	-	+	-	-	-	-	+	+	+

**Figure 93.** Photo des différents profils des plasmides pNCK1 sur un gel d'agarose 0,8 % avec 0,5  $\mu\text{g/ml}$  de chloroquine. L'ADN traité par la topoisomérase I, donc relâché, a été présenté dans le puits 6. Le puits 1, contient pNCK1 sans traitement à la topoisomérase I qui a été visualisé. Le puits 2, contient pNCK1 avec PepA. Les puits 3, 4 et 5, contiennent pNCK1 avec respectivement ArgRwt, ArgR149 et ArgR5aa. Les puits 7, 8 et 9 contiennent le plasmide pNCK1 qui a été observé avec PepA dans chaque puits, et respectivement avec ArgRwt, ArgR149 et ArgR5aa. Les concentrations en protéines sont les suivantes, [PepA]=18 nM (0,05  $\mu\text{g}$ ) de monomères, [ArgRwt]=120 nM de monomères (0,1  $\mu\text{g}$ ), [ArgR149]=120 nM de monomères (0,1  $\mu\text{g}$ ), [ArgR5aa]=120 nM de monomères (0,1  $\mu\text{g}$ ). La quantité de plasmide pNCK1 est de 500 ng dans chaque expérience.

Lors de ces expériences, la topoisomérase I relâche le plasmide pNCK1 et dans le gel avec chloroquine, le fait migrer plus vite, c'est-à-dire plus bas, que le plasmide non traité (figure 91). Lorsque les protéines ArgRWt, ArgR149 et ArgR5aa, ont été ajoutées sans ajout de PepA, moins de relâchement du plasmide en présence de topoisomérase I a été visualisé. Lorsque PepA est la seule protéine mise en présence de l'ADN, peu de bandes ont été détectées, donc moins de relaxation du plasmide par rapport aux protéines ArgR (en effet PepA se lie moins facilement à l'ADN que ArgR) et au plasmide seul traité par la topoisomérase I. Si les différentes versions d'ArgR sont présentes avec PepA, le profil de relaxation est différent selon que l'on est en présence de la version sauvage d'ArgR ou des versions mutantes. Avec un mélange ArgRWt avec PepA, l'ADN est plus relâché qu'avec PepA seul. Avec les mutants ArgR149 et ArgR5aa, des profils de relaxation montrant plus de surenroulements négatifs sont visibles.

**CHAPITRE IV**  
**DISCUSSION**

## CHAPITRE IV. DISCUSSION

Les études portant sur la recombinaison site spécifique en *cer* ont toutes eu lieu dans différentes souches d'*E.coli*.

### IV.1 Travaux précédents

Certains de ces travaux rassemblent des résultats visant à définir chez la protéine ArgR les résidus d'acides-aminés situés dans la région C-terminale permettant la liaison à la protéine PepA, et ce dans l'optique de générer un complexe de protéines accessoires capables de permettre la recombinaison en *cer* grâce aux recombinases XerC et XerD.

Précédemment des mutants d'ArgR avaient déjà été générés dans la région Ct de cette protéine, entraînant, par exemple, un mutant super répresseur en remplaçant le codon codant le glutamate en position 154 par un stop (Tian and Maas, 1994).

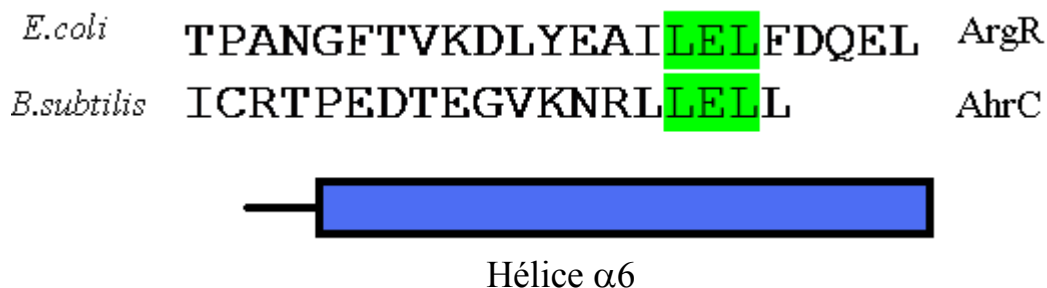
La région Ct d'ArgR a été l'objet de mutations, qui ne gênaient pas l'interaction entre la protéine et l'ADN, sachant que la région Nt d'ArgR interagit avec l'ADN (Tian and Mass, 1994 ; Grandori *et al*, 1995). On peut en conclure une certaine importance de la région Ct dans la recombinaison aux sites *cer* (Sénéchal *et al*, 2010).

On avait précédemment remarqué des similitudes structurales intéressantes entre ArgR et AhrC dans la région C-terminale (North *et al*, 1989 ; Smith *et al*, 1989), qui ont le même nombre d'hélices  $\alpha$  et de feuillets  $\beta$  (soit six feuillets  $\beta$  et six hélices  $\alpha$ ) (Sénéchal *et al*, 2010), de plus, il faut se rappeler qu'AhrC complète des *E.coli* déficientes en ArgR. De surcroît, dans cette région, il existe un triplet d'acides-aminés identiques (leucine-acide glutamique-leucine, soit LEL dans la nomenclature conventionnelle) au niveau des positions 149 à 151 de la protéine ArgR (Sénéchal *et al*, 2010).

Ce triplet est également présent dans la protéine AhrC de *B.subtilis* et cela pourrait expliquer pourquoi, malgré une faible homologie de séquence entre ces deux protéines, la protéine AhrC est capable de compléter les souches d'*E.coli* ne synthétisant pas ArgR. Il se peut aussi que les similitudes structurales entre ArgR et AhrC soit suffisantes pour la complémentation, mais le triplet mérite notre attention.

Il est à noter que lorsque l'on compare les séquences entre *E.coli* et *B.subtilis*, on remarque que ce triplet est précédé par l'isoleucine chez *E.coli* et par une leucine chez *B.subtilis*. Ces acides-aminés sont de fonction similaire, et donc les séquences sont ILEL et LLEL respectivement chez *E.coli* et *B.subtilis*.

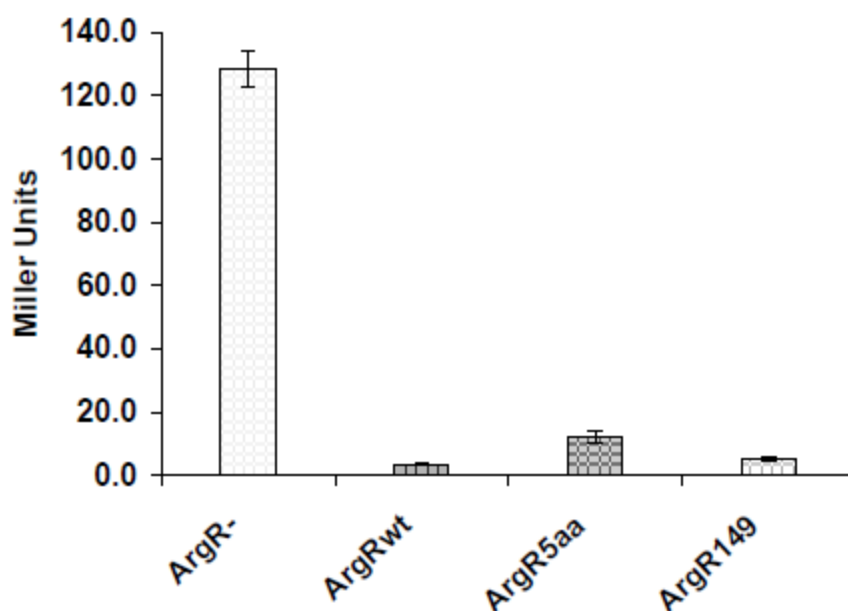
Or lors des expériences menées dans notre laboratoire, c'est ce triplet LEL d'ArgR qui a été ciblé par les différentes mutations (délétions, insertions et substitutions) lors de nos études sur l'interaction entre les deux protéines ArgR et PepA. Il est important de noter que ce triplet LEL, compris dans l'hélice  $\alpha_6$ , correspond respectivement aux acides-aminées en position de 149, 150, et 151 (figure 92).



**Figure 94.** Détails des différences entre les séquences en Ct d'ArgR et d'AhrC. Ici, il a été mis en évidence les triplets LEL que les deux protéines partagent. Il est important de remarquer que la protéine ArgR fait 156 aa et AhrC 149 aa.

Lors de mes travaux j'ai également tenté de voir quelle région de PepA pouvait être impliquée dans l'assemblage du complexe ArgR-PepA.

Si dans notre laboratoire des travaux avaient déjà été menés avec la fabrication des mutants ArgR149, ArgR5aa et ArgRE150A, les plasmides qui permettaient leurs expressions (c'est-à-dire, pHSArgR149, pHSArgR5aa et pHSArgRE150A, tous dérivés du pGS38) contenaient le promoteur *lac* (Sénéchal *et al*, 2010). Outre les tests de recombinaison menés sur le plasmide pCS210 avec les protéines ArgR149, ArgR5aa et ArgRE150A produites avec les dérivés du pGS38 cités plus haut, des tests de l'activité  $\beta$ -galactosidase dans les cellules EC146 $\lambda$  ( $\lambda$ AZ-7) avaient précédemment été réalisés dans notre laboratoire. On avait utilisé différents clones de ces cellules que l'on avait indépendamment transformées par pGS38, pHSArgR5aa et pHSArgR149. Les résultats qui ont été obtenus montrent que par rapport au niveau d'expression basal de l'activité  $\beta$ -galactosidase, la fusion *argA::lacZ* a été réprimée en présence des protéines ArgRWt, ArgR5aa et ArgR149. Cela démontre que les mutations dans les régions C-terminales ne sont pas impliquées dans la répression de la fusion *argA::lacZ* (figure 93).



**Figure 95.** Répression de l'activité  $\beta$ -galactosidase dans les cellules EC146 $\lambda$  ( $\lambda$ AZ-7). Des clones de ces cellules ont été transformés respectivement avec les plasmides pGS38, pHSArgR5aa et pHSArgR149. Les cellules EC146 $\lambda$  ( $\lambda$ AZ-7) contiennent une fusion *argA::lacZ* et possèdent une mutation dans le gène *argR*. La première colonne montre l'expression constitutive de la  $\beta$ -galactosidase quand il n'y a pas de protéine ArgR dans la cellule (128 unités Miller en moyenne). La seconde colonne montre l'expression en présence d'ArgRWt. La troisième montre l'expression en présence d'ArgR5aa et la quatrième montre l'expression en présence d'ArgR149 (avec permission de reproduction, Sénéchal *et al*, 2010).

Ces tests montrent qu'il n'y a pas de différences majeures dans les capacités de ces protéines à se lier à l'ADN pour la répression de la fusion *argA::lacZ in vivo*. Et ce malgré les différences *in vitro* remarquées pour les protéines ArgR149, ArgR5aa comparées au type sauvage.

#### IV.2. Protéines ArgR générées

Suite à ces expériences, dans l'optique d'obtenir une plus grande surexpression de mes protéines, j'ai d'abord généré des plasmides contenant le promoteur *tac* provenant du plasmide pDB169 (surexprimant ArgRWt). Il s'agissait des plasmides pDBArgR149 et pDBArgR5aa surexprimant respectivement les protéines ArgR149 et ArgR5aa.

Si les résultats de purification partielle pour les formes ArgRWt et ArgR5aa étaient satisfaisants au niveau de leurs surexpressions, ceux pour ArgR149 étaient particulièrement décevants. En effet, il y avait une pureté et un rendement final trop faibles (figures 57, 58 et 59). Il se pourrait que lors de la surexpression des protéines ArgR149 natives avec le promoteur *tac*, les protéines soient moins solubles et que l'on ait eu une perte des protéines surexprimées lors de la précipitation au sulfate d'ammonium ou en présence de la L-arginine utilisés dans ce protocole.

Ce qui m'amena à changer de stratégie de purification pour toutes les formes d'ArgR. Une purification des protéines fusionnées à 6-HIS en N-terminale fut choisie. Le plasmide utilisé pour ces nouvelles purifications fut le plasmide pQE31 de Qiagen, vecteur d'expression à promoteur fort, qui après clonage des formes ArgRWt, ArgR149 et ArgR5aa, permet avec un milieu auto-inductible, une bonne purification des protéines sauvages et mutantes (figure 60, 61 et 62).

C'est à partir de ces mêmes plasmides vecteurs pQE31 que les protéines ArgRL149A et ArgRL151A ont été purifiées sur colonne Ni-NTA, leur capacité de

recombinaison site spécifique en *cer* ayant été précédemment testée *in vivo*, comme pour toutes les versions d'ArgR.

La plupart de ces vérifications de la capacité à toujours effectuer la recombinaison ont d'abord été faites après transformation des cellules DS956 contenant le plasmide pCS210. Comme dans les expériences précédemment réalisées avec les protéines ArgRWt, ArgR149 et ArgR5aa, et les plasmides dérivés du plasmide pGS38 et du plasmide pDB169. Ces protéines surexprimées par les vecteurs pQE31 donnent des résultats similaires *in vivo* (Sénéchal *et al*, 2010), c'est-à-dire une absence de recombinaison de pCS210 en présence d'ArgR149 et ArgR5aa. Par contre, les formes délétées ArgR150 et ArgR151, sont toujours capables de faire cette recombinaison, ce qui semblerait indiquer que la leucine et l'acide glutamique en position 149 et 150 (et le restant de la protéine) suffiraient à permettre la recombinaison. On peut donc se poser la question qui va suivre.

Les deux résidus leucines (en position 149 et 151) auraient-ils donc une importance non négligeable dans le phénomène de recombinaison en *cer* ? Il faudrait pour le démontrer substituer en même temps les deux résidus leucines situés en position 149 et 151. Il se pourrait aussi que les résidus touchés par ces mutations soient impliqués dans une possible interaction entre la partie Ct et le DBD (*DNA Binding Domain*) pour la protéine ArgR (Ni *et al*, 1999; Dennis *et al*, 2001).

On pourrait également se poser d'autres questions, y-a-t'il une longueur minimale requise pour la recombinaison ou est-ce la séquence qui importe, voire les deux ?

D'autres mutants générés cette fois par substitution dans les plasmides dérivés du pGS38 ont été testés, les mutants ArgRL149A et ArgRL151A qui permettent toujours cette même recombinaison *in vivo*, quoique pas toujours aussi rapidement (figure 63).



Il a été remarqué qu'entre les mutants ArgR150, ArgR151 ArgRL149A et ArgRL151A, il y a toujours au moins les acides-aminés, leucine et acide glutamique, la version ArgR151 comprenant tout le triplet LEL.

Donc, chez la protéine ArgR, il semblerait qu'il faille au moins un couple LE ou EL (si comparé à AhrC, on constate que ArgR possède une isoleucine en position 148 et AhrC une leucine au niveau du résidu correspondant après alignement des séquences), qui ont des fonctions similaires. En position 152, par contre, les acides-aminés sont la phénylalanine pour ArgR et la leucine pour la protéine homologue AhrC (toujours par rapport à l'alignement des deux séquences). Ces deux acides-aminés ne sont pas de fonction similaire pour que la protéine soit fonctionnelle à première vue. Mais le mutant ArgRE150A était toujours capable de faire les recombinaisons, cela semble indiquer que le couple d'acides-aminés LE qui reste dans cette version tronquée est suffisant pour la recombinaison se fasse. Le couple EL qui lui est plus en position Ct que le couple LE, permet lui aussi la recombinaison (Sénéchal *et al*, 2010). Ce qui suggérerait qu'une leucine jouxtant un acide glutamique suffirait à réaliser la recombinaison. Cependant, le double mutant ArgRL149AL151A de séquence AIAEAFDQEL en Ct permet toujours la recombinaison comme ArgRWt, ce qui suggérerait que l'isoleucine en position 148 pourrait être importante ou remplacer la leucine en position 149 (communication de Gina UGAZ). Du côté de la fixation des mutants sur l'ADN, les formes ArgRWt, ArgR149 et ArgR5aa ont été testées *in vivo* pour examiner leur capacité à réprimer les fusions *argA::lacZ*, et *in vitro* par retard sur gel. Toutes ces formes peuvent se lier à la boîte arginine (figures 66, 67 et 68).

### **IV.3. Retards sur gel et *crosslinking***

Lors de cette discussion, nous allons tout d'abord voir les expériences de retard sur gel et de *crosslinking* au glutaraldéhyde pratiquées sur les protéines ArgRWt, ArgR5aa et ArgR149.

Pour les expériences de retard sur gel, il faut préciser qu'il nous faut travailler à 0,25X de tampon TBE pour pouvoir visualiser nettement dans le même puits une bande retardée ADN+ArgR et une bande d'ADN libre, si cet ADN est la séquence nommée boîte arginine. C'est lorsque des migrations de deux séquences *cer* ont été pratiquées que nous avons des conditions de migration avec 0,5X de tampon TBE et une migration à température ambiante pour que les retards soient visibles. De telles différences seraient probablement dues au degré de pureté des protéines ArgR utilisées, des purifications de protéines fusionnées à 6-HIS (90 % de pureté) ont été utilisées pour les retards sur gel avec la boîte arginine. Pour les retards sur gel des séquences *cer*, ce sont des protéines natives obtenues par purification partielle qui ont été utilisées (79 %, 82 % et 52 % de pureté respectivement pour ArgRWt, ArgR5aa et ArgR149). Il est envisageable que ces purifications de protéines natives aient pu être accompagnées de certaines protéines contaminantes qui auraient stabilisé les complexes ADN-protéine lors des migrations à 0,5X de TBE, surtout lorsqu'il y a eu retards des boîtes arginine dans Sénéchal *et al* (2010).

Il faut également savoir que lors de ces expériences de retards sur gel, on part du principe que la partie Nt des hexamères d'ArgR va se lier à la boîte arginine (Tian and Mass, 1994 ; Grandori *et al*, 1995; Chen *et al*, 1997) et causer le retard de l'ADN marqué par une sonde fluorescente. Avec chaque type de protéine ArgR testée (c'est-à-dire ArgRWt, ArgR5aa et ArgR149), les retards de la séquence boîte arginine (figures 66, 67 et 68) ou de deux séquences *cer* ont été obtenues (figures 69, 70 et 71). En utilisant soit des protéines ArgR avec un groupement 6-HIS en partie Nt de la protéine (permettant une grande pureté des échantillons de protéines) avec les boîtes arginine, soit en utilisant des protéines natives pour les retards sur gels avec les deux séquences *cer*. La présence des 6-HIS ne gênant en rien la formation des hexamères d'ArgR ni la fixation des multimères (comme démontré dans les expériences de *crosslinking* et les retards sur gel avec la boîte arginine), les résultats ont des interprétations similaires entre les formes natives et les formes 6-HIS des protéines ArgR.

Pour mieux étudier les différences observées avec nos mutants au niveau de leurs liaisons avec l'ADN, nous avons posé la question suivante. Est-ce que ces différences sont dues à la capacité, ou l'incapacité de ces mutants à former des hexamères ? C'est pour cela que l'on a procédé à des expériences de *crosslinking* au glutaraldéhyde avec les trois formes suivantes d'ArgR, ArgRWt, ArgR5aa et Arg149 (figures 73, 74 et 75). Les tests s'inspirent de la méthode utilisée par Chung-dar Lu *et al*, 1992. Comme dans cet article, où la protéine ArgR de *S.typhimurium* avait été testée, les résultats avec la forme sauvage d'ArgR d'*E.coli* sont comparables à ceux obtenus avec l'ArgR de *S.typhimurium*, bien que ce soit plutôt avec des pourcentages autour de 0,02 et 0,04 % que le plus grand nombre d'hexamères a été visualisé pour ArgR d'*E.coli*, dans des tampons identiques, entre cet article et mes expériences. Il faut rappeler que le glutaraldéhyde stabilise une liaison entre protéines qui sont généralement déjà capables d'interagir entres-elles. Cette stabilisation a lieu par formation d'une liaison covalente. Les deux ArgR issues de *S.typhimurium* et d'*E.coli* font donc partie de ces protéines stabilisables par le glutaraldéhyde (Chung-dar lu *et al*, 1992). Il est également remarquable que dans l'article, le pourcentage maximum de glutaraldéhyde soit de 0,01%.

Pour les résultats des expériences de retards sur gel avec ArgRWt et les *crosslinking* au glutaraldéhyde, on a constaté les choses suivantes, ArgRWt donne des bandes nettes lors de retards sur gel réalisés avec la boîte arginine, lorsque les protéines se lient à l'ADN. Lors des retards sur gel avec les deux séquences *cer* séparées de 600 pb, la protéine ArgRWt native, s'est liée avec l'ADN des deux séquences *cer*. Plus il y a eu de protéines, plus le retard a été élevé, jusqu'à un certain seuil de 4920 nM où les séquences ne migrent plus, traduisant une liaison de tous les ADN avec les protéines ArgRWt présentes (cette liaison pouvant se faire avec des formes trimériques ou hexamérique de la protéine). Lors des expériences de *crosslinking* au glutaraldéhyde, la formation de multimères jusqu'à la forme hexamérique d'ArgRWt a été constatée, mais à partir de 0,02% de glutaraldéhyde, et pour des pourcentages plus élevés, des agrégats de protéines se forment facilement pour la forme ArgRWt (figure 73).

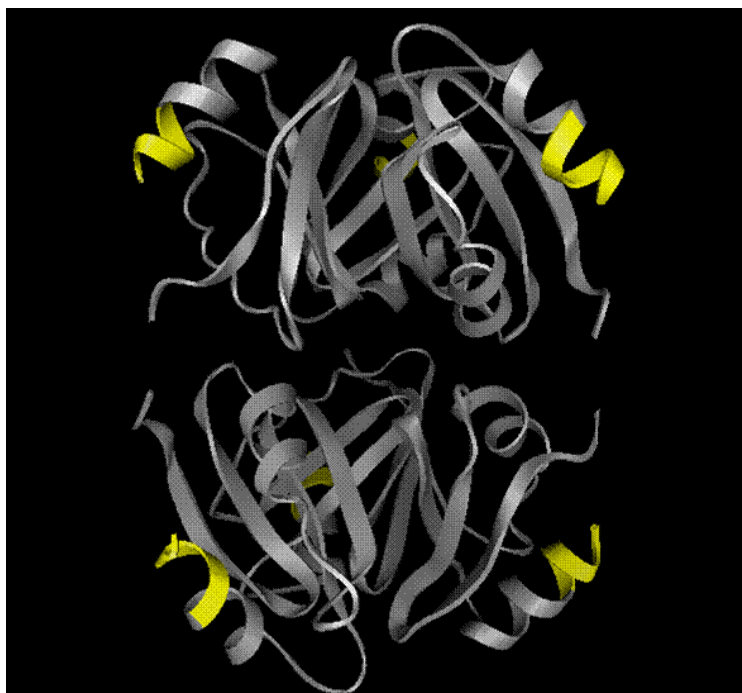
En ce qui concerne la protéine ArgR5aa, il y a toujours formations de bandes retardées nettes avec la boîte arginine, ce qui indique que la protéine est stable, et que la mutation par insertion de 5aa ne gêne pas la fixation des hexamères d'ArgR5aa sur les séquences d'ADN, que ce soit avec la boîte arginine ou la séquence *cer* (figures 69 et 70). Les résultats comparés avec ArgRWt et ArgR5aa lors des retards sur gel générés avec les deux séquences *cer*, montrent que des concentrations proches (4920 nM pour ArgRWt, 4998 nM pour ArgR5aa) ont pu entraîner les retards avec formations de complexes ADN-protéines stables. Un retard des bandes maximal à 6180 nM d'ArgRWt et 7500 nM pour ArgR5aa a également été observé (figures 69 et 70). Pour la protéine ArgR5aa, les résultats des expériences de *crosslinking* au glutaraldéhyde sont assez proches de ceux obtenus avec la protéine ArgRWt. En effet, la multimérisation des protéines se fait de manière évidente à des concentrations de 0,01 et 0,02 % de glutaraldéhyde où des pentamères et des hexamères ont été observés. Il y a également formation d'agrégats très visibles à 0,02% de glutaraldéhyde (figure 74). Cette forme mutante d'ArgR est donc toujours capable de former des hexamères. Qui plus est ces hexamères sont capables de se lier à l'ADN, tout comme ceux d'ArgRWt. Le problème de recombinaison *in vivo* qui a été observé n'est donc pas dû à un problème de formation des hexamères d'ArgR5aa comme on peut l'envisager pour ArgR149. Pour cette forme ArgR5aa, il est envisageable de penser qu'à ce stade les problèmes rencontrés pour la recombinaison en *cer* soient des problèmes d'interaction entre les hexamères de PepA et d'ArgR5aa qui n'empêchent pas forcément la liaison ArgR5aa-PepA, mais qui empêcheraient la recombinaison *in vivo* même si le contact entre PepA et ArgR5aa a lieu.

Pour la protéine mutante ArgR149, lors des retards sur gel, les concentrations d'ArgR149 nécessaires pour obtenir un retard étaient plus importantes comparées à celles d'ArgRWt et ArgR5aa pour obtenir un retard. Si les concentrations étant proche du simple au double pour ArgRWt et ArgR5aa (144 nM d'ArgRWt et 363,6 nM pour ArgR5aa) pour obtenir un retard complet avec la boîte arginine, on a constaté avec ArgR149 l'apparition d'une trainée traduisant une dissociation partielle du complexe ADN-protéine qui se fait facilement à ces concentrations de la protéine ArgR149. Des

bandes dont l'intensité était moindre (moins visible) ont même été constatées lors de la migration des fragments d'ADN contenant les deux séquences *cer* (figure 71), même en augmentant la concentration en protéine. Ce qui peut s'interpréter comme une autre preuve du manque de stabilité des complexes formés par ArgR149, ou d'une affinité réduite. Lors des expériences de *crosslinking* au glutaraldéhyde, il a été visualisé qu'ArgR149 forme moins d'hexamères que les autres formes testées, même avec 0,08% de glutaraldéhyde. Les formes les plus courantes étant la forme dimérique et la forme tétramérique de cette protéine (la forme dimérique étant la plus présente), accréditant un problème de formation des hexamères de la protéine ArgR149 *in vitro* (figure 71 et 75) (Ni *et al*, 1999). En effet, ce problème de stabilité fut aussi visualisé dans les retards sur gel avec la boîte arginine et les deux séquences *cer*. Dans ces expériences, il fallait utiliser des quantités et des concentrations plus importantes d'ArgR149 que pour les deux autres formes de la protéine, ArgRWt et ArgR5aa. De plus, il y a peu d'agrégats de cette protéine même à forte concentration de glutaraldéhyde. Même si théoriquement, un tétramère pourrait se lier à l'ADN (Sunnerhagen *et al*, 1997), la forme tétramérique reste rare et difficile à générer d'après les *crosslinking* au glutaraldéhyde (figure 75). Le problème de formation des hexamères peut être envisageable au regard de tous ces résultats.

#### **IV.4. Possibles causes des problèmes de multimérisations.**

Ce problème d'hexamérisation pourrait être engendré par un problème d'interaction Ct-DBD, sachant que le DBD (*DNA binding domain*) correspond à la partie Nt, (Ni *et al*, 1999) pour la protéine mutante ArgR149. Cette dernière possibilité est tout à fait plausible après comparaison entre les hexamères d'ArgRBst et d'AhrC évoquée plus haut (figure 42) (même si pour l'instant, des problèmes d'interaction avec PepA ne sont pas à exclure).



**Figure 96.** Représentation d'un hexamère des parties Ct d'ArgR d'*E.coli*. Les parties en jaune sont celles où on a réalisé les différentes mutations au cours des travaux présentés dans cette thèse. Cette région Ct, en temps normal dénuée de mutation, pourrait représenter un site d'interaction avec une autre protéine comme PepA (logiciel utilisé RasMol).

Selon la modélisation tridimensionnelle présentée en figure 94, les hexamères des parties Ct de la protéine montrent que la partie mutée lors de nos expériences est rejetée en périphérie des faces externes de l'hexamère, comme c'est également le cas avec les hexamères d'AhrC, d'ArgRBst et de *MtbArgR* (Ni *et al*, 1999; Dennis *et al*, 2002 ; Cherney *et al*, 2010). Ce qui pourrait également laisser supposer deux possibilités, connaissant les interactions Ct et Nt chez les homologues d'ArgR. Soit les mutations empêchant la recombinaison ont rendu la protéine plus dégradable, mais ce qui est improbable car la répression du gène *lacZ* dans la fusion *argA :: lacZ* a été observée à des niveaux équivalents à ArgRWt et ArgR5aa et ces expériences ont lieu *in vivo*. Il est donc bien plus logique de supposer que les mutations empêchent une interaction normale entre la partie Ct d'ArgR et la partie Nt d'une autre sous-unité de cette même protéine, ce qui semble probable après comparaison avec les structures connues des hexamères d'AhrC (capables de compléter des souches d'*E.coli* déficientes en ArgR), d'ArgRBst et de *MtbArgR* (Ni *et al*, 1999; Dennis *et al*, 2002 ;

Cherney, *et al* (2010), et les résultats obtenus par Holtham *et al* (1999), avec des protéines chimériques composées des parties Nt et Ct d'ArgR et d'AhrC. De plus des expériences menées sur la protéine ArgR provenant de bactérie *Bacillus stearothermophilus* ont montré que justement la partie Nt (au niveau des acides-aminés I8, R11 et M15) provenant d'une sous-unité interagit avec la partie Ct (au niveau des acides-aminés L146 et L149) provenant d'une autre sous-unité située sur l'autre trimère d'ArgR de *B.stearothermophilus* (Ni *et al*, 1999) (figure 42). Ce qui laisse à penser que ce serait également le cas avec la protéine ArgR d'*E.coli*. De plus, la protéine ArgR de la bactérie *B.stearothermophilus* a une structure monomérique et hexamérique similaire à celle de la protéine ArgR d'*E.coli* (Ni *et al*, 1999; Dennis, *et al*, 2002), mais surtout d'AhrC. Les résidus Nt interagissant avec la partie Ct dans la protéine ArgR sont probablement différents de ceux de ses homologues précédemment cités, bien que probablement de structure et charges proches.

#### **IV.5. Comparaisons expérimentales pour les retards sur gel.**

Si l'on revient sur les expériences avec les deux séquences *cer* et les protéines ArgRWt, ArgR149 et ArgR5aa natives, on a bien eu des retards sur gel. Pour la protéine ArgR5aa native, comme avec ArgRWt, il y a une concentration pour laquelle tous les ADN sont retardés, donc la partie Ct, comme avec la séquence boîte arginine, n'intervient pas dans la liaison à l'ADN. C'est seulement la partie Nt non mutée qui rentre en contact avec l'ADN (Grandori *et al*, 1995). Pour la protéine ArgR149 native, il y a bien un retard des fragments d'ADN contenant les deux séquences *cer*, comme avec les deux autres formes natives. On remarquera cependant, que la forme ArgR149 montre des retards qui forment des bandes dont l'intensité est moindre malgré l'augmentation de la concentration d'ArgR149. Si ArgR149 réprime autant la fusion *argR :: lacZ* que la forme ArgRWt *in vivo*, cela peut toujours s'interpréter comme une autre preuve du manque de stabilité d'ArgR149 *in vitro*, bien que les résultats de Sénéchal *et al* (2010) montrent des liaisons aux boîtes arginine, probablement stabilisées par d'autres protéines, car la purification d'ArgR149 présentée est native.

Toutes ces expériences de retards sur gel peuvent être comparées à d'autres précédemment réalisées avec ArgRWt et d'autres versions mutantes. Pour ce qui est des retards sur gel avec les deux séquences *cer*, les concentrations de protéines ArgRWt sont plus importantes que dans l'article publié par Stirling *et al* (1988). Les facteurs qui expliquent cette différence sont, le degré de pureté des protéines entre les purifications partielles et les purifications sur colonne Ni-NTA, le fait que les migrations des séquences *cer* se sont faites avec du TBE 0,5X à température ambiante et celles des boîtes arginine avec du TBE 0,25X à 4°C. La composition des tampons de liaison, qui ne contiennent pas les mêmes concentrations en sels et glycérol sont également à évoquer, et finalement, le fait que la purification partielle utilise une précipitation des protéines à la L-arginine alors que la purification sur colonne utilise l'affinité du groupement 6-HIS pour la résine Ni-NTA. Toutefois, malgré ces différences, les retards sont observés, et ce de manières similaires dans le sens que les protéines ArgR n'entraînent pas la formation de plusieurs bandes, c'est-à-dire un fragment d'ADN libre et un fragment lié à la protéine comme on a pu le visualiser avec la boîte arginine si toutes les protéines ne sont pas liées à la séquence d'ADN. En effet, avec les séquences *cer*, plus on met de protéines ArgR, plus le retard devient important sur le gel, comme il a été visualisé sur figures 69, 70 et 71, et dans les travaux de Stirling *et al*, 1988. Pour ce qui est des retards sur gel avec la boîte arginine et les protéines ArgRWt, ArgR5aa et ArgR149, les concentrations pour ArgRWt et ArgR5aa utilisées sont moins élevées que pour les retards sur gel publiés dans Sénéchal *et al* (2010) et Chen *et al* (1997). Les différences principales sont que les tampons de liaisons ne sont pas identiques, mais surtout que les protéines que j'ai utilisées dans mes retards sur gel avec la boîte arginine sont fusionnées avec 6-HIS, surexprimées par un vecteur pQE31 et purifiées en conséquence, alors que dans les deux articles cités plus en amont, les protéines ArgR sont natives et surexprimées à partir de pGS38 ou pDB169. De plus, les moyens de détections ne sont pas non plus les mêmes, en effet, si dans l'article de Chen *et al* se sont des ADN marqués par un radioélément (le <sup>32</sup>P), dans l'article de Sénéchal *et al*, c'est une méthode de détection à la digoxygénine qui a été utilisée pour les boîtes arginine. Dans cette thèse, ce sont les retards sur gel des séquences *cer* qui ont été détectés à la digoxygénine, et les retards sur gel avec les boîtes arginine, quant à elles,



ont été marquées par HEX et détectées par fluorescence. Ces différences dans les méthodes de détections pourraient aussi entraîner des différences de sensibilité lorsque l'on visualise les retards. Mais les sensibilités entre le marquage par HEX et par digoxygénine sont proches et les quantités d'ADN qui ont été utilisées pour le marquage à la digoxygénine des séquences *cer* sont identiques à celles publiées dans Sénéchal *et al*, et les concentrations proches, donc il ne devrait pas y avoir de différences majeures sur le niveau de détection des ADN liés aux ArgRs. Il est important de noter que lors de mes expériences, les boîtes arginines marquées par HEX ont été faites migrées à 4°C et avec un tampon TBE de concentration 0,25 X, ce qui rend les complexes ADN-ArgR plus visibles. Il est remarquable que la protéine ArgR149, dans le cadre de mes expériences, ait dû être utilisée à de grandes concentrations pour obtenir un retard complet (de l'ordre de 3675 nM). Par contre les concentrations sont assez proches pour retarder environ la moitié des brins d'ADN (aux alentours de 918 nM de protéines monomériques) lors de ces expériences.

#### **IV.6. Expériences avec PepA.**

Pour la protéine PepA, on a aussi testé sa capacité à se multimériser car, comme ArgRWt, PepA peut s'hexamériser (Sträter *et al*, 1999 ; Reijns *et al*, 2005) et ce de la même manière que les différentes formes de la protéine ArgR. De plus, outre sa capacité à s'hexamériser, PepA fait partie du complexe de recombinaison XerC-XerD-ArgR-PepA qui, dans notre cas, permet la recombinaison en *cer*. L'hexamère de PepA étant un considéré comme un dimère de trimères (Sträter *et al*, 1999 ; Reijns *et al*, 2005), l'hexamérisation de PepA passerait d'abord par une trimérisation des sous-unités. Ce qui semble vérifié quand on regarde les résultats des expériences de *crosslinking* au glutaraldéhyde où on obtient très facilement des trimères de PepA dès 0,005% de glutaraldéhyde avant l'apparition des hexamères (figure 68).

Par contre, les formes les plus lourdes sont surtout visibles à de forts pourcentages de glutaraldéhyde (surtout vers 0,08%). Et de plus, on peut surtout voir

des trimères à faible pourcentage comme à fort pourcentage de glutaraldéhyde, ce qui tendrait à aller dans le sens que l'hexamérisation de PepA passe d'abord par une trimérisation des sous-unités (Ni *et al*, 1999 ; Dennis *et al*, 2002 ; Cherney *et al*, 2010).

Lors des expériences de *crosslinking* réversibles au DSP, les versions d'ArgR testées avec PepA furent les suivantes : ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A. Le but de cette expérience étant de voir si toutes les versions ou non de la protéine ArgR sont capables d'interagir avec PepA, le DSP permettant de créer une liaison d'environ 12 Angströms entre deux protéines qui interagissent naturellement (figures 85 et 86).

Cette liaison faite par le DSP est réversible en présence de DTT ou de  $\beta$ -mercaptoéthanol, deux produits rajoutés lors de la dénaturation des protéines à chaque migration SDS-PAGE suivant les *crosslinking* entre les deux protéines.

Les résultats pour une migration avec ou sans présence de L-arginine (figures 85 et 86) furent assez surprenants. Les deux types de migrations indiquent une interaction entre chaque version d'ArgR, même avec ArgR149 et ArgR5aa, et PepA. Il existait un contrôle PepA seul avec les billes magnétiques Ni-NTA montrant que PepA n'interagissait pas avec les billes en présence de DSP, après traitement des protéines PepA au DSP. En présence de L-arginine, la protéine ArgR s'hexamérise de manière stable, permettant une liaison de deux trimères d'ArgR (Van Duyne *et al*, 1996). L'absence de L-arginine devrait en théorie rendre ArgR incapable de s'hexamériser de manière stable, mais n'empêchant pas la liaison ou la formation de trimères.

L'une des possibilités est que dans chaque cas, en présence de L-arginine ou non, les protéines ArgR multimères soient capables de se lier à un monomère ou un trimère de la protéine PepA en présence de DSP (la liaison à un hexamère devant être surtout favorisée par la liaison à la L-arginine), qui permet des liaisons entre protéines avec une séparation de 12 Angströms. L'autre possibilité est que le DSP soit un

*crosslinker* qui soit capable de favoriser la liaison des deux protéines trimériques ArgR et PepA, même sans L-arginine.

Une autre possibilité est que sans ou bien avec L-arginine, un trimère d'ArgR, soit capable de se lier avec un trimère de PepA en présence de DSP. On peut aussi avancer qu'un trimère d'ArgR ait pu interagir avec un hexamère de PepA en présence de L-arginine ou non.

Lors des expériences menées dans notre laboratoire, on avait avancé que les mutants ArgR149 et ArgR5aa avaient des problèmes d'interaction avec PepA, or les expériences de *crosslinking* au DSP (figures 85 et 86) montrent que toutes les formes mutantes testées étaient capables d'interagir avec PepA. La remarque importante dans ces expériences de *crosslinking* au DSP est que l'on voit une intensité similaire des bandes correspondant à PepA dans les expériences avec et sans L-arginine avec les formes d'ArgR testées (ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A). Ce qui nous fait penser que l'interaction aurait très bien pu se faire avec des trimères d'ArgR qui n'ont pas besoin de L-arginine pour se former (Van Duyne *et al*, 1996), malgré le fait qu'ArgR149 semble faire peu de trimères si l'on regarde les résultats du *crosslinking* à la figure 65. C'est ensuite possible que des trimères de chaque forme d'ArgR aient pu interagir soit avec des trimères de PepA ou des hexamères de PepA, les conditions d'expériences utilisées n'empêchant pas la formation des hexamères de PepA.

Un autre point que ces *crosslinking* au DSP soulèvent, est la raison pour laquelle toutes les formes testées d'ArgR (mutantes ou non) ont été capables d'interagir avec PepA, alors que l'une des hypothèses avancées dans notre laboratoire était que les formes ArgR5aa et ArgR149 pouvaient ne pas interagir correctement avec PepA. La forme ArgR5aa pouvant former des hexamères, l'insertion des cinq acides-aminés aurait pu entraîner une gêne stérique qui nuirait à l'interaction avec PepA, mais on voit qu'il y a visiblement toujours une interaction possible. La forme ArgR149 quant à elle, de part sa délétion, aurait eu un problème d'interaction à cause d'un manque d'acides-

aminés. Ces acides-aminés manquants seraient ceux qui favoriseraient la stabilité du complexe ArgR-PepA de part une possible interaction entre les parties Ct et Nt de la protéine ArgR. Il se peut que ce soit aussi ceux qui interagiraient avec PepA. Le fait que l'on ait pu voir des interactions entre ArgR et PepA pour chaque forme de la protéine ArgR laisse à supposer que les mutations n'ont pas entraîné un manque total d'interaction, ce qui expliquerait les bandes correspondant à PepA. Après tout, il y a toujours un faible pourcentage de recombinaison *in vivo* avec les formes ArgR149 et ArgR5aa (Sénéchal *et al*, 2010), bien que ce dernier ne soit pas équivalent à ce que l'on a pu visualisé *in vitro* avec les expériences au DSP. Il se pourrait donc qu'en présence de DSP, dès qu'il y a la moindre interaction entre les protéines PepA et ArgR (sous ses différentes formes) ou approche à 12 Angströms des deux protéines, les complexes soient très rapidement stabilisés peut-être à cause de la longueur du *linker arm*.

Ce qui finalement nous amène à penser qu'un autre *crosslinker* réversible serait plus approprié pour mieux déterminer s'il y a des problèmes d'interaction. Il faudrait essayer le DST (Disuccinimidyl tartarate) qui permet une liaison entre protéines avec un *linker arm* permettant une séparation de 6,4 Angströms, limitant les problèmes possiblement liés à la longueur du *crosslinker*. Si on veut utiliser des *crosslinker* non réversibles, on pourrait très bien se servir de produits tels que l'AMAS (N-(α-maleïmidoacétoxy) succinimide ester) (4 Angströms), le DFDNB (1,5-Difluoro-2,4-dinitrobenzène) (3 Angströms) ou l'EDC (1-Ethyl-3-[3-diméthylaminopropyl] carbodiimide Hydrochloride) (0 Angström).

Pour essayer de voir quelles régions de PepA interagissent avec ArgR, nous avons décidé de « tronquer » PepA en deux parties. C'est ainsi que les protéines PepACtMBP et PepANtMBP furent générées dans ce but, pour voir s'il y aurait liaison d'une de ces parties Nt ou Ct de PepA avec ArgR. On n'a pas remarqué de multimérisation de ces formes tronquées PepANtMBP et PepACtMBP en présence de glutaraldéhyde. À la vue de ces derniers résultats les formes tronquées de PepA ne sont visiblement pas capables de former d'hexamères, ni de trimères, mais en revanche, on a visualisé très rapidement une agrégation des formes tronquées (figures 81 et 82). De

plus, avec la faible quantité de PepA observée lors des expériences d'interaction ArgR-PepA avec les billes magnétiques Ni-NTA, nous avons décidé de faire l'expérience « inverse », c'est-à-dire isoler PepA (ou les parties de PepA) via des billes magnétiques amylose, après la fusion générée entre le MBP et chaque partie de PepA (Nt et Ct), et de voir si la protéine ArgR pouvait être récupérée avec les protéines PepANtMBP et PepACtMBP. Selon les résultats obtenus, il semble qu'il n'y ait pas d'interaction entre ArgR et les formes tronquées PepACtMBP et PepANtMBP qui soit visualisée par cette approche.

Avec les interactions de protéines à protéines complétant les résultats précédents, ni PepACtMBP ni PepANtMBP n'ont interagi avec la protéine ArgRWt (figure 87). Pris ensemble, ces résultats semblent montrer que les deux régions N-terminale et C-terminale de la protéine PepA pourraient interagir toutes les deux avec la protéine ArgR et seraient impliquées dans la formation du complexe ArgR-PepA, ce qui permettrait la recombinaison en *cer*. Les deux régions Nt et Ct rentreraient donc en contact avec la protéine ArgR, rendant le complexe de recombinaison avec les protéines XerC et XerD actif.

De plus, quand on teste l'activité aminopeptidase des protéines PepACtMBP et PepANtMBP comparée à celle de PepA, seule PepA entière est encore capable de réaliser la réaction de restriction des peptides à partir de leur extrémité possédant le groupement amine. Montrant par là-même que son activité aminopeptidase, générée par la partie Ct de la protéine, était intacte contrairement aux deux formes tronquées de la protéine. Il se peut également que le fait d'avoir tronqué de telle manière la protéine PepA en séparant les deux régions Nt et Ct ait rendu les deux régions inactives. En effet, il ne semble pas que les deux régions séparées puissent se multimériser comme le montre les expériences de *crosslinking* au glutaraldéhyde où on ne voit que des agrégats de ces deux protéines. Il a été tenté de voir si l'activité aminopeptidase est restituée si on mélange les deux protéines tronquées PepANtMBP et PepACtMBP et si ledit mélange est mis en présence de L-leucine p-nitroamilide hydrochloride (LPNA). Il n'y a pas eu de restitution de l'activité aminopeptidase, montrant une absence d'interaction

entre les deux parties tronquées (ce qui pourrait être impliqué à un blocage dû au groupement MBP, ou à un problème de repliement des protéines tronquées à cause de la fusion avec le groupement MBP). Quoi qu'il en soit, ces expériences ne permettent pas de trancher entre le modèle de Sträter modifié et le modèle proposé par Reijns *et al* en 2005 (Sträter *et al*, 1999 ; Reijns *et al*, 2005), pour connaître les résidus et parties de PepA interagissant avec les séquences *cer*.

#### **IV.7. Expériences de recombinaison *in vitro*.**

Les expériences de recombinaison *in vitro* des plasmides montrent, quant à elles, des profils de recombinaison différents selon la version de la protéine ArgR présente dans le mélange. Pour cela trois versions de la protéine ArgR ont été étudiées (figure 90). La protéine ArgRWt, qui permet la recombinaison et la formation des jonctions d'Holliday *in vitro*. La protéine ArgR149, qui ne permet pas la recombinaison *in vitro*, ni la formation de la jonction d'Holliday qui l'accompagne. La protéine ArgR5aa, qui permet une formation de jonction d'Holliday, mais ne permet pas la recombinaison de pCS210 *in vivo*.

Ce dernier résultat *in vitro* est assez remarquable du fait que contrairement aux recombinaisons *in vivo*, on a pu visualiser, faiblement certes, la formation de jonction d'Holliday avec ArgR5aa *in vitro*, alors que la forme mutante ArgR5aa est incapable de réaliser la recombinaison *in vivo*. Ces résultats suggèrent qu'ArgR149 est incapable de former un complexe de recombinaison, et qu'ArgR5aa *in vitro* en soit capable, mais il n'y a pas de résolution de la jonction d'Holliday pour obtenir une recombinaison finie grâce aux recombinases XerC et XerD et aux résolvases cellulaires *in vivo* qui sont nécessaires pour compléter la recombinaison de pCS210 vers le plasmide pCS211 (McCulloch *et al*, 1996). Ce qui semble indiquer que pour ArgR5aa, on retourne vers la forme de départ, c'est-à-dire les deux séquences d'ADN non recombinées. On aurait donc perte de la directionnalité de la réaction avec ArgR5aa.

Ces résultats ont été comparés avec ceux de Villion et Szatmari (2003), on visualise beaucoup plus de formes relaxées du plasmide pCS210, et les bandes correspondant aux jonctions d'Holliday sont tout à fait visibles sur ces deux expériences de Southern blot, ce qui prouve que les résultats sont reproductibles et montrent qu'il y a formation d'une jonction d'Holliday avec la forme ArgRWt suivi *in vivo* par la recombinaison. Si les profils avec ArgRWt sont ceux attendus, pour ArgR149 par contre, on a aucune formation de jonction d'Holliday, ce qui semble indiquer une absence de formation de complexes ArgR-PepA stables, donc il n'y aurait pas suffisamment intervention des recombinases XerC et XerD. Ce qui semble s'accorder avec les résultats obtenus avec les expériences menées avec le glutaraldéhyde où des problèmes de formation des hexamères d'ArgR149 ont été visualisés.

Les résultats obtenus avec ArgR5aa semblent mettre en avant le fait que la jonction d'Holliday se forme (mais les bandes sont moins intenses), mais que les résolvas cellulaires ne peuvent agir, il y aurait peut-être une difficulté d'alignement des sites due à l'encombrement stérique que les 5 acides-aminés entraînent, compromettant la recombinaison, la jonction d'Holliday se formant malgré les 5 acides-aminés supplémentaires, mais cette dernière jonction ne pouvant être résolue, il y a retour vers le produit de départ de la réaction, c'est-à-dire un plasmide sans recombinaison. Le problème dû aux 5 acides-aminés supplémentaires de ce mutant ne serait pas un problème de formation du complexe ADN-protéine, ni de formation des jonctions d'Holliday, mais de la résolution de ces dernières. Ce qui s'accorde aussi avec les résultats obtenus avec les expériences de *crosslinking* au glutaraldéhyde où ArgR5aa s'hexamérise, et les expériences de *crosslinking* au DSP où les protéines ArgR5aa pouvaient toujours interagir avec PepA. Les résolvas cellulaires évoquées plus haut sont encore inconnues (Mc Culloch *et al*, 1994).

Si l'on compare les résultats avec ArgR149 avec les données de l'article de Minh et al (2009) où il a été avancé que PepA pourrait éventuellement permettre la recombinaison en *cer* sans intervention d'ArgR, on peut avancer avec nos résultats de

recombinaison *in vitro* que PepA ne peut effectuer la recombinaison avec ArgR149, et que la protéine ArgR fonctionnelle (ArgRWt) est en fait requise pour cette recombinaison au niveau du site *cer*. Une expérience qui pourra être envisagée, c'est de mettre la protéine PepA aux concentrations utilisées seulement avec XerC et XerD pour vérifier si la recombinaison peut avoir lieu sans la présence d'ArgR comme il a été avancé par Minh *et al* (2009). Or les concentrations de PepA utilisées sont déjà très élevées, c'est-à-dire 10 fois plus importantes que celles des différentes versions d'ArgR. Il est donc possible que PepA sans ArgR n'entraîne pas de formation de jonction d'Holliday comme lorsque la protéine PepA est en présence d'ArgR149.

Les versions ArgRE150A, ArgR150, ArgR151, ArgRL149A et ArgRL151A n'ont pas été testées pour la recombinaison *in vitro* car elles permettaient déjà la recombinaison *in vivo*.

De plus, il a été suggéré, que la phénylalanine présente en position 152 qui est un acide-aminé aromatique et l'acide aspartique, acide-aminé polaire chargé négativement et situé en position 153 auraient permis la stabilisation de la dernière hélice  $\alpha$  (rapport de stage de Chloé Jaubert).

De plus, en position 153, l'introduction d'un stop a généré un mutant super-répresseur, renforçant la thèse d'une stabilisation de la dernière hélice  $\alpha$ . Mais les protéines tronquées ArgR150 et ArgR151 permettant toujours la recombinaison, cette supposition ne semblent pas être la plus vraisemblable.

Dans le tableau suivant, les effets des différentes protéines ArgR testées et mises au point dans notre laboratoire sur la liaison à l'ADN ont été résumés, que ce soit sur les séquences boîte arginine pour les tests *in vitro* de retards sur gel, ou sur la séquence *cer* pour les tests *in vivo* pour la recombinaison plasmidique.

On remarquera que la séquence *cer* a été utilisée dans les expériences de recombinaisons aussi bien *in vivo* qu'*in vitro*.



protéines	Partie Ct	Liaison ADN		Recombinaison	Jonction d'Holliday
		<i>in vivo</i>	<i>in vitro</i>	<i>in vivo</i>	
ArgRWt	AILELFDQEL	+++	+++	+++	+++
ArgR149	AIL	+++	+	-	-
ArgR5aa	AILGVPLLELFDQEL	+++	+++	-	+
ArgRE150A	AILALFDQEL	+++	+++	+++	nd
ArgR150	AILE	+++	nd	+++	nd
ArgR151	AILEL	+++	nd	+++	nd
ArgRL149A	AIAELFDQEL	+++	nd	++	nd
ArgRL151A	AILEAFDQEL	+++	nd	++	nd

**Tableau 13.** Comparaison des différentes versions d'ArgR. Les différentes parties Ct Terminales des différentes versions d'ArgR obtenues sont indiquées. Les capacités de liaisons à l'ADN ainsi que la capacité pour permettre une recombinaison à la fois *in vitro* et *in vivo* ont été indiquées, **nd** voulant dire **non déterminé** à la fois pour les jonctions d'Holliday (qui restent fortement probables) et les liaisons à l'ADN. La liaison à l'ADN *in vivo* est exprimée par le degré de répression d'une fusion *argA :: lacZ*.

#### IV.8. Expériences de topologie.

Pour vérifier l'interaction ArgR-PepA, nous avons visualisés les changements topologiques du plasmide pNCK1, contenant les sites *cer* et *nmr* (Zakova and Szatmari, 1995). On a ainsi vérifié les effets des protéines PepA, ArgRWt, ArgR149, ArgR5aa sur la topologie de l'ADN en présence de topoisomérase 1 et dans des gels pourvus ou non de chloroquine.

Pour les gels de chloroquine (Alén *et al*, 1997 ; Mc Culloch *et al*, 1994), les profils sont un peu différents de ceux obtenus par Alén *et al* en 1997. En effet nous avons utilisé un tampon différent afin de permettre à la topoisomérase 1 de fonctionner. C'est donc après action de la topoisomérase 1 que le plasmide est le plus relâché, et se retrouve dans sa position la plus basse sur le gel (figure 93), et que l'ADN surenroulé non traité présente plusieurs bandes sur le gel de chloroquine. C'est une des différences

majeures avec les résultats d'Alén *et al* (1997) (figure 93). Lorsque chaque protéine a été séparément rajoutée, les protéines PepA, ArgRWt, ArgR149 et ArgR5aa provoquent des changements de topologie visibles, avec une augmentation des surenroulements négatifs, confirmant qu'il y a bien liaison de ces protéines sur l'ADN et une protection apportée à l'ADN contre l'action de la topoisomérase 1, ce qui semble à première vue contradictoire avec les résultats d'Alén *et al* (1997). Mais n'oublions pas que le tampon utilisé pour la réaction impliquant la topoisomérase 1 est différent. De plus, mes protéines ArgRWt, ArgR149 et ArgR5aa sont fusionnées avec 6-HIS et ont été purifiées sur colonne Ni-NTA.

Dans le gel avec chloroquine, PepA a empêché un peu l'action de la topoisomérase 1 sur l'ADN, entraînant tout de même une certaine protection qui n'avait pas été précédemment visualisée, avec plus de surenroulements négatifs. Il est important de souligner que les résultats de la figure 72 (un retard sur gel) semblaient montrer que PepA ne se liait pas seul à l'ADN, mais les résultats de Minh *et al* (2009) et ceux de la figure 93 montrent que PepA peut très bien se lier à l'ADN. Comment de telles différences sont possibles ? En fait les tampons utilisés dans ces deux expériences sont très différents et le taux de pureté de PepA est meilleur dans le cas des expériences de topologies. De plus, les expériences de topologies permettent mieux de visualiser des formes transitoires du plasmide en présence de la protéine PepA, que les expériences de retards sur gel. Et ce que la protéine se fixe moins bien à l'ADN qu'ArgR ou non. De plus les deux séquences *cer* ont une façon de migrer très différente de la boîte arginine, ce qui peut rendre la visualisation des retards difficile avec PepA.

De plus, ArgRWt protège plus l'ADN de l'action de la topoisomérase 1 que PepA (figure 93). Les profils obtenus avec ArgR149 et ArgR5aa sont comparables à ceux d'ArgRWt. Ces données semblent indiquer que les différentes formes d'ArgR se fixent mieux à l'ADN que PepA.

Lors d'un mélange ArgRWt avec PepA, l'ADN a été plus relâché qu'avec PepA seul, voire avec la topoisomérase seule, ce qui montre que lorsque PepA peut former un

complexe avec ArgR, PepA et ArgRWt ensemble semblent moins bien protéger l'ADN que lorsqu'elles sont seules. Ce qui peut s'expliquer par les effets de PepA sur l'ADN. En effet PepA entraîne la formation de « tours » d'ADN autour du complexe ArgR-PepA qui sont en sens inverse, ne protégeant plus l'ADN de la topoisomérase 1 (figures 47, 48 et 49).

Pour ce qui est des formes mutantes ArgR149 et ArgR5aa, les profils sur les deux types de gels ont indiqué qu'il y avait protection contre l'action de la topoisomérase 1 comme avec le type sauvage, car les mutants sont capables de se fixer à l'ADN. Par contre, un plus important surenroulement négatif de l'ADN semble avoir été remarqué avec les mutants d'ArgR qu'avec PepA. Ce qui a montré que les trois formes d'ArgR apportent cette protection à l'encontre de l'action de la topoisomérase 1 sur l'ADN.

Lorsque PepA est la seule protéine mise en présence de l'ADN, peu de bandes ont été détectées, ce qui montre moins de relaxation du plasmide par rapport aux protéines ArgR (en effet PepA se lie moins facilement à l'ADN que ArgR). Il a été également visualisé que lorsque le plasmide pNCK1 est mis en présence à la fois des protéines PepA et ArgR (versions sauvage et mutantes), des profils différents par rapport à la version sauvage ont été obtenus avec les versions mutantes d'ArgR.

Lorsque les deux protéines ArgRWt et PepA ont été mises en même temps, les profils semblent indiquer une prévalence de l'effet topologique de PepA sur ArgRWt, donnant un profil très relâché pour le gel, plus qu'avec PepA seule (figure 93). Quoiqu'il en soit, lorsque l'on met les formes mutantes d'ArgR (les formes ArgR149 et ArgR5aa) avec la protéine PepA, les profils ressemblent plus à ceux de l'ADN avec la protéine PepA seule. En effet, les profils de relaxation montrent plus de surenroulements négatifs et se rapprochent de ceux des expériences où l'ADN est uniquement mis en présence de PepA. Ces deux derniers profils ont montré que si les mutants sont toujours capables de se lier à l'ADN, ils n'interagissent pas de la même façon avec PepA qui semble mieux protéger l'ADN.

L'apport de ces expériences sur la topologie de l'ADN est le suivant, les mutants ArgR149 et ArgR5aa interagissent moins avec PepA. Ces derniers peuvent par contre toujours conserver tout ou partie de leur effet topologique sur pNCK1

Ces expériences de topologies montrent de plus l'effet de PepA (se fixant également de manière hexamérique) sur les profils de migration, notamment lorsque les protéines ArgR et PepA sont présentes en même temps. Cela montre que lorsque PepA peut former un complexe avec ArgRWt et l'ADN, PepA empêcherait ArgRWt de donner à l'ADN la même topologie que si ArgRWt était seule. En effet, comme dit plus haut, lorsque les deux protéines ArgR et PepA sont présentes, on a vu des profils qui semblaient plus relâchés que ceux de PepA seul avec l'ADN. Il est tout à fait concevable et probable que lorsque PepA se lie à l'ADN et à ArgRWt, le ou les hexamères de PepA imposeraient leur effet topologique sur l'ADN même si ArgRWt est présent et lié au plasmide pNCK1, rendant l'ADN plus accessible à la topoisomérase 1 lors de ces expériences, car PepA empêcherait ArgR de changer la topologie de l'ADN de la même manière que si l'hexamère d'ArgRWt était seul. Il y a donc des effets antagonistes entre les protéines PepA et ArgRWt sur la topologie de l'ADN.

Si les expériences de topologie de l'ADN et de recombinaison *in vitro* ont montré que PepA se liait à l'ADN sans que la présence d'ArgR soit nécessaire à cette liaison (figure 93), on peut tout de même avancer qu'un hexamère d'ArgR fonctionnel est nécessaire à la recombinaison (figure 90). Même si les résultats de ces expériences ne permettent pas de trancher entre le modèle de Sträter modifié et le modèle proposé par Reijns *et al* (2005), ils permettent de dire que la présence d'ArgR est requise quel que soit le modèle, la formation du complexe ArgRWt-PepA permettant à l'ADN d'adopter probablement une topologie favorable à la recombinaison. On peut toutefois avancer, vu ces résultats, qu'ArgR serait la première protéine à se fixer sur l'ADN et que PepA viendrait ensuite se fixer sur la séquence par un recrutement dû à ArgR, le complexe ainsi formé serait alors stabilisé par ArgR. De plus, le mutant ArgR149, s'il est toujours capable de se lier à l'ADN (figure 93), ne permet pas la formation de

jonction d'Holliday ni recombinaison. Les résultats apportés par les *crosslinking* au glutaraldéhyde (figure 75) montrent que les hexamères d'ArgR149 sont minoritaires, ce qui devrait empêcher la formation d'un complexe de recombinaison, même si PepA pourrait se fixer. De plus, le mutant ArgR149 n'ayant pas le même effet topologique qu'ArgRWt, cela contribuerait également à empêcher la recombinaison de se faire. Pour se qui est d'ArgR5aa, on a pu distinguer la formation de jonction d'Holliday (figure 87) et cette protéine est tout à fait capable de s'hexamériser (figure 74). De plus, ArgR5aa peut très bien se fixer sur l'ADN (figure 93). Le problème de résolution des jonctions d'Holliday, donc de la recombinaison, pourrait provenir d'une modification incorrecte de la topologie de l'ADN qui est normalement apportée par le complexe ArgRWt-PepA.

Lors d'une expérience de substitution menée au laboratoire, un mutant nommé ArgR180 a été accidentellement obtenu lors d'une tentative pour obtenir le mutant ArgRL149A. Ce mutant ArgR180 possède certes la substitution en position 149 de la protéine ArgR, mais un glissement de cadre de lecture a engendré une protéine longue de 180 acides-aminés au lieu de 156 comme c'est le cas de la forme sauvage d'ArgR. Il y a eu donc un ajout de 24 acides-aminés. Ce qui est important de signaler avec ce mutant ArgR180, c'est qu'il serait toujours capable d'effectuer la recombinaison au niveau des deux sites *cer* de pCS210. Donc une augmentation de la longueur de la protéine après la position 151 ne semble pas avoir d'effets trop importants pour le phénomène de recombinaison (figure 97).



Lorsque l'on regarde les résultats des expériences de recombinaison *in vitro* et de topologie avec les expériences de crosslinking et que l'on regarde la sixième hélice  $\alpha$ , on peut proposer les interprétations suivantes.

```

ArgRWt           TPANGFTVKDLYEAILLELFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

ArgR5aa          : TPANGFTVKDLYEAILGVPLLELFDQEL
Jnet             : -----HHHHHHHHH---HHHHH-----
jhmm            : -----HHHHHHHHH---HHHHH-----
jpssm           : -----HHHHHHHHH---HHHHH-----

ArgR149          : TPANGFTVKDLYEAIL
Jnet             : -----HHHHHHHH--
jhmm            : -----HHHHHHHH--
jpssm           : -----HHHHHHHH--

```

**Figure 99.** Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgR5aa et ArgR149. Pour chaque protéine un alignement Jnet, jhmm et jpssm a été réalisé. On peut voir la séquence d'une hélice  $\alpha$  indiquée par une lettre H. On peut constater l'hélice  $\alpha$  d'ArgRWt. Pour ArgR5aa, on voit que cette hélice est séparée et qu'elle reprend après le résidu proline (P). Pour ArgR149, l'hélice est tronquée.

Comme on peut le voir sur la figure 99, Le programme Jpred3 montre l'hélice  $\alpha$  6 comme prédit et connue par les modèles et la cristallographie de cette partie Ct pour ArgRWt. Par contre, pour ArgR5aa, cette hélice a été écourtée par l'insertion des cinq acides-aminés, mais elle reprend après le résidu proline, créant une septième hélice  $\alpha$ . On peut donc se demander si tous les problèmes de recombinaison rencontrés avec cette version mutante d'ArgR serait dut à cette interruption de l'hélice sur trois acides-aminés (GVP) (mais pas forcément à un encombrement stérique) donnant une certaine flexibilité qui n'existait précédemment pas dans la région qui est sensée interagir avec PepA. Pour ArgR149, on voit très bien que l'hélice a été écourtée, ce qui peut expliquer la perte de la capacité à réaliser la recombinaison au niveau de deux sites *cer* directement répétés. La question de l'importance de la structure et de la longueur de la sixième hélice  $\alpha$  par rapport à la composition en acides-aminés de la partie Ct peut donc se poser.

```

ArgRWt           TPANGFTVKDLYEAILLELFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

ArgRE150A       : TPANGFTVKDLYEAILALFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

ArgRL149A       : TPANGFTVKDLYEAI AELFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

ArgRL151A       : TPANGFTVKDLYEAILLEAFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

ArgRL149AL151A : TPANGFTVKDLYEAI AEAFFDQEL
Jnet             : -----HHHHHHHHHHHHHH-----
jhmm            : -----HHHHHHHHHHHHHH-----
jpssm           : -----HHHHHHHHHHHHHH-----

```

**Figure 100.** Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgRE150A, ArgRL149A, ArgRL151A et ArgRL149AL151A. Pour chaque protéine un alignement Jnet, jhmm et jpssm a été réalisé. On peut voir la séquence d'une hélice  $\alpha$  indiquée par une lettre **H**. On peut constater que pour chaque protéine permettant toujours la recombinaison, la structure en hélice  $\alpha$  est conservée et possède une longueur similaire.

La figure 100 montre les protéines mutantes obtenues par substitution et toujours capables de permettre la recombinaison au niveau de deux séquences *cer* répétées directement. On peut voir que pour chaque protéine obtenue par substitution d'un ou deux résidus leucine, la structure reste très semblable, ce qui permettrait toujours la recombinaison comme il a été observé. Il en est de même pour la substitution de l'acide glutamique en position 150. On peut donc supposer l'importance de la structure en hélice  $\alpha$  en Ct. Mais comme on peut le voir sur la figure 101, la restitution d'un seul acide-aminé en position 150 a permis de reconstituer la capacité de recombinaison de la protéine pour ArgR150 et il en est de même pour ArgR151. On peut donc constater que



si dans les protéines tronquées ArgR150 et ArgR151, la longueur de l'hélice  $\alpha$  est plus courte que pour ArgRWt, elle est à chaque fois plus grande, même d'un acide-aminé.

```

ArgRWt          TPANGFTVKDLYEAILELFDQEL
Jnet           : -----HHHHHHHHHHHHHH-----
jhmm          : -----HHHHHHHHHHHHHH-----
jpssm         : -----HHHHHHHHHHHHHH-----

ArgR149        : TPANGFTVKDLYEAIL
Jnet           : -----HHHHHHHH-----
jhmm          : -----HHHHHHHH-----
jpssm         : -----HHHHHHHH-----

ArgR150        : TPANGFTVKDLYEAILE
Jnet           : -----HHHHHHHHHH-----
jhmm          : -----HHHHHHHHHH-----
jpssm         : -----HHHHHHHHHH-----

ArgR151        : TPANGFTVKDLYEAILEL
Jnet           : -----HHHHHHHHHHHH-----
jhmm          : -----HHHHHHHHHHHH-----
jpssm         : -----HHHHHHHHHHHH-----

```

**Figure 101.** Analyse Jpred3 de la séquence Ct des protéines ArgRWt, ArgR149, ArgR150, ArgR151. Pour chaque protéine un alignement Jnet, jhmm et jpssm a été réalisé. On peut voir la séquence d'une hélice  $\alpha$  indiquée par une lettre **H**. Mis à part la protéine ArgR149, toutes les autres protéines présentées dans cette figure sont capable de générer une recombinaison en *cer*.

Cet acide-aminé de différence dans l'hélice  $\alpha$  semble suffisant pour rétablir le phénomène de recombinaison pour ArgR150. On peut donc s'interroger sur le rôle des acides-aminés dans l'établissement de la structure en hélice  $\alpha$  rencontrer dans la partie Ct, et sur l'importance de leur nature. De plus, quand on compare la partie Ct d'ArgR et d'ArgRBst, on peut se poser la question de l'importance des résidus leucine pour ArgR par rapport à ceux présents en Ct d'ArgRBst et qui stabilisent l'hexamère d'ArgRBst. Sachant qu'ArgRBst est issu d'un organisme thermophile (Dennis *et al*, 2001), il se peut qu'ArgR n'ai pas besoin d'être aussi stable, mais la région homologe entre ArgR et ArgRBst en Ct mérite tout de même d'être étudiée. Lorsque l'on compare les protéines ArgR151 et AhrC, on peut constater que la séquence est quasiment identique, il ne manque qu'un résidu leucine à la fin du mutant pour retrouver la séquence d'AhrC

dans cette région, le triplet LEL a même été reconstitué. Ce qui laisse à penser que les structures entre ArgR151 et AhrC au niveau de l'hélice  $\alpha$  6 sont comparables d'où la complémentarité permise par AhrC chez E.coli (Stirling *et al*, 1988b).

**CHAPITRE V**  
**CONCLUSION ET PERSPECTIVES**

## CHAPITRE V. CONCLUSION ET PERSPECTIVES

### V.1.Récapitulatif des résultats

Les expériences ont confirmé les résultats suivants déjà obtenues par des expériences antérieures à ma thèse.

1) Il y a une importance de la région Ct de la protéine ArgR dans la recombinaison site *cer* spécifique, de part l'incapacité des mutants ArgR149 et ArgR5aa à effectuer une recombinaison *in vivo*.

2) Les régions mutées dans ArgR149 et ArgR5aa n'affectent pas la capacité de ces protéines à réprimer une fusion *argA :: lacZ in vivo*.

3) Les protéines ArgR149 et ArgR5aa se lient à la boîte arginine.

Par contre, mes expériences ont apporté les nouvelles données et renseignements suivants :

1) Il existe des problèmes d'hexamérisation pour le mutant ArgR149, peut-être dus aux acides-aminés touchés par la délétion. La majorité de ces protéines ArgR149 montre des formes dimériques et tétramériques et non hexamériques. Les résultats pour ArgR149 mettent en lumière ceux pour les retards sur gel de la même protéine, on aurait besoin d'une plus grande quantité de cette protéine pour avoir un retard de la boîte arginine ou de *cer*. Des problèmes de stabilité ont été devinés avec le complexe ArgR149-ADN. Pour le mutant ArgR5aa, on voit qu'il est aussi bien capable de se lier à la séquence *cer* qu'ArgRWt. De plus, ArgR5aa est tout aussi capable de s'hexamériser qu'ArgRWt.

2) PepA est capable d'interagir avec toutes les formes d'ArgR à une distance de 12 Angströms, que se soient avec les formes trimériques d'ArgR ou ses formes hexamériques. Le *crosslinker* choisi, à savoir le DSP, peut très bien avoir stabilisé les liaisons les plus instables comme celles avec ArgR149.

3) La forme mutante ArgR149 est incapable de permettre une réaction de recombinaison site spécifique menant à la formation d'une jonction d'Holliday, et la forme ArgR5aa permet la formation d'une jonction d'Holliday, mais a une efficacité réduite. La jonction d'Holliday formée par ArgR5aa ne peut pas être résolue pour générer les monomères. Ce qui permettrait aussi de dire que toutes les protéines du complexe ArgR-PepA-XerC-XerD-ADN doivent être fonctionnelles pour réaliser la recombinaison site spécifique en *cer*.

4) ArgR provoque un changement de topologie du plasmide pNCK1, lorsqu'il s'agit du type sauvage ou des formes mutantes ArgR149 et ArgR5aa. Lorsqu'il y a présence de PepA avec ArgR, il a été observé que l'ADN était au moins aussi relâché qu'avec la topoisomérase 1 seule, ce qui serait dû à l'effet topologique du complexe ArgR-PepA qui rend l'ADN accessible à la topoisomérase 1. Par contre, les profils de migration de l'ADN avec les mutants combinés à PepA montrent un patron de topologie différent, en effet les patrons de migration sont plus surenroulés négativement. On peut donc proposer que PepA soit capable de se lier à l'ADN au niveau des séquences *cer* sans l'intervention d'ArgR. Les deux formes mutantes ArgR149 et ArgR5aa ne pouvant visiblement pas générer la même topologie avec PepA car elles se lient moins facilement avec PepA, mais toujours à l'ADN. Ces dernières expériences permettent de suggérer que lors de la formation du complexe de recombinaison, ArgR serait la première protéine à se fixer sur l'ADN (comme dans le modèle de Sträter), et il y aurait un recrutement de PepA suivi d'une stabilisation du complexe. Mais, même si l'ordre de fixation des protéines rappelle le modèle de Sträter à deux hexamères de PepA, on ne peut totalement exclure qu'il n'y ait qu'un seul hexamère recruté par ArgR.

Si les résultats et leurs analyses nous ont fourni de nombreuses indications sur les zones d'interactions entre ArgR et PepA, d'autres expériences peuvent être planifiées en vue de répondre à certaines questions que les résultats peuvent soulever sur cette interaction entre ArgR et PepA nécessaire à la formation du complexe de recombinaison au niveau du site *cer*.

## V.2.Mutagenèse d'ArgR.

Nous pourrions générer des mutants d'ArgR dont le triplet LEL a été remplacé par trois alanines (LEL → AAA), ce qui donnerait alors la séquence suivante en position Ct, AIAA<sup>AAA</sup>FDQEL. Le but de cette expérience serait de voir si la recombinaison au niveau de deux sites *cer* aurait toujours lieu, ce qui nous renseignerait sur l'importance des deux résidus leucines voir du résidu isoleucine I en position 148.

On pourrait également envisager de substituer une leucine et une alanine en même temps, en position 149 et 150 (ce qui donnerait la séquence AIAA<sup>LL</sup>FDQEL). Ces expériences permettront de voir l'importance du triplet LEL d'ArgR<sup>Wt</sup> et peut-être, de définir une fonction précise de ce triplet, voire des acides-aminés qui le composent. Il serait également tentant de faire un mutant où tous les résidus entre les positions 148 et 151 incluses seraient substitués par des résidus alanine (ce qui donnerait la séquence AAAAA<sup>LL</sup>FDQEL).

De plus les mutants ArgRL149A et ArgRL151A permettent toujours la recombinaison comme la forme sauvage de la protéine ArgR. Pour AhrC, c'est une leucine qui précède le triplet LEL. Toutefois, une seule substitution en position 149 ou 150 n'a pas empêché la recombinaison.

ArgR5aa est incapable de faire la recombinaison qui semblait indiquer que les 5 acides aminés ont changés suffisamment la structure à ce niveau de la protéine pour la rendre inactive pour le phénomène de recombinaison, ou que certains résidus tel que la

proline (qui fait partie des 5 acides aminés rajoutés) entraînent un encombrement stérique qui empêche à la protéine d'être active pour la recombinaison. Il se peut que les deux phénomènes soient présents.

D'autres mutants restent à faire à partir d'ArgR5aa dont la séquence en Ct est AILGVPLLELFDQEL. Ces mutants qui restent à générer à partir du gène donnant ArgR5aa nous permettraient de voir lesquels d'entre eux seraient capables ou non de générer la recombinaison site spécifique en *cer* comme ArgRWt.

Le tableau suivant montre une liste de mutants réalisables à partir d'ArgR5aa.

Protéine	Séquence Ct de la protéine
ArgRWt	AILELFDQEL
ArgR5aa	AILGVPLLELFDQEL
Autres mutants ArgR à générer à partir d'ArgR5aa	AILGVPELFDQEL AILGVPELFDQEL AILGVELFDQEL AILGELFDQEL AILVPLLELFDQEL AILPLLELFDQEL AILLELFDQEL AILLELFDQEL

**Tableau 14.** Présentation de mutants à générer à partir d'ArgR5aa. Les protéines mutantes sont générées par délétions successives à partir du mutant ArgR5aa.

Lorsque l'on regarde la protéine ArgR de *B.stearothermophilus*, on peut envisager d'autres expériences qui viseraient cette fois la partie Nt de la protéine ArgR d'*E.coli*.

En effet, lorsque l'on regarde la partie Nt de ArgRBst, plusieurs acides-aminoés sont sensés interagir avec la partie Ct de la protéine, en L146 et L149, pour stabiliser l'hexamère (Ni *et al*, 1999). Ces acides-aminoés en Nt, compris dans la première hélice

$\alpha$ , sont les suivants chez ArgRBst, I8, R11 et M15 (figure 102) et avaient été mis en avant dans la figure 42.

ArgRNt	MRSSAKQEEELVKAFKALLKEEKFFSS
ArgRBstNt	----LNKGQRHIKIREIIMSNDIET
AhrCNt	----MNKGQRHIKIREIITSNEIET

Hélice  $\alpha$  1

**Figure 102.** Alignement de séquence Nt entre les protéines ArgR d'*E.coli* (ArgRNt) de *B. stearothermophilus* (ArgRBstNt) et de *B.subtilis* (AhrCNt). Ce schéma représente un alignement des parties Nt des deux protéines homologues peu après la fin de la première hélice  $\alpha$ 1. Les résidus en bleu correspondent à ceux compris dans l'hélice  $\alpha$ 1 qui interagissent avec la partie Ct de *B. stearothermophilus*. Les résidus de la même hélice  $\alpha$ 1 de *B.subtilis* sont également montrés en orange. Les résidus en violet d'*E.coli* sont ceux qui pourraient avoir le même rôle dans la stabilisation des hexamères (l'alignement a été obtenu par le programme ClustalW, et inspiré de Dennis *et al*, 2001).

Étant donné qu'il y a une différence de séquence entre les protéines ArgR issues de *B.stearothermophilus* et d'*E.coli* (Dennis *et al*, 2001), les résidus compris dans l'hélice  $\alpha$ 1 d'*E.coli* qui pourraient éventuellement générer le même type d'interaction (malgré la différence de séquence) pour la stabilisation des hexamères d'ArgR chez *E.coli*, sont peut-être en Nt L10, V11 et K15, en sachant que L10 et V11 chez *E.coli* pourrait avoir le même rôle que le résidu I8 chez *B.stearothermophilus*, et que K15 chez *E.coli* pourrait avoir le même rôle que le résidu R11 chez *B.stearothermophilus*, et tout cela en raison de leur structure. À noter que M15 n'a pas d'équivalent chez *E.coli* ni chez *B.subtilis*. Ces mutations restent à tenter, même si seulement la partie Ct d'ArgR a été cristallisée de façon hexamérique et non la partie Nt. Des délétions ou des substitutions par l'alanine au niveau de ces résidus compris dans l'hélice  $\alpha$ 1 d'ArgR sont envisageables, car l'article de Sunnerhagen *et al* (1997) suggérait déjà cette interaction Nt-Ct, et c'est dans cette structure que se font plusieurs des interactions chez *B.stearothermophilus* (Ni *et al*, 1999). Il est important de souligner que dans les travaux de Tian *et al*, 1994, la valine en position 11 avait été remplacée par l'isoleucine dont la



structure est proche de la valine et un mutant super-répresseur avait été obtenu, mais une substitution par l'alanine n'a pas été tentée. Une autre donnée qui nous permet de choisir cette région, est que chez la protéine de *B.subtilis*, AhrC, qui est capable de compléter les *E.coli*, possède la même longueur qu'ArgRBst, c'est-à-dire 149 acides-aminés, et possède les mêmes résidus en position 8 et 11, soit I8 et R11 (figure 102). De plus la partie Ct d'AhrC est identique pour les résidus 146 et 149, soit L146 et L149 (Dennis *et al*, 2001). Il est de plus important de signaler que des protéines chimériques réalisées à partir d'AhrC et ArgR avaient déjà montré qu'elles pouvaient toujours se lier à la boîte arginine (Holtman *et al*, 1999), que ce soit la partie Nt d'ArgR avec la partie Ct d'AhrC ou inversement. Mais ce que l'on envisage de faire est quelque peu différent. En effet, on désire faire des mutations précises dans l'hélice  $\alpha 1$  pour distinguer les résidus importants dans la formation des hexamères d'ArgR.

Même si toutes ces mutations sont envisageables, il ne faut pas oublier que la partie Nt est responsable de la liaison d'ArgR à l'ADN, il faudra vérifier la liaison de ces mutants à la boîte arginine par retard sur gel, s'ils présentent une déficience de recombinaison d'un plasmide pCS210. Si on n'a pas de retard de la boîte arginine, on aura généré non pas un problème d'interaction Nt-Ct, mais un problème d'interaction entre la protéine mutante et la séquence boîte arginine. Ceci étant, il est probable que l'interaction avec l'ADN ne soit pas touchée, car si on touche les résidus présents dans l'hélice  $\alpha 1$ , ce sont les résidus de l'hélice  $\alpha 3$  qui auraient une interaction avec l'ADN selon les modèles développés avec ArgRBst, AhrC et la structure ArgR déterminées par spectroscopie (Sunnerhagen *et al*, 1997). On pourrait également faire des substitutions et délétions d'I8 et R11 chez AhrC qui ont probablement la même fonction, et compléter des souches d'*E.coli* déficientes en ArgR par ces mutants et voir si la capacité de recombinaison est toujours présente, avec un plasmide pCS210 comme pour les expériences précédemment citées avec ArgR d'*E.coli*. Bien évidemment, des retards sur gel avec la séquence boîte arginine doivent être réalisés pour vérifier si les protéines mutantes sont toujours capables de se lier à l'ADN.

Comme signalé précédemment, il y a de nombreuses similitudes entre les protéines ArgR d'*E.coli*, AhrC de *B.subtilis* et ArgRBst de *B.stearothermophilus*. En effet entre AhrC et ArgRBst, les acides-aminés situés en position 146 et 149 sont les mêmes, et des leucines dans chaque cas (figure 103), (Dennis *et al*, 2001 ; Ni *et al*, 1999). Outre les mutations envisagées en Nt, on peut aussi encore envisager de réaliser des mutations dans la partie Ct de la protéine AhrC pour voir l'importance des résidus cités plus haut dans la formation de l'hexamère. Les mutations à envisager seraient de type délétions et substitutions par l'alanine sur les leucines en position 146 et 149, et on chercherait à voir leurs effets sur la recombinaison de pCS210 dans des souches d'*E.coli* déficientes en ArgR, complémentées par ces AhrC mutantes. Il serait intéressant de faire ces mêmes expériences de délétion et de substitution sur la l'isoleucine en position 148 d'ArgR, car si on compare l'alignement des séquences entre ArgR et AhrC, c'est cette isoleucine qui pourrait peut-être remplir le rôle de la leucine en position 146. Il est vrai que les résultats présentés dans cette thèse avec ArgR149 laissent entrevoir les résidus impliqués dans la stabilité de l'hexamère d'ArgR (figure 75). De plus, une substitution de la leucine en position 145 chez AhrC serait à envisager, pour vérifier si cette dernière pourrait éventuellement jouer un rôle dans la stabilisation de l'hexamère d'AhrC, donc dans la recombinaison en *cer*. Il serait aussi important de voir si les versions mutantes d'AhrC en Ct sont toujours capables de se lier à l'ADN par un retard sur gel avec la séquence boîte arginine.

ArgR	DTI FTTPANGFTVKDL YEAI L E L F D Q E L
AhrC	DTIL I I C R T P E D T E G V K N R L E L L - - - -
ArgRBst	DTCL I I C R T P K D A K K V S N Q L L S M L - - - -

Hélice  $\alpha 6$

**Figure 103.** Alignement des séquences des hélices  $\alpha 6$  des protéines ArgR d'*E.coli*, d'AhrC de *B.subtilis* et ArgRBst de *B.stearothermophilus*. Les acides-aminés qui sont situés au même endroit avec des fonctions similaires sont en orange, en bleu, ceux en commun aux trois protéines. En violet, les acides-aminés en commun entre ArgR et AhrC. En vert, on a les acides-aminés en commun entre AhrC et ArgRBst (l'alignement a été obtenu par le programme ClustalW, et inspiré de Dennis *et al*, 2001).

De plus, si des protéines chimériques comprenant la partie Nt d'AhrC et Ct d'ArgR soit la partie Nt d'ArgR et la partie Ct d'AhrC avaient déjà été générées (Holtham *et al*, 1999) et montraient que les chimères obtenues étaient toujours capables de se lier à des séquences de types boîtes arginine et de s'hexamériser. Par contre, leurs effets sur la recombinaison *in vivo* n'ont pas été testés, et c'est une chose qui serait envisageable de faire avec le plasmide pCS210 dans des *E.coli* déficientes en ArgR. On pourrait aussi tenter de réaliser des chimères avec seulement l'hélice  $\alpha 1$  ou l'hélice  $\alpha 6$  de modifiée. On aurait donc AhrC avec soit l'hélice  $\alpha 1$  d'ArgR ou l'hélice  $\alpha 6$  d'ArgR, soit ArgR avec l'hélice  $\alpha 1$  d'AhrC ou l'hélice  $\alpha 6$  d'AhrC. Il faudrait ensuite regarder les recombinaisons *in vivo* et *in vitro* permises par ces chimères, et faire des retards sur gel de la boîte arginine avec les différentes chimères obtenues.

### V.3.Mutagenèse de PepA.

Si on envisage des mutations sur ArgR, on peut très également tenter de faire des mutations sur PepA. En effet, de nombreuses mutations ont déjà été faites sur cette partie (Reijns *et al*, 2005), mais les études menées par ce groupe ont pu montrer que les acides-aminés que l'on considérait précédemment responsables de l'interaction ArgR-PepA ne jouaient finalement pas le rôle que l'on pensait. Par contre, les mutations générées par cette équipe ont montré que les résidus I217, L338, R484, P485 et G483, s'ils subissaient une substitution (dans ce cas, I217N, L338S, R484G et G483D), entraînaient perte de la recombinaison en *psi* impliquant les protéines ArcA et PepA. On peut donc tenter de faire dans un premier temps des substitutions ou des délétions dans les acides-aminés comprises entre les positions 330 et 503 de la protéine PepA. Sachant que la plupart des mutations citées impliquent les résidus 483 à 485, on pourrait donc réaliser des mutations dans la région de la quinzième hélice  $\alpha$ , région qui suit ces résidus précédemment substitués. Les résidus mutés sont en fait compris entre l'hélice  $\alpha 15$  et le feuillet  $\beta 16$ . Cette région comprise entre ces deux structures de PepA mérite aussi notre attention. Il est en effet probable que ces régions comprennent des résidus responsables de l'interaction entre ArgR et PepA. Il faudrait aussi penser à muté

des résidus dans les feuillets  $\beta 11$  et  $\beta 12$ , car cette région est rejeté en périphérie de l'hexamère et pourrait interagir avec ArgR (Sträter *et al*, 1999). On peut se demander si les hexamères de PepA subissent une modification de conformation lorsqu'ils entrent en contact avec l'hexamère d'ArgR voire avec l'ADN (les deux phénomènes sont possibles en même temps). En effet, si l'étude de Reijns *et al* (2005) ne présente pas de déformation des hexamères, les résultats sur AhrC et *Mtb*ArgR (Garnett *et al*, 2010 ; Cherney *et al*, 2010) laissent présager la même chose pour ArgR, donc on peut légitimement se poser la question pour PepA. De plus mes résultats pour les versions tronquées de PepA (PepACtMBP et PepANtMBP) suggèrent soit que les deux régions de PepA sont nécessaires à l'interaction avec ArgR (qu'il y ait changement de conformation ou non), soit que le fait d'avoir tronqué la protéine ait rendu totalement inactive chacune des deux parties, ou encore que la partie MBP bloque les interactions entre les deux parties de la protéine (ou la partie MBP rend ces interactions incorrectes dans le cadre de l'activité aminopeptidase). On peut aussi penser au fait qu'avoir tronqué la protéine ait engendré des problèmes de repliement des régions de PepA. La région centrale de la protéine PepA, avec l'hélice  $\alpha 4$ , pourraient très bien avoir des interactions avec les régions Ct et Nt (Sträter *et al*, 1999). On pourrait aussi envisager des mutations au sein de cette hélice  $\alpha$ , qui pourrait rentrer en contact avec ArgR ou stabiliser l'hexamère de PepA. Les mutations à envisager dans ce dernier cas seraient des substitutions de résidus par l'alanine. On pourrait également tenter de connaître le possible changement de conformation de l'hexamère de PepA évoqué plus haut en analysant la fixation de cet hexamère avec ses séquences de fixation PEPA1 et PEPA2 au sein du promoteur *carAB* après co-cristallisation. Le protocole à suivre serait inspiré de celui pour *Mtb*ArgR décrit dans Cherney *et al* (2010). On pourrait utiliser cette technique en associant la protéine PepA avec l'ADN (notamment la séquence *cer*) et successivement avec les protéines ArgRWt, ArgR5aa et ArgR149, voire ArgR150 pour la co-cristallisation. On pourrait également tenter une co-cristallisation des protéines ArgRWt et PepA, avec ou sans ADN. Le problème étant la cristallisation complète d'ArgR d'*E.coli* qui n'a pas encore été réalisée.

#### V.4. Autres expériences

Pour de futur *crosslinking*, il serait possible de tenter de les réaliser avec le DST (Disuccinimidyl tartarate), un *crosslinker* avec un « bras » de 6,4 Angströms. Il serait intéressant de vérifier quels sont les effets de ce cross-linker sur les formes d'ArgR déjà testées (à savoir, ArgRWt, ArgR149, ArgR5aa, ArgRL149A et ArgRL151A). En effet, les profils et intensités des bandes pourraient être différents de ceux déjà visualisés au DSP (Dithiobis succinimidyl propionate).

Une autre possibilité est de tenter une trypsinisation de PepA, une digestion partielle, voire totale, libèrera des fragments qui devraient garder leur structure tridimensionnelle. Les fragments seront, séparément, mis en présence d'ArgRWt-6-HIS et subiront un *crosslinking* réversible au DSP. S'il y a interaction entre les fragments de PepA et l'ArgR fusionné avec 6-HIS, on pourra récupérer les fragments par billes magnétiques Ni-NTA, et les purifier par SDS-PAGE pour ensuite les identifier par spectrométrie de masse. Une autre façon de procéder avec la trypsinisation, est de comparer les patrons de migrations SDS-PAGE des fragments protéiques obtenus avec ArgR et PepA séparément, puis lorsque ArgR et PepA sont mélangées (avec L-arginine). La trypsine pourrait ne pas cliver les endroits où les protéines sont en contact, donnant un patron de migration des fragments obtenu différent (technique dite de *protein footprinting*).

Pour trouver la partie de PepA qui pourrait interagir avec ArgR, on peut tenter la technique de *phage display* (New England Biolabs<sup>TM</sup>) où une librairie de peptides exprimés à la surface de phages serait mise en présence de la protéine ArgR fixée sur une plaque à 96 puits. Après passage d'une librairie de peptides (exprimés à la surface du phage), les phages qui sont capables d'interagir avec ArgR sont élués et recueillis et on répète plusieurs fois l'opération avec d'autres librairies. Les phages ainsi recueillis sont amplifiés et analysés sur trois « rondes » par séquençage d'ADN. Les peptides ainsi retrouvés seront comparés à la séquence de PepA et on pourra déterminer quelles séquences interagissent avec ArgR et où elles se trouvent sur la protéine PepA.

Un autre type d'expérience mettrait en application la technique de centrifugation analytique avec les protéines ArgR149 et ArgR5aa (cette technique ayant été employée sur ArgRWt par Holtham *et al*, 1999), ce qui permettrait de connaître de possibles dissociations plus rapides des hexamères d'ArgR149 et de vérifier de possibles changements dans la structure des mutants par rapport à ArgRWt. On peut également tenter cette technique avec PepA mise en présence respectivement des protéines ArgRWt, ArgR5aa et ArgR149.

On pourrait également envisager les techniques de FRET (*Fluorescence resonance energy transfer*) avec les protéines ArgR et PepA fusionnées chacune par un fluorophore différent (généralement des GFP ou BFP), l'un étant le donneur et l'autre l'accepteur. Cette technique consiste à observer le transfert d'énergie, c'est-à-dire une différence d'émission lumineuse, qui a lieu entre le fluorophore donneur et le fluorophore accepteur des deux protéines. Des techniques *in vitro*, sont applicables par analyse informatique de mélanges contenant les deux protéines fusionnées à un fluorophore, de nombreux types de plaques existent, dont des plaques à 96 puits pour l'analyse (Chuck *et al*, 2010; Yengo *et al*, 2010). On aurait donc l'occasion de voir si les complexes se forment bien avec les versions mutantes d'ArgR, surtout avec ArgR149, où la jonction d'Holliday n'a pas été remarquée lors des tests de recombinaison *in vitro*.

La technique de *fluorescence lifetime imaging*, une variante du FRET, pourrait être tentée si l'on veut suivre les interactions d'ArgR et PepA au sein d'une cellule d'*E.coli* (Sekar *et al*, 2003). Mais à l'heure actuelle, les microscopes à fluorescence utilisés dans les études actuelles manqueraient de résolution. Si la technique est utilisable, on visualiserait le lieu de formation des complexes de recombinaison dans la bactérie, possiblement dans le temps.

On pourrait aussi tenter d'effectuer des mutations sur la demie-boîte arginine contenue dans la séquence *cer*, par exemple, par insertion de nucléotides dans cette séquence sur laquelle se lie la protéine ArgR, mais aussi PepA, quand

cette séquence est présente en deux exemplaires. Une telle démarche nous permettrait de voir quels effets ces mutations auraient sur la recombinaison et la topologie de l'ADN.

Finalement, une autre technique à tenter serait celle de *surface plasmon resonance* (SPR). Dans les expériences qui la mettraient en pratique, une biopuce servirait de moyen de détection des interactions entre ArgR et PepA. C'est sur cette puce que serait fixée ArgR ou PepA qui pourrait ensuite capturer l'autre protéine, puisque ArgR et PepA sont censées interagir, mises en solution. Le problème de cette technique est que la protéine qui doit être capturée par la biopuce doit être biotinylée ou fusionnée avec 6-HIS pour autoriser une chélation générant une liaison covalente stabilisant le complexe à analyser. Pour le moment, seule ArgR a été fusionnée à 6-HIS sans perte de fonction. La protéine PepA hexamérisée est donc la protéine qui devrait être fixée sur la biopuce, pour être mise en présence des différentes versions d'ArgR (également multimérisées) afin de vérifier leurs interactions (avec présence ou non de L-arginine). Il est à noter que des expériences utilisant la SPR ont déjà été réalisées avec ArgRWt et AhrC en présence et absence de L-arginine avec la boîte arginine qui était liée aux biopuces (Stockley *et al*, 1998).

Au final, toutes ces expériences nous permettront de mettre en lumière les principaux résidus assurant la stabilité du complexe protéique tant au niveau de la protéine ArgR que de la protéine PepA.

Les expériences proposées plus haut pourront également nous renseigner plus en détail sur les résidus impliqués dans l'interaction entre les protéines ArgR et PepA. Il est même envisageable de suggérer que les résidus qui seront proposés pour l'interaction entre les deux protéines pourront permettre de préciser quel est le modèle d'interaction entre ArgR et PepA lors de la formation du complexe de recombinaison. Il serait également possible de vérifier si l'interaction ArgR-PepA ne pourrait pas avoir une importance dans d'autres fonctions cellulaires, notamment

au niveau de l'opéron *carAB* qui possède des sites de liaison pour les deux protéines  
au niveau de son promoteur.



## BIBLIOGRAPHIE

- Abeles, A.L., Friedman, S.A., and Austin, S.J. (1985) Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the partition region. *J Mol Biol* **185** : 261-272.
- Alèn, C., Sherratt, D.J., and Colloms, S.D. (1997) Direct interaction of aminopeptidase A with recombination site DNA in Xer site-specific recombination. *Embo J* **16** : 5188-5197.
- Arciszewska, L.K., and Sherratt, D.J. (1995) Xer site-specific recombination in vitro. *Embo J* **14** : 2112-2120.
- Arciszewska, L.K., Grainge, I., and Sherratt, D.J. (1997) Action of site-specific recombinases XerC and XerD on tethered Holliday junctions. *Embo J* **16** : 3731-3743.
- Aussel, L., Barre, F.X., Aroyo, M., Stasiak, A., Stasiak, A.Z., and Sherratt, D. (2002) Ftsk is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* **108** : 195-205.
- Barre, F.X., Aroyo, M., Colloms, S.D., Helfrich, A., Cornet, F., and Sherratt, D.J., (2000) FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. *Genes Dev* **14** : 2976-2988.
- Barre, F.X., and Sherratt, D.J., (2002). Xer site specific recombination: promoting chromosome segregation, p.149-161. In Graig, N.L., Craigie R., Gellert M., Lambowitz, A.M. (ed.), *Mobile DNA II*. ASM Press, Washington, D.C.

Bartling, D., and Weiler, E.W. (1992) Leucine aminopeptidase from *Arabidopsis thaliana*. Molecular evidence for a phylogenetically conserved enzyme of protein turnover in higher plants. *Eur J Biochem* **205** : 425-431.

Bigot, S., Corre, J., Louarn, J.M., Cornet, F., and Barre, F.X. (2004). FtsK activities in Xer recombination, DNA mobilization and cell division involve overlapping and separate domains of the protein. *Mol Microbiol* **54** (4) : 876-886.

Bigot, S., Saleh, O.A., Lesterlin, C., Pages, C., Karoui, M.E., Dennis, C., Grigoriev, M., Allemand, J.F., and Cornet, F. (2005). KOPS : DNA motifs that control *E.coli* chromosome segregation by orienting the FtsK translocase. *EMBO J.* **24** : 3370-3780.

Blaby, I.K., and Summers, D.K. (2009) The role of FIS in the Rcd checkpoint and stable maintenance of plasmid ColE1. *Microbiology* **155** : 2676-2682.

Blakely, G., Colloms, S., May, G., Burke, M., and Sharratt, D. (1991) *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *New Biol* **3** : 789-798.

Blakely, G.W., Davidson, A.O., and Sherratt, D.J. (1997) Binding and cleavage of nicked substrates by site-specific recombinases XerC and XerD. *J Mol Biol* **265** : 30-39.

Borek, E., and Ryan, A. (1958) The transfer of irradiation-elicited induction in a lysogenic organism. *Proc. Natl. Acad. Sci. USA* **44** : 374-377.

Bouet, J.Y., and Funnell, B.E. (1999) P1 ParA interacts with the P1 partition complex at parS and an ATP-ADP switch controls parA activities. *EMBO J* **18** : 1415-1424.

Bouet, J.Y., Surtees, J.A., and Funnell, B.E. (2000) Stoichiometry of P1 plasmid partition complexes. *J Biol Chem* **275** : 8213-8219.

Bregu, M., Sherratt, D.J., and Colloms, S.D. (2002) Accessory factors determine the order of strand exchange in Xer recombination at psi. *EMBO J* **21** : 3888-3897.

Burke, M., Merican, A.F., and Sherratt, D.J. (1994) Mutant Escherichia coli arginine repressor proteins that fail to bind L-arginine, yet retain the ability to bind their normal DNA-binding sites. *Mol Microbiol* **13** : 609-618.

Burley, S.K., David, P.R., Sweet, R.M., Taylor, A., and Lipscomb, W.N. (1992) Structure determination and refinement of bovine lens leucine aminopeptidase and its complex with bestatin. *J Mol Biol* **224** : 113-140.

Caldara, M., Charlier, D., Cunin, R. (2006) The arginine regulon of Escherichia coli : whole system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. *Microbiology* **152** (Pt 11) : 3343-54.

Caldara, M., Minh, P.M., Bostoën, S., Massant, J., Charlier, D. (2007) ArgR-dependent repression of arginine and histidine transport genes in Escherichia coli K-12. *J Mol Biol* **373** (2) : 251-67.

Capiaux, H., Lesterlin, C., Pérals, K., Louarn, J.M., Cornet, F. (2002) A dual role for the FtsK protein in Escherichia coli chromosome segregation. *EMBO Rep* **3** (6): 532-6

Chao, G., Shen, J., Tseng, C.P., Park, S.J., and Gunsalus, R.P. (1997) Aerobic regulation of isocitrate dehydrogenase gene (icd) expression in Escherichia coli by the arcA and fnr gene products. *J Bacteriol* **179** : 4299-4304.

Charlier, D., Gigot, D., Huysveld, N., Roovers, M., Pierard, A., and Glansdorff, N. (1995a) Pyrimidine regulation of the Escherichia coli and Salmonella typhimurium carAB operons : CarP and integration host factor (IHF) modulate the methylation status of a GATC site present in the control region. *J Mol Biol* **250** : 383-391.

Charlier, D., Hassanzadeh, G., Kholti, A., Gigot, D., Pierard, A., and Glansdorff, N. (1995b) *carP*, involved in pyrimidine regulation of the *Escherichia coli* carbamoylphosphate synthetase operon encodes a sequence-specific DNA-binding protein identical to XerB and PepA, also required for resolution of Co1E1 multimers. *J Mol Biol* **250** : 392-406.

Charlier, D., Kholti, A., Huysveld, N., Gigot, D., Maes, D., Thia-Toong, T.L., Glansdorff, N. (2000) Mutational analysis of *Escherichia coli* PepA, a multifunctional DNA-binding aminopeptidase. *J. Mol. Biol.* **302** : 411-426.

Chen, S.H., Merican, A.F., and Sherratt, D.J. (1997) DNA binding of *Escherichia coli* arginine repressor mutants altered in oligomeric state. *Mol. Microbiol.* **24** (6) : 1143-1156.

Cherney, L.T., Cherney, M.M., Garen, C.R., and James, M.N.G. (2010) Crystal Structure of the intermediate complex of the arginine repressor from *Mycobacterium tuberculosis* bound with its DNA operator reveals detailed mechanism of arginine repression. *J. Mol. Biol.* **399**: 240-254.

Chuck, C.P., Chong, L.T., Chen, C., Chow, H.F., Wan, D.C.C., and Wong, K.B. (2010) Profiling of Substrate specificity of SARS-CoV 3CL<sup>pro</sup>. *PLoS ONE.* **5** (10) e13197 : 1-7.

Colloms, S.D., Sykora, P., Szatmari, G., and Sherratt, D.J. (1990) Recombination at Co1E1 *Cer* requires the *Escherichia coli* XerC gene product, a member of the lambda integrase family of site-specific recombinases. *J Bacteriol* **172** : 6973-6980.

Colloms, S.D., Mc Culloch, R., Grant, K., Neilson, L., and Sherratt, D.J. (1996) Xer mediated site-specific recombination in vitro. *EMBO J* **15** : 1172-1181.

Colloms, S.D., Alèn, C., and Sherratt, D.J. (1998) The ArcA/ArcB two-component regulatory system of Escherichia coli is essential for Xer site-specific recombination at psi. *Mol Microbiol* **28** : 521-530.

Cornet, F., Mortier, I., Patte, J., and Louarn, J.M. (1994) Plasmid pSC101 harbors a recombination site, psi, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal Escherichia coli dif. *J Bacteriol* **176** : 3188-3195.

Cornet, F., Louarn, J., Patte, J., and Louarn, J.M. (1996) Restriction of the activity of the recombination site dif to a small zone of the Escherichia coli chromosome. *Genes Dev* **10** : 1152-1161.

Cornet, F., Hallet, B., and Sherratt, D.J. (1997) Xer recombination in Escherichia coli. Site specific DNA topoisomerase activity of the XerC and XerD recombinases. *J Biol Chem* **272** : 21927-21931.

Corre, J., and Louarn, J.M. (2002) Evidence from terminal recombination gradients that FtsK uses replicore polarity to control chromosome terminus positioning at division in Escherichia coli. *J. Bacteriol* **184** : 3801-3807.

Cotter, P.A., and Gunsalus, R.P. (1992) Contribution of the fnr and arcA gene products in coordinate regulation of cytochrome o and d oxydase (cyoABCDE and cydAB) genes in Escherichia coli. *FEMS Microbiol Lett*, **70** (1) : 31-6.

Craigie, R., and Mizuuchi, K. (1986) Role of DNA topology in Mu transposition : mechanism of sensing the relative orientation of two DNA segments. *Cell* **45** : 793-800.

Cunin, R., Eckhardt, T., Piette, J., Boyen, A., and Glansdorff, N. ( 1983) Molecular basis for modulated regulation of gene expression in the arginine regulon of Escherichia coli K-12. *Nucleic Acids Res* **11** : 5007-5019.

- Cuypers, H.T., van Loon-Klaassen, L.A., Egberts, W.T., de Jong, W.W., and Bloemendal, H. (1982) The primary structure of leucine aminopeptidase from bovine eye lens. *J Biol Chem* **257** : 7077-7085.
- Das, B., Bischerour, J., Val, M.E., and Barre, F.X. (2010) Molecular keys of the tropism of integration of the cholera toxin phage. *Proc. Natl. Acad. Sci. USA* Epub ahead of print.
- De Massy, B., Béjar, S., Louarn, J.M., and Bouché, J.P. (1987). Inhibition of replication forks exiting the terminus region of the E.coli chromosome occurs at two loci separated by 5 min. *Proc. Natl. Acad. Sci. USA* **84** : 1758-1763.
- Dennis, C.A., Glycos, N.M., Parsons, M.M., and Phillips, S.E.V. (2002). The structure of AhrC, the Arginine repressor/activator protein from *Bacillus subtilis*. *Acta Cryst.* (2002). **D58** : 421-430
- Drlica, K. (1990) Bacterial topoisomerases and the control of DNA supercoiling. *Trends Genet* **6** : 433-437.
- Dubarry, N., and Barre, F.X. (2010) Fully efficient chromosome dimer resolution in *Escherichia coli* cells lacking the integral membrane domain of FtsK. *EMBO J* **29** (3) :597-605.
- Dubarry, N., Possoz, C., and Barre, F.X. (2010) Multiple regions along the *Escherichia coli* FtsK protein are implicated in cell division. *Mol Microbiol.* **78** (5) : 1088-1100.
- Ebersbach, G., and Gerdes, K. (2005) Plasmid segregation Mechanism. *Annu. Rev. Genet.* **39** : 453-79
- Eckhardt, T. (1980) Isolation of plasmids carrying the arginine repressor gene argR of *Escherichia coli* K12. *Mol Gen Genet* **178**: 447-452.

Ferreira, H., Butler-Cole, B., Burgin, A., Baker, R., Sherratt, D.J., and Arciszewska, L.K. (2003) Functional analysis of the C-terminal domains of the site-specific recombinases XerC and XerD. *J. Mol. Biol.* **330**: 15-27.

Flinn, H., Burke, M., Stirling, C.J., and Sherratt, D.J. (1989) Use of gene replacement to construct *Escherichia coli* strains carrying mutations in two genes required for stability of multicopy plasmids. *J. Bacteriol* **171**: 2241-2243.

Funnell, B.E. (1991) The P1 plasmid partition complex at *parS*. The influence of *Escherichia coli* integration host factor and of substrate topology. *J Biol Chem* **266** : 14328-14337.

Funell, B.E., and Gagnier, L. (1993) The P1 plasmid partition complex at *parS*. II. Analysis of ParB protein binding activity and specificity. *J Biol Chem* **268**: 3616-3624.

Funell, B.E., and Slavcev, R.A. (2004) Partition system of Bacterial Plasmids. *Plasmid biology ASM Press*, Washington, D.C.

Garnett, J.A., Marincs, F., Baumberg, S., Stockley, P.G., Phillips, S.E. (2008) Structure and function of the arginine repressor-operator complex from *Bacillus subtilis*. *J Mol Biol* **379** (2) : 284-98.

Gellert, M., Mizuuchi, K., O'Dea, M.H., and Nash, H.A. (1976) DNA gyrase : an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci USA* **73** : 3872-3876.

Gerdes, K., Rasmussen, P.B., and Molin, S. (1986) Unique type of plasmid maintenance function: post-segregational killing of plasmid-free cells. *Proc Natl Acad Sci USA* **83** : 3116-3120.

Gerdes, K., Gulyaev, A.P., Franch, T., Pedersen, K., and Mikkelsen, N.D. (1997) Antisense RNA-regulated programmed cell death. *Annu. Rev. Genet.* **31** : 1-31.

Ghosh, S.K., Hajra, S., Paek, A., Jayaram, M. (2006) Mechanism for chromosome and plasmid Segregation. *Annu. Rev. Biochem* **75** : 211-41.

Glansdorff, N. (1996) Biosynthesis of arginine and polyamines. In *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. (Neidhardt, F.C.I., editor in chief.) Washington, D.C. : *American Society for Microbiology*, pp 408-433.

Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W., and Wright, A. (1997) Chromosome and low copy plasmid segregation in *Escherichia coli* : visual evidence for distinct mechanisms. *Cell* **90** : 1113-1121.

Gourlay, S.C., and Colloms, S.D. (2004) Control of Cre recombination by regulatory elements from Xer recombination Systems. *Mol Microbiol* **52** (1) : 53-65.

Grandory, R., Lavoie, T.A., Pflumm, M., Tian, G., Niersbach, H., Maas, W.K., Fairman, R., and Carey, J. (1995) The DNA-binding domain of the hexameric arginine repressor. *J Mol Biol* **254**: 150-162.

Grant, S.G., Jesse, J., Bloom, F.R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *P Natl Acad Sci USA* **87**: 4645-4649.

Graham, J.E., Sivanathan, V., Sherratt, D.J., Arciszewska, L.K. (2009) FtsK translocation on DNA stops at XerCD-dif. *Nucleic Acid Research* **38** (1): 72-81.

Grindley, N.D.F., Whiteson, K.L., Rice, P.A. (2006) Mechanism of site-specific recombination. *Annu. Rev. Biochem* **75** : 567-605.



Hallet, B., and Sherratt, D.J. (1997) Transposition and site-specific recombination : adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol Rev* **21** : 157-178.

Hallet, B., Sherratt, D.J., and Hayes, F. (1997) Pentapeptide scanning mutagenesis : random insertion of a variable five amino acid cassette in a target protein. *Nucleic Acids Res* **25** : 1866-1867.

Hallet, B., Arciszewska, L.K., and Sherratt, D.J. (1999) Reciprocal control of catalysis by the tyrosine recombinases XerC and XerD : an enzymatic switch in site-specific recombination. *Mol Cell* **4** : 949-959.

Harwood, C.R., Baumberg, S. (1977) Arginine Hydroxamate-resistant mutants of *Bacillus subtilis* with altered control of arginine metabolism. *J Gen Microbiol* **100** (1) : 177-88.

Hazan, R., Sat, B., Reches, M., and Engelberg-Kulka. (2001) Postsegregational killing mediated by the P1 phage “addiction module” *phd-doc* requires the *Escherichia coli* programmed cell death system *mazEF*. *J Bacteriol.* **183** : 2046-2050.

Hiraga, S., Jaffe, A., Ogura, T., Mori, H., Takahashi, H. (1986) F plasmid *ccd* mechanism in *Escherichia coli*. *J Bacteriol.* **166** : 100-104.

Hiraga, S., Niki, H., Ogura, T., Ichinose, C., Mori, H., Ezaki, B., and Jaffe, A. (1989) Chromosome partitioning in *Escherichia coli* : novel mutants producing anucleate cells. *J Bacteriol* **171** : 1496-1505.

Hiraga, S. (1993) Chromosome partition in *Escherichia coli*. *Curr Opin Genet Dev* **3** : 789-801.

Hiraga, S. (2000) Dynamic localization of bacterial and plasmid chromosomes. *Annu Rev Genet* **34** : 21-59.

Holmes, V.F., and Cozzarelli, N.R. (2000) Closing the ring : links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci U S A* **97** : 1322-1324.

Horiuchi, T., Fujimura, Y., Nishitani, H., Kobayashi, T., Hidaka, M. (1994) The DNA replication fork blocked at the *ter* site may be an entrance for the RecBCD enzyme into duplex DNA. *J Bacteriol* **176** (15) : 4656-63.

Holtham, C.A.M., Jumel, K., Miller, C.M., Harding, S.E., Baumberg, S., and Stockley, P.G. (1999) Probing Activation of the prokaryotic arginine transcriptional regulator using chimeric proteins. *J.Mol.Biol.* **289** : 707-727

Hu, K.H., Liu, E., Dean, K., Gingras, M., DeGraff, W., Trun, N.J. (1996) Overproduction of three genes leads to camphor resistance and chromosome condensation in *Escherichia coli*. *Genetics* **143** (14) : 1521-32.

Jaffe, A., Ogura, T., and Hiraga, S. (1985) Effects of the *ccd* function of the F plasmid on bacterial growth. *J.Bacteriol.* **163** : 841-849.

Jensen, R.B., and Gerdes, K. (1997) Partitioning of plasmid R1. The ParM protein exhibits ATPase activity and interacts with the centromere-like ParR-parC complex. *J Mol Biol* **269** : 505-513.

Jensen, R.B., Lurz, R., and Gerdes, K., (1998) Mechanism of DNA segregation in prokaryotes : replicon pairing by *parC* of plasmid R1. *Proc Natl Acad SCI USA* **95** : 8550-8555.

Jeon, Y., Lee, Y.S., Han, J.S., Kim, J.B., and Hwang, D.S. (2001) Multimerization of phosphorylated and non-phosphorylated ArcA is necessary for the response regulator function of the Arc two-component signal transduction system. *J Biol Chem* **276** : 40873-40879.

Kennedy, S.P., Chevalier, F., Barre, F.X. (2008) Delayed activation of Xer recombination at dif by FtsK during septum assembly in *Escherichia coli*. *Mol.Microbiol* **68** (4) : 1018-28.

Kobayashi, I. (2004) Genetic Addiction : a principle of gene symbiosis in a genome. *Plasmid Biology* ASM Press, Washington, D.C.

Kupakis, A.K., and Reitzer, L. (2002) ArgR-Independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli*. *J.Bacteriol.* **184** (11) : 2940-2950.

Krogh, B.O., and Shuman, S. (2000a) Catalytic mechanism of DNA topoisomerase IB. *Mol Cell* **5** : 1035-1041.

Krogh, B.O., and Shuman, S. (2000b) DNA strand transfer catalyzed by vaccinia topoisomerase : peroxidolysis and hydroxylaminolysis of the covalent protein-DNA intermediate. *Biochemistry* **39** : 6422-6432.

Kuempel, P.L., Henson, J.M., Dircks, L., Tecklenburg, M., and Lim, D.F. (1991) Dif, a recA-independent recombination site in the terminus region of the chromosome of *Escherichia coli*. *New Biol* **3** : 799-811.

Le Bourgeois, P., Bugarel, M., Campo, N., Daveran-Mingot, M.L., Labonté, J., Lafranchi, D., Lautier, T., Pagès, C., and Ritzenhaler, P. (2007) Unconventional Xer recombination machinery of *Streptococci/Lactococci*. *PLOS Genet* **3** (7) : 1190-1199.

Levy, O., Ptacin, J.L., Pease, P.J., Gore, J., Eisen, M.B., Bustamante, C., Cozzarelli, N.R. (2005) Identification of nucleotide sequences that direct the movement of the *Escherichia coli* FtsK translocase. *Proc Natl Acad Sci USA* **102** (49) : 17618-17623.

Li, Y., Stewart, N.K., Berger, A.J., Vos, S., Schoeffler, A.J., Berger, J.M., Chait, B.T., and Oakley, M.G. (2010) *Escherichia coli* condensin MukB stimulates topoisomerase IV activity by a direct physical interaction. *Proc Natl Acad Sci USA* **107** (44) : 18832-18837.

Lim, D.B., Oppenheim, J.D., Eckhardt, T., and Maas, W.K. (1987) Nucleotide sequence of the *argR* gene of *Escherichia coli* K-12 and isolation of its product, the arginine repressor. *Proc Natl Acad Sci USA* **84** : 6697-6701.

Liu, G., Draper, G.C., and Donachie, W.D. (1998). FtsK is a bifunctional protein involved in cell division and chromosome localization in *Escherichia coli*. *Mol Microbiol* **29** : 893-903.

Lovett, S.T., and Kolodner, R.D. (1991) Nucleotide sequence of the *Escherichia coli* *recJ* chromosomal region and construction of *recJ*-overexpression plasmids. *J Bacteriol* **173** : 353-364.

Löwe, J., Ellonen, A., Allen, M.D., Atkinson, C., Sherratt, D.J., and Grainge, I. (2008). Molecular Mechanism of sequence-directed DNA loading and translocation by FtsK. *Mol Cell* **31** (4) : 498-509.

Luchi, S., and Lin, E.C. (1988) *arcA* (*dye*), a global regulatory gene in *Escherichia coli* mediating repression of enzymes in aerobic pathways. *Proc Natl Acad Sci USA* **85** : 1888-1892.

Luchi, S., and Lin, E.C. (1992) Mutational analysis of signal transduction by ArcB, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathways in *Escherichia coli*. *J Bacteriol* **174** : 3972-3980.

Luchi, S., and Lin, E.C. (1993) Adaptation of *Escherichia coli* to redox environments By gene expression. *Mol Microbiol* **9** : 9-15.

Lynch, A.S., and Lin, E.C. (1996) Transcriptional control mediated by the ArcA two-component response regulator protein of *Escherichia coli* : characterization of DNA binding at target promoters. *J Bacteriol* **178** (21) : 6238-6249.

Maas, W.K. (1994) The arginine repressor of *Escherichia coli*. *Microbiol Rev* **58** : 631-640.

Masuda, Y., Miyakawa, K., Nishimura, Y., and Ohtsubo, E. (1993) chpA and chpB, *Escherichia coli* chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. *J Bacteriol*. **175** : 6850-6856.

Mc Culloch, R., Coggins, L.W., Colloms, S., and Sherratt, D.J. (1994) Xer mediated site-specific recombination at *cer* generates Holliday junctions *in vivo*. *EMBO J* **13** (8) : 1844-1855.

MacDonald, Alasdair, I., Lu, Y., Kilbride, E.A., Akopian, A., and Colloms, S.D. (2008) PepA and ArgR do not regulate Cre recombination at the bacteriophage P1 loxP site. *Plasmids* **59** (2) : 119-129.

McLeod, S.M., and Waldor, M.K. (2004) Characterization of XerC and XerD dependent CTX phage integration in *Vibrio cholerae*. *Mol Microbiol* **54** (4) : 935-347.

Massey, T.H., Aussel, L., Barre, F.X., Sherratt, D.J. (2004) Asymmetric activation of Xer site-specific recombination by FtsK. *EMBO Rep* **5** (4) : 399-404.

Massey, T.H., Mercogliano, C.P., Yast, J., Sherrat, D.J., Löwe, J. (2006) Double-Stranded DNA translocation: structure and mechanism of hexameric FtsK. *Mol Cell* **23** (4) : 457-469.

Migliarese, M.R., Richardson, A.F., Aziz, N., and Bender, T.P. (1996) Differential regulation of c-Myb-induced transcription activation by a phosphorylation site in the negative regulatory domain. *J Biol Chem* **271** (37): 22697-22705.

Minh, P.N., Devroede, N., Massant, J., Maes, D., Charlier, D. (2009) Insight into the architecture and stoichiometry of Escherichia coli PepA\*DNA complexes involved in transcriptional control and site-specific DNA recombination by atomic force microscopy. *Nucleic Acids Res* **37** (5) : 1463-76.

Mori, H., Morin, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A ;, and Hiraga, S. (1989) Purification and characterization of SopA and SopB proteins essential for F plamid Partitioning. *J Biol Chem* **264** : 15535-15541.

Morin, A., Huysveld, N., Braun, F., Dimova, D., Sakanyan, V., Charlier, D. (2003) Hyperthermophilic Thermototoga arginine repressor binding to full-length cognate and heterologous arginine operators and to half-site targets. *J Mol Biol* **332** (3) : 537-53.

Nielsen, A.K., Thorsted, P., Thisted, T., Wagner, E.G., and Gerdes, K. (1991) The rifampicin-inducible genes *srnB* from F and *pnd* from R483 are regulated by antisense RNAs and mediate plasmid maintenance by killing of plasmid-free segregants. *Mol Microbiol.* **5** : 1961-1973.

Neilson, L., Blakely, G., and Sherrat, D.J. (1999) Site-specific recombination at dif by Hemophilus influenzae XerC. *Mol Microbiol* **31** : 915-926.

Ni, J., Sakanyan, V., Charlier, D., Glandorff, N., and Van Duyne, G.D. (1999) Structure of the arginine repressor from *Bacillus stearothermophilus*. *Nature Struct Biology* **6** : 427-432.

Niki, H., Imamura, R., Kitaoka, M., Yamanaka, K., Ogura, T., and Hiraga, S. (1992) E.coli MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO J* **11** : 5101-5109.

North, A.K., Smith, M.C., and Baumberg, S. (1989) Nucleotide sequence of a *Bacillus subtilis* arginine regulatory gene and homology of its product to the *Escherichia coli* arginine repressor. *Gene* **80** : 29-38.

Novilos, S., Pages, C., Rousseau, P., Le Bourgeois, P., and Cornet, F. (2010) Are two better than one? Analysis of an FtsK/Xer recombination system that uses a single recombinase. *Nucleic Acids Res* **38** (19) : 6477-6489.

Ogura, T., and Hiraga, S. (1983) Partition mechanism of F plasmid : two plasmid-gene encoded products and a cis-acting region are involved in partition. *Cell* **32** : 351-360.

Patient, M.E., and Summers, D.K. (1993) ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. *Mol Microbiol* **9** : 1089-1095.

Paul, S., and Summers, D. (2004) ArgR and PepA, accessory proteins for XerC-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. *Plasmid* **52** (1) : 63-8.

Pease, P.J., Levy, O., Cost, G.J., Gore, J., Ptacin, J.L., Sherratt, D., Bustamante, C., Cozzarelli, N.R. (2005) Sequence-directed DNA translocation by purified FtsK. *Science* **307** (5709) : 586-590.

Perals, K., Cornet, F., Merlet, Y., Delon, I., and Louarn, J.M. (2000) Functional polarization of the *Escherichia coli* chromosome terminus : the *dif* site acts in chromosome dimer resolution only when located between long stretches of opposite polarity. *Mol Microbiol* **36** : 33-43.

Perals, K., Capiiaux, H., Vincourt, J.B., Louarn, J.M., Sherratt, D.J., Cornet, F. (2001) Interplay between recombination, cell division and chromosome structure during chromosome dimer resolution in *Escherichia coli*. *Mol Microbiol* Feb (39) **4** : 904-913

Recchia, G.D., Aroyo, M., Wolf, D., Blakely, G., and Sherratt, D.J. (1999) FtsK-dependent and-independent pathways of Xer site-specific recombination. *EMBO J* **18** : 5724-5734.

Recchia, G.D., and Sherratt, D.J. (1999) Conservation of *xer* site-specific recombination genes in bacteria. *Mol Microbiol* **34** : 1146-1148.

Redinbo, M.R., Champoux, J.J., and Hol, W.G. (2000) Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA. *Biochemistry* **39** : 6832-6840.

Réhel, N., and Szatmari, G. (1996) Characterization of the stable maintenance of the *Shigella flexneri* plasmid pHS-2. *Plasmid* **36** (3) : 183-90.

Reijns, M., Lu, Y., Leach, S., Colloms, S.D. (2005) Mutagenesis of PepA suggests a new model for the Xer/cer synaptic complex. *Mol Microbiol* **57** (4) : 927-941.

Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D.J. (2008) Independent positioning and action of *Escherichia coli* replisomes in live cells. *Cell* **133** (1) : 90-102.



- Roovers, M., Charlier, D., Feller, A., Gigot, D., Holemans, F., Lissens, W., Pierard, A., and Glansdorff, N. (1988) *carp*, a novel gene regulating the transcription of the carbamoylphosphate synthetase operon of *Escherichia coli*. *J Mol Biol* **204** : 857-865.
- Sawers, G., and Suppmann, B. (1992) Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. *J Bacteriol* **174** (11) : 3474-3478.
- Sawitzke, J.A., and Austin, S. (2000) Suppression of chromosome segregation defects of *Escherichia coli* *muk* mutants by mutations in topoisomerase I. *Proc Natl Acad Sci USA* **97** :1671-1676.
- Scher, W.I., Vogel, H.J. (1957) Occurrence of ornithine delta-transaminase : a dichotomy. *Proc Natl Acad Sci USA*, **43** (9) : 796-803.
- Schneider, B.L., Kuipakis, A., and Reitzer, L.J. (1998) Arginine catabolism and arginine succinyltransferase pathway in *Escherichia coli*. *J. Bacteriol.* **180** (16) : 4278-4286.
- Sekar, R.B., and Periasamy, A. (2003) Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *J. Cell. Biol* **160** (5) : 629-633.
- Sénéchal, H., Delesques, J., Szatmari, G. (2010) *Escherichia coli* ArgR mutants defective in *cer*/Xer recombination, but not in DNA binding. *FEMS Microbiol Lett*, **305** : 162-169.
- Sherratt, D.J., Arciszewska, L.K., Blakely, G., Colloms, S., Grant, K., Leslie, N., and McCulloch, R. (1995) Site-specific recombination and circular chromosome segregation. *Philos Trans R Soc Lond B Biol Sci* **347** : 37-42.

Sirois, S., and Szatmari, G. (1995) Detection of XerC and XerD recombinases in Gram-negative bacteria of family *Enterobacteriaceae*. *Journal of bacteriology* **177** (14) : 4183-4186.

Smith, M.C., Czaplewski, L., North, A.K., Baumberg, S., and Stockley, P.G. (1989) Sequences required for regulation of arginine biosynthesis promoters are conserved between *Bacillus subtilis* and *Escherichia coli*. *Mol Microbiol* **3** : 23-28.

Spiers, A.J., and Sherratt, D.J. (1999) C-terminal interactions between the XerC and XerD site-specific recombinases. *Mol Microbiol* **32** : 1031-1042.

Steiner, W., Liu, G., Donachie, W.D., and Kuempel, P. (1999) The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. *Mol Microbiol* **31** : 579-583.

Stirling, C.J., Colloms, S.D., Collins, J.F., Szatmari, G., and Sherratt, D.J. (1989) *xerB*, an *Escherichia coli* gene required for plasmid ColE1 site-specific recombination, is identical to pepA, encoding aminopeptidase A, a protein with substantial similarity to bovine lens leucine aminopeptidase. *EMBO J* **8** : 1623-1627.

Stirling, C.J., Stewart, G., and Sherratt, D.J. (1988a) Multicopy plasmid stability in *Escherichia coli* requires host-encoded functions that lead to plasmid site-specific recombination. *Mol. Gen. Genet.* **214** : 80-84.

Stirling, C.J., Szatmari, G., Stewart, G., Smith, M.C., and Sherratt, D.J. (1988b) The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the ColE1 *cer* locus. *EMBO J* **7**: 4389-4395.

Stockley, P.G., Baron, A.J., Wild, C.M., Parsons, I.D., Miller, C.M., Holtham, C.A.M., and Baumberg, S. (1998) Dissecting the molecular details of prokaryotic transcriptional

control by surface plasmon resonance: the methionine and arginine repressor proteins. *Biosensors & bioelectronics* **13**: 637-650.

Sträter, N., Sherratt, D.J., and Colloms, S.D. (1999) X-ray structure of aminopeptidase A from *Escherichia coli* and a model for the nucleoprotein complex in Xer site-specific recombination. *EMBO J* **18**: 4513-4522.

Strawn, R., Melichercik, M., Green, M., Stockner, T., Carey, J., and Ettrich, R. (2010) Symmetric allosteric mechanism of *Escherichia coli* arginine repressor exploits competition between L-arginine ligands and resident arginine residues. *PLOS comput. Biol.* **6** (6) : e1000801.

Subramanya, H.S., Arciszewska, L.K., Baker, R.A., Bird, L.E., Sherratt, D.J., Wigley, D.B. (1997) Crystal structure of the site-specific recombinase, XerD. *EMBO J* **16** : 5178-5187.

Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein production and purification* **41** (1): 207-234.

Summers, D.K., and Sherratt, D.J. (1984) Multimerization of high copy number plasmids causes instability : ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell* **36** : 1097-1103.

Summers, D.K., and Sherratt, D.J. (1988) Resolution of ColE1 dimers requires a DNA sequence implicated in the three-dimensional organization of the *cer* site. *EMBO J* **7** : 851-858.

Sunnerhagen, M., Nilges, M., Otting, G., and Carey, J. (1997) Solution structure of the DNA-binding domain and model for the complex of multifunctional hexameric arginine repressor with DNA. *Nat Struct Biol* **4** : 819-826.

Tardat, B., and Touati, D. (1993) Iron and oxygen regulation of *Escherichia coli* MnSOD expression : competition between the global regulators Fur and ArcA for binding to DNA.. *Mol Microbiol* **9** : 53-63.

Tecklenburg, M., Naumer, A., Nagappan, O., Kuempel, P. (1995) The dif resolvase locus of the *Escherichia coli* chromosome can be replaced by a 33-bp sequence, but function depends on location. *Proc Natl Acad Sci USA* **92** (5) : 1352-1356.

Tian, G., Lim, D., Carey, J., and Maas, W.K. (1992) Binding of the arginine repressor of *Escherichia coli* K12 to its operator sites. *J Mol Biol* **226** : 387-397.

Tian, G., and Maas, W.K. (1994) Mutational analysis of the arginine repressor of *Escherichia coli*. *Mol Microbiol* **13** : 599-608.

Toro-Roman, A., Mack, T.R., and Stock, A.M., (2005) Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the  $\alpha 4$ - $\beta 5$ - $\alpha 5$  face. *J. Mol. Biol* **349** : 11-26.

Trigueros, S., Tran, T., Sorto, N., Newmark, J., Colloms, S.D., Sherratt, D.J. and Tolmasky, M.E. (2009) *mwr* Xer site-specific recombination is hypersensitive to DNA supercoiling. *Nucleic Acids Res* **37** (11) : 3580-3587.

Val, M.E., Bouvier, M., Campos, J., Sherratt, D., Cornett, F., Mazel, D., Barre, F.X. (2005) The single stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol Cell* **19** (4) : 559-566.

Van Duyne, G.D., Ghosh, G., Maas, W.K., and Sigler, P.B. (1996) Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J Mol Biol* **256** : 377-391.

- Vazquez, M., Colloms, S.D., Summers, D.W. (2005) Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three-dimensional topological pathway. *J. Mol. Biol.* **346** : 493–504.
- Villion, M., and Szatmari, G. (1998) Cloning and characterisation of the *Proteus mirabilis* *xerD* gene. *FEMS Microbiol Lett.* **164** : 83-90.
- Villion, M., and Szatmari, G. (2003) The XerC recombinase of *Proteus mirabilis* : characterization and interaction with other tyrosine recombinase. *FEMS Microbiol Lett.* **226** : 65-71.
- Vogt, A.M. (1970) Purification and properties of an aminopeptidase from *Escherichia coli*. *Journal of biological chemistry* **245**: 4760-4769.
- Wang, H., Glansdorff, N., and Charlier, D. (1998) The arginine repressor of *Escherichia coli* K-12 makes direct contacts to minor and major groove determinants of the operators. *J Mol Biol* **277** : 805-824.
- Webb, C.D., Graumann, P.L., Kahana, J.A., Teleman, A.A., Silver, P.A., and Losick, R. (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* **28** : 883-892.
- Yates, J., Zhekov, I., Baker, R., Eklund, B., Sherratt, D.J., Arciszewska, L.K. (2006). Dissection of a functional interaction between the DNA translocase, FtsK, and the XerD recombinase. *Mol Microbiol* **59** (6) : 1754-66.
- Yengo, C.M., and Berger, C.L. (2010) Fluorescence anisotropy and resonance energy transfer: powerful tools for measuring real time protein dynamics in a physiological environment. *Curr Opin Pharmacol* **10** (6) : 731-7

Yu, X.C., Weihe, E.K., and Margolin, W. (1998) Role of the C terminus of FtsK in *Escherichia coli* chromosome segregation. *J.Bacteriol.* **180** : 6424-6428.

Zakova, N., and Szatmari, G. (1995) Site specific recombination between ColE1 *cer* and NTP16 *nmr* sites in vivo. *Mol Gen Genet* **247** : 509-514.

Zechiedrich, E.L., and Cozzarelli, N.R. (1995) Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in *Escherichia coli*. *Genes Dev.* Nov **15** ; 9 (22) :2859-2869.

## REMERCIEMENTS

Je tiens tout d'abord à remercier tout particulièrement le Docteur George SZATMARI qui m'a accueilli dans son laboratoire et donné la chance de vivre cette longue et passionnante aventure que fut mon doctorat à l'Université de Montréal. Grâce à vous, j'ai découvert de nouveaux horizons, je vous en serai à jamais reconnaissant. Je vous souhaite santé et bonheur. Je tiens également à vous remercier pour votre patience et vos précieux conseils qui m'ont permis d'avancer dans la bonne voie, tout au long de ces années, tant sur le plan personnel que sur le plan professionnel et expérimental.

Je souhaite également remercier le Docteur Marc DROLET, qui lui aussi m'a permis de venir au Québec et tenter ma chance. Je vous remercie aussi pour vos conseils avisés sur certaines des expériences que j'ai menées et sur tant d'autres sujets.

J'adresse mes plus sincères remerciements aux autres membres du jury, à savoir le Docteur Patrick HALLENBECK, Madame le Docteur Chantal AUTEXIER et le Docteur Dindial RAMOTAR le représentant du doyen, qui ont bien voulu me consacrer un peu de leur temps pour évaluer cette thèse.

J'adresse mes plus sincères remerciements à de nombreux membres de l'équipe des Docteurs de ce département, dont les Docteurs France DAIGLE, Guy LEMAY, Pierre BELHUMEUR, Jacques THIBODEAU, qui ont su être à l'écoute de mes questions et me prodiguer leurs conseils avisés.

Je tiens aussi à remercier nombre de personnes membres du secrétariat dont Lucy LEVESQUES, Lise BABINEAU, Johanne MERCIER, Karine LANGLOIS, Richard ALIX, Pauline PHOENIX, Daniel MORENCY, Lamercie YOUANCE, Céline MILLETTE, pour m'avoir secouru dans le dédale administratif, qui ont toujours été à mon écoute, toujours disponibles, et bien plus encore. Je tiens également à remercier les membres du secrétariat qui sont partis en retraite dont Mesdames Carole ARSENAULT,

Lise BABINEAU, Nicole FALLECKER, et Lise DAVID qui ont remué ciel et terre pour me renseigner et faciliter mon existence.

Je tiens également à remercier les membres (passés et présents) du laboratoire du Docteur George SZATMARI; Hua LIU, dont le sérieux et la rigueur furent de véritables exemples pour moi. Hélène SÉNÉCHAL et Manuela VILLION, qui m'ont mis le pied à l'étrier lors de mes débuts au Canada et Fuli JIA, qui m'a prodigué d'excellents conseils.

Parmi les stagiaires que j'aie côtoyés, j'adresserais plus particulièrement une pensée pour Chantal FOREST, connue lors d'un stage d'initiation, dont j'aie vu le potentiel et que j'ai encouragée dans la voie scientifique. C'est un véritable plaisir de la voir désormais en doctorat chez le docteur France DAIGLE. Les remerciements qu'elle m'a adressés en arrivant au doctorat m'ont mis du baume au cœur.

Parmi les membres des autres équipes, je tiens tout particulièrement à remercier mademoiselle Flora NOLENT du laboratoire du docteur Marc DROLET, qui m'a toujours accueilli avec le sourire et qui surtout, a réussi à comprendre mon sens de l'humour farfelu et hors-normes, tout en étant à l'écoute de mes dires. Je tiens également à remercier Valentine USONGO également membre du laboratoire du docteur Marc DROLET, qui a été très chaleureux et ouvert d'esprit.

Je tiens également à remercier parmi l'équipe du Docteur Jacques THIBODEAU, mesdemoiselles Laetitia GENÈVE et Marie-Claude BOURGEOIS pour leur gentillesse et leur patience. Je tiens à remercier aussi Mahamad PEZESKI membre du laboratoire du Docteur Jacques THIBODEAU, pour sa sympathie.

Et enfin, pour terminer, je tiens à remercier mes parents qui m'ont toujours soutenu lors de ces années de doctorat à 6000 Kms de France, et remonté le moral lors des moments de galère qui ont parsemé ma route. Merci d'avoir encouragé votre « zigoto » de fils, et d'avoir toujours cru en mes capacités. Une pensée émue pour ma grand-mère qui malheureusement nous a quittés.